

شبکه های کامپیوتری

نویسنده

اندرواس.تنن باوم

فهرست مطالب

۱۱	مقدمه	۱۱
۱-۱	کاربردهای شبکه های کامپیوتری	۱۲
۱-۱-۱	کاربردهای تجاری	۱۲
۲-۱-۱	کاربردهای خانگی	۱۵
۳-۱-۱	کاربران سیار	۱۸
۴-۱-۱	تباعات اجتماعی	۲۰
۲-۱	سخت افزار شبکه	۲۲
۱-۲-۱	شبکه های محلی	۲۳
۲-۲-۱	شبکه های شهری	۲۵
۳-۲-۱	شبکه های گسترده	۲۶
۴-۲-۱	شبکه های بیسیم	۲۷
۵-۲-۱	شبکه های خانگی	۲۹
۶-۲-۱	شبکه های	۳۱
۳-۱	نرم افزار شبکه	۳۲
۱-۳-۱	سلسله مراتب پروتکل ها	۳۲
۲-۳-۱	ملاحظات در طراحی لایه ها	۳۶
۳-۳-۱	سرویسهای اتصال-گرا و غیرمتصل	۳۷
۴-۳-۱	عملکردهای پایه سرویس	۳۹
۵-۳-۱	رابطه سرویس و پروتکل	۴۱
۴-۱	مدلهای مرجع	۴۲
۱-۴-۱	مدل مرجع OSI	۴۲
۲-۴-۱	مدل مرجع TCP/IP	۴۵
۳-۴-۱	مقایسه مدل های OSI و TCP/IP	۴۷
۴-۴-۱	نگاهی انتقادی به مدل OSI و پروتکل های آن	۴۹
۵-۴-۱	نگاهی انتقادی به مدل TCP/IP	۵۱
۵-۱	شبکه های نمونه	۵۲
۱-۵-۱	ایترنت	۵۲
۲-۵-۱	شبکه های اتصال-گرا: X.25, Frame Relay	۵۲
۶-۰	ATM	۶۰
۳-۵-۱	ایترنت	۶۵
۴-۵-۱	شبکه های محلی بیسیم: 802.11	۶۷
۶-۱	استانداردهای شبکه	۷۰
۱-۶-۱	مراجع مسئول استانداردهای مخابرات	۷۰
۲-۶-۱	مراجع مسئول استانداردهای بین المللی	۷۲
۳-۶-۱	مراجع مسئول استانداردهای اینترنت	۷۴
۷-۱	واحدهای اندازه گیری	۷۵
۸-۱	طرح کلی مباحث کتاب	۷۶
۹-۱	خلاصه	۷۸
۷۹	مسائل	۷۹
۲	لایه فیزیکی	۸۳
۱-۲	مبانی نظری مخابرات داده	۸۳
۱-۱-۲	آنالیز فوریه	۸۳
۲-۱-۲	محدودیت پهنای باند	۸۴
۳-۱-۲	حداکثر نرخ داده در یک کانال	۸۶
۲-۲	رسانه انتقال هدایت پذیر	۸۷
۱-۲-۲	رسانه مغناطیسی	۸۷
۲-۲-۲	زوج تابیده	۸۸
۳-۲-۲	کابل کواکسیال	۸۹
۴-۲-۲	فیبر نوری	۹۰
۳-۲	انتقال بیسیم	۹۶
۱-۳-۲	طیف الکترومغناطیس	۹۶
۲-۳-۲	مخابرات رادیویی	۹۹
۳-۳-۲	مخابرات مایکروویو	۱۰۰
۴-۳-۲	امواج مادون قرمز و میلیمتری	۱۰۲
۵-۳-۲	مخابرات امواج نوری	۱۰۳
۴-۲	ماهواره های مخابراتی	۱۰۳
۱-۴-۲	ماهواره های زمین ثابت	۱۰۵
۲-۴-۲	ماهواره های مدار متوسط	۱۰۸

فهرست مطالب	
۱۹۶	۴-۳ پروتکل های پنجره لغزنده
۱۹۸	۱-۴-۳ پروتکل پنجره لغزنده ای بی تی
۲۰۱	۲-۴-۳ پروتکل «N تا به عقب برگرد»
۲۰۸	۳-۴-۳ پروتکل تکرار انتخابی
۲۱۴	۵-۳ ارزیابی پروتکل ها
۲۱۴	۱-۵-۳ مدل ماشین حالت محدود
۲۱۷	۲-۵-۳ مدل شبکه پتری
۲۱۹	۶-۳ چند نمونه از پروتکل های لینک داده
۲۱۹	۱-۶-۳ HDLC - کنترل سطح بالای لینک داده
۲۲۲	۲-۶-۳ لایه پیوند داده در اینترنت
۲۲۷	۷-۳ خلاصه
۲۲۷	مسائل
۴ زیر لایه دسترسی به لایه انتقال .. ۲۳۱	
۲۳۲	۱-۴ مسئله تخصیص کانال
۲۳۵	۲-۴ پروتکل های دسترسی چندگانه
۲۳۵	۱-۲-۴ ALOHA
۲۳۹	۲-۲-۴ پروتکل های دسترسی چندگانه با قابلیت شنود سیگنال حامل (CSMA)
۲۴۲	۳-۲-۴ پروتکل های بدون تصادم
۲۴۵	۴-۲-۴ پروتکل های با رقابت محدود
۲۴۹	۵-۲-۴ پروتکل های دسترسی چندگانه مبتنی بر تقسیم طول موج
۲۵۲	۶-۲-۴ پروتکل های بی سیم برای شبکه محلی
۲۵۵	۳-۴ اترنت
۲۵۶	۱-۳-۴ کابل کشی اترنت
۲۵۹	۲-۳-۴ کدینگ منجستر
۲۶۰	۳-۳-۴ پروتکل زیر لایه MAC در اترنت
۲۶۲	۴-۳-۴ الگوریتم عقب گرد نمایی
۲۶۳	۵-۳-۴ کارائی (بازده) اترنت
۲۶۵	۶-۳-۴ اترنت مبتنی بر سوئیچ
۲۶۷	۷-۳-۴ اترنت سریع
۲۷۱	۸-۳-۴ اترنت گیگابیت
۲۷۵	۹-۳-۴ IEEE 802.2: کنترل منطقی لینک
۲۷۶	۱۰-۳-۴ نگاهی به گذشته اترنت
۲۷۶	۴-۴ شبکه های محلی بی سیم
۲۷۷	۱-۴-۴ پشته پروتکلی 802.11
۱۰۸	۳-۴-۲ ماهواره های مدار پائین
۱۱۱	۴-۴-۲ ماهواره یا فیبر؟
۱۱۲	۵-۲ شبکه تلفن عمومی
۱۱۲	۱-۵-۲ ساختار سیستم تلفن
۱۱۵	۲-۵-۲ تلفن و سیاست
۱۱۷	۳-۵-۲ مدارهای پایانی: مودم، ADSL، و بیسیم
۱۲۹	۴-۵-۲ ترانک ها و مالتی پلکس کردن
۱۳۷	۵-۵-۲ سوئیچینگ
۱۴۲	۶-۲ شبکه تلفن همراه
۱۴۳	۱-۶-۲ تلفن های همراه نسل اول: صدای آنالوگ
۱۴۳	۲-۶-۲ تلفن های همراه نسل دوم: صدای دیجیتال
۱۴۷	۳-۶-۲ تلفن های همراه نسل سوم: صدای دیجیتال و داده
۱۵۵	۷-۲ تلویزیون کابلی
۱۵۸	۱-۷-۲ تلویزیون با آنتن مرکزی
۱۵۸	۲-۷-۲ اینترنت کابلی
۱۶۰	۳-۷-۲ تخصیص طیف فرکانسی
۱۶۱	۴-۷-۲ مودم های کابلی
۱۶۳	۵-۷-۲ مودم کابلی یا ADSL؟
۱۶۴	۸-۲ خلاصه
۱۶۵	مسائل
۳ لایه پیوند داده .. ۱۷۱	
۱۷۱	۱-۳ ملاحظات طراحی لایه پیوند داده
۱۷۱	۱-۱-۳ سرویسهایی که به لایه شبکه داده می شود
۱۷۲	۲-۱-۳ فریم بندی
۱۷۴	۳-۱-۳ کنترل خطا
۱۷۷	۴-۱-۳ کنترل جریان
۱۷۸	۲-۳ کشف و تصحیح خطا
۱۷۹	۱-۲-۳ گدهای تصحیح خطا
۱۸۱	۲-۲-۳ گدهای کشف خطا
۱۸۵	۳-۳ چند پروتکل ساده لینک داده
۱۸۹	۱-۳-۳ پروتکل یکطرفه نامقید
۱۹۱	۲-۳-۳ پروتکل توقف-انتظار یکطرفه
۱۹۳	۳-۳-۳ پروتکل یکطرفه برای کانال های نوبت دار

۳۳۴	دیتاگرام	۲۷۸	۲-۴-۴ لایه فیزیکی در 802.11
۳۳۶	۲-۵ الگوریتمهای مسیریابی	۲۸۰	۳-۴-۴ پروتکل زیر لایه MAC در 802.11
۳۳۸	۱-۲-۵ اصل بهینگی	۲۸۴	۴-۴-۴ ساختار فریم 802.11
۳۳۹	۲-۲-۵ مسیریابی مبتنی بر کوتاهترین مسیر	۲۸۶	۵-۴-۴ خدمات
۳۴۱	۳-۲-۵ الگوریتم سیل آسا (Flooding)	۲۸۷	۵-۴ بی سیم با باند گسترده
۳۴۳	۴-۲-۵ مسیریابی بردار فاصله	۲۸۸	۱-۵-۴ مقایسه 802.11 با 802.16
۳۴۷	۵-۲-۵ مسیریابی حالت لینک	۲۸۹	۲-۵-۴ پشته پروتکلی 802.16
۳۵۳	۶-۲-۵ مسیریابی سلسله‌مراتبی	۲۹۰	۳-۵-۴ لایه فیزیکی در 802.16
۳۵۵	۷-۲-۵ مسیریابی فراگیر (Broadcast Routing)	۲۹۲	۴-۵-۴ پروتکل زیر لایه MAC در 802.16
۳۵۷	۸-۲-۵ مسیریابی چندپخش	۲۹۴	۵-۵-۴ ساختار فریم در 802.16
۳۵۹	۹-۲-۵ مسیریابی برای ماشینهای متحرک	۲۹۵	۶-۴ بلوتوث (Bluetooth)
۳۶۳	۱۰-۲-۵ مسیریابی در شبکه‌های ویژه	۲۹۶	۱-۶-۴ معماری بلوتوث
۳۶۸	۱۱-۲-۵ جستجوی گره در شبکه‌های هم‌تابه	۲۹۷	۲-۶-۴ کاربردهای بلوتوث
۳۶۸	همتابه	۲۹۸	۳-۶-۴ پشته پروتکلی بلوتوث
۳۷۳	۳-۵ الگوریتمهای کنترل ازدحام	۳۰۰	۴-۶-۴ لایه رادیویی در بلوتوث
۳۷۵	۱-۳-۵ اصول کلی در کنترل جریان	۳۰۰	۵-۶-۴ لایه باند پایه در بلوتوث
۳۷۷	۲-۳-۵ سیاستهای پیشگیری از ازدحام	۳۰۱	۶-۶-۴ لایه L2CAP در بلوتوث
۳۷۷	۳-۳-۵ کنترل ازدحام در زیر شبکه‌های مدار	۳۰۱	۷-۶-۴ ساختار فریم در بلوتوث
۳۷۸	مجازی	۳۰۳	۷-۴ هدایت در سطح لایه پیوند داده‌ها
۳۸۰	۴-۳-۵ کنترل ازدحام در زیر شبکه‌های دیتاگرام	۳۰۵	۱-۷-۴ پلهانی از 802.x به 802.y
۳۸۳	۵-۳-۵ دور ریختن بار (Load Shedding)	۳۰۷	۲-۷-۴ بهم‌بندی شبکه‌ها به صورت محلی
۳۸۵	۶-۳-۵ کنترل لرزش (Jitter Control)	۳۰۹	۳-۷-۴ پلهای مبتنی بر درخت پوشا
۳۸۶	۴-۵ کیفیت خدمات (Quality of Service)	۳۱۰	۴-۷-۴ پلهای راه دور (Remote Bridges)
۳۸۶	۱-۴-۵ نیازها	۳۱۰	۵-۷-۴ تکرارکننده، هاب، پل، سوئیچ، مسیریاب و
۳۸۷	۲-۴-۵ راهکارهای دستیابی به کیفیت خوب	۳۱۱	دروازه
۳۸۷	خدمات	۳۱۴	۶-۷-۴ شبکه‌های محلی مجازی (Virtual LANs)
۴۰۰	۳-۴-۵ خدمات مجتمع (Integrated Services)	۳۲۲	۸-۴ خلاصه
۴۰۲	۴-۴-۵ خدمات متمایز	۳۲۴	مسائل
۴۰۵	۵-۴-۵ سوئیچ برچسب و MPLS		
۴۰۹	۵-۵ بهم‌بندی شبکه‌ها (Internetworking)	۳۲۹	۵ لایه شبکه
۴۱۰	۱-۵-۵ شبکه‌ها از چه دیدگاهی متفاوتند؟	۳۲۹	۱-۵ مسائل طراحی لایه شبکه
۴۱۱	۲-۵-۵ چگونگی اتصال شبکه‌ها به یکدیگر	۳۲۹	۱-۵ هدایت (سوئیچینگ) بسته به روش ذخیره و
۴۱۳	۳-۵-۵ مدارات مجازی الحاق شده	۳۲۹	هدایت
۴۱۴	۴-۵-۵ بهم‌بندی شبکه‌های بدون اتصال	۳۳۰	۲-۱-۵ خدمات ارائه شده برای لایه انتقال
۴۱۶	۵-۵-۵ ایجاد تونل (Tunneling)	۳۳۱	۳-۱-۵ پیاده‌سازی خدمات بی‌اتصال
۴۱۷	۶-۵-۵ مسیریابی بین شبکه‌های بهم‌متصل	۳۳۳	۴-۱-۵ پیاده‌سازی خدمات اتصال گرا
۴۱۸	۷-۵-۵ قطعه‌قطعه‌سازی بسته‌ها	۵-۱-۵	مقایسه زیر شبکه‌های مدار مجازی و

۵۲۶	۳-۴-۶ پروتکل انتقال بی درنگ	۴۲۲	۶-۵ لایه شبکه در اینترنت
۵۳۰	۵-۶ پروتکل‌های لایه انتقال در اینترنت: TCP	۴۲۴	۱-۶ پروتکل IP
۵۳۰	۱-۵-۶ مقدمه‌ای بر TCP	۴۲۸	۲-۶-۵ آدرسهای IP
۵۳۱	۲-۵-۶ مدل خدمات TCP	۴۴۱	۳-۶-۵ پروتکل‌های کنترل اینترنت
۵۳۳	۳-۵-۶ پروتکل TCP	۴-۶-۵ OSPF: پروتکل مسیریابی برای دروازه‌های	
۵۳۴	۴-۵-۶ سرآیند قطعه TCP	درونی	
۵۳۸	۵-۵-۶ برقراری اتصال TCP	۵-۶-۵ BGP: پروتکل مسیریابی برای دروازه	
۵۳۹	۶-۵-۶ خاتمه دادن به اتصال TCP	خارجی	
۵۴۰	۷-۵-۶ مدل‌سازی فرآیند مدیریت اتصال در TCP	۶-۶-۵ ارسال چندپخشی در اینترنت	
۵۴۲	۸-۵-۶ سیاستهای انتقال در TCP	۷-۶-۵ IP متحرک (Mobile IP)	
۵۴۶	۹-۵-۶ کنترل ازدحام در TCP	۸-۶-۵ IPv6	
۵۴۹	۱۰-۵-۶ مدیریت تایمرها در TCP	۷-۵ خلاصه	
۵۵۲	۱۱-۵-۶ TCP و UDP بی سیم	مسائل	
۵۵۴	۱۲-۵-۶ TCP تراکنشی (Transactional TCP)		
۵۵۶	۶-۶ مسائل مرتبط با کارایی	۶ لایه انتقال	
۵۵۷	۱-۶ مشکلات کارایی در شبکه‌های کامپیوتری	۱-۶ خدمات انتقال (The Transport Service)	
۵۵۹	۲-۶ اندازه گیری کارایی شبکه	۱-۱-۶ خدمات ارائه شده به لایه‌های بالاتر	
۵۶۱	۳-۶ طراحی سیستم برای کارایی بهتر	۲-۱-۶ عملکردهای اولیه و توابع بنیانی لایه	
۵۶۵	۴-۶ پردازش سریع TPDU	انتقال	
۵۶۹	۵-۶ پروتکل‌هایی برای شبکه‌های گیگابیتی	۳-۱-۶ سوکت‌های برکلی (Berkeley Socket)	
۵۷۳	۷-۶ خلاصه	۴-۱-۶ مثالی از برنامه‌نویسی سوکت: یک	
۵۷۳	مسائل	سرویس دهنده اینترنتی فایل	
		۲-۶ مؤلفه‌های هر پروتکل انتقال	
۵۷۹	۷ لایه کاربرد	۱-۲-۶ آدرس دهی	
۵۷۹	۱-۷ سیستم نام ناحیه DNS	۲-۲-۶ برقراری اتصال	
۵۸۰	۱-۱-۷ فضای نام DNS	۳-۲-۶ خاتمه اتصال	
۵۸۲	۲-۱-۷ رکوردهای منابع	۴-۲-۶ کنترل جریان و بافرسازی	
۵۸۵	۳-۱-۷ سرویس دهنده نام	۵-۲-۶ مالتی پلکسینگ (تسهیم)	
۵۸۷	۲-۷ پُست الکترونیک	۶-۲-۶ جبران از کارافتادگی (Crash Recovery)	
۵۸۸	۱-۲-۷ معماری و سرویسها	۳-۶ یک پروتکل ساده انتقال	
۵۸۹	۲-۲-۷ عامل کاربر	۱-۳-۶ توابع اولیه ارائه خدمات در مثال فوق	
۵۹۲	۳-۲-۷ فرمت پیامها	۲-۳-۶ واحد انتقال در مثال فوق	
۵۹۸	۴-۲-۷ انتقال پیام	۳-۳-۶ بررسی مثال فوق از دید «ماشین حالت	
۶۰۱	۵-۲-۷ تحویل نهایی	محدوده»	
۶۰۶	۳-۷ تارنمای جهانی - وب	۴-۶ پروتکل‌های لایه انتقال در اینترنت: UDP	
۶۰۷	۱-۳-۷ بررسی ساختاری	۱-۴-۶ مقدمه‌ای بر UDP	
۶۲۲	۲-۳-۷ سندهای وب استاتیک	۲-۴-۶ فراخوانی پروسیجرهای راه دور (RPC)	

۳۳۷	سندهای وب دینامیک.....	۶۳۴
۳۳۷	پروتکل انتقال ابرمتن - HTTP.....	۶۴۱
۵۳۷	بهبود کارایی.....	۶۴۵
۶۳۷	وب بیسیم.....	۶۵۱
۴۷	چند رسانه ای.....	۶۶۱
۱۴۷	مقدمه ای بر صدای دیجیتال.....	۶۶۲
۲۴۷	فشرده سازی صدا.....	۶۶۳
۳۴۷	صدای جویباری.....	۶۶۵
۴۴۷	رادیوی اینترنتی.....	۶۶۹
۵۴۷	صدا روی IP.....	۶۷۱
۶۴۷	مقدمه ای بر ویدئو.....	۶۷۷
۷۴۷	فشرده سازی ویدئو.....	۶۸۱
۸۴۷	پخش فیلم بر حسب تقاضا.....	۶۸۷
۹۴۷	ستون فقرات چندپخش - Mbone.....	۶۹۳
۵۷	خلاصه.....	۶۹۷
	مسائل.....	۶۹۷
	امنیت شبکه ۷۰۳	
۱۸	رمزنگاری.....	۷۰۶
۱۱۸	مقدمه ای بر رمزنگاری.....	۷۰۷
۲۱۸	رمزهای جانشینی (Substitution Cipher).....	۷۱۰
۳۱۸	رمزنگاری جایگشتی (Transposition).....	۷۱۲
۴۱۸	رمز One-Time Pads.....	۷۱۳
۵۱۸	دو اصل اساسی در رمزنگاری.....	۷۱۸
۲۲۱	الگوریتمهای رمزنگاری با کلید متقارن.....	۷۲۱
۱۲۸	رمزنگاری DES.....	۷۲۲
۲۲۸	استاندارد پیشرفته رمزنگاری: AES.....	۷۲۵
۳۲۸	حالات رمز (Cipher Modes).....	۷۲۹
۴۲۸	رمزهای دیگر.....	۷۳۵
۵۲۸	تحلیل رمز (رمز شکنی).....	۷۳۵
۳۸	الگوریتمهای کلید عمومی (Public Key).....	۷۳۶
۱۳۸	RSA.....	۷۳۷
۲۳۸	الگوریتمهای کلید عمومی دیگر.....	۷۳۹
۴۸	امضاهای دیجیتالی.....	۷۴۰
۱۴۸	امضاهای دیجیتالی با کلید متقارن.....	۷۴۱
۲۴۸	امضاهای با کلید عمومی.....	۷۴۲
۳۴۸	خلاصه پیامها (Message Digests).....	۷۴۳
۴۸	حمله روز تولد (The birthday Attack).....	۷۴۷
۵۸	مدیریت کلیدهای عمومی.....	۷۵۰
۱۵۸	گواهینامه ها (Certificates).....	۷۵۰
۲۵۸	X.509.....	۷۵۲
۳۵۸	زیرساخت کلید عمومی.....	۷۵۳
۶۸	امنیت ارتباطات.....	۷۵۷
۱۶۸	IPsec.....	۷۵۷
۲۶۸	دیوارهای آتش (Firewalls).....	۷۶۲
۳۶۸	شبکه های خصوصی مجازی (VPN).....	۷۶۵
۴۶۸	امنیت شبکه های بی سیم.....	۷۶۶
۷۸	پروتکل های احراز هویت.....	۷۷۱
۱۷۸	احراز هویت بر اساس کلید مشترک و سزی.....	۷۷۲
۲۷۸	ایجاد کلید مشترک: مبادله کلید به روش «دیفی-هلمن».....	۷۷۷
۳۷۸	احراز هویت توسط مرکز توزیع کلید.....	۷۷۹
۴۷۸	احراز هویت با استفاده از Kerberos.....	۷۸۲
۵۷۸	احراز هویت با استفاده از رمزنگاری با کلید عمومی.....	۷۸۵
۸۸	امنیت نامه های الکترونیکی.....	۷۸۶
۱۸۸	PGP (Pretty Good Privacy).....	۷۸۶
۲۸۸	PEM (Privacy Enhanced Mail).....	۷۹۱
۳۸۸	S/MIME.....	۷۹۱
۹۸	امنیت وب.....	۷۹۲
۱۹۸	تهدیدها.....	۷۹۲
۲۹۸	نامگذاری مطمئن.....	۷۹۳
۳۹۸	SSL: لایه سوکت های امن.....	۸۰۱
۴۹۸	امنیت کدهای متحرک.....	۸۰۵
۱۰۸	زمینه ها و پی آمدهای اجتماعی.....	۸۰۸
۱۰۸	حریم خصوصی افراد (Privacy).....	۸۰۸
۲۰۸	آزادی بیان.....	۸۱۱
۳۰۸	مالکیت معنوی (Copyright).....	۸۱۳
۱۱۸	خلاصه.....	۸۱۶
	مسائل.....	۸۱۷
	واژه نامه ۸۲۳	
	محتویات دیسک فشرده همراه کتاب ۸۲۹	

مقدمه

هر یک از سه قرن گذشته را با یک تکنولوژی خاص بعنوان نماد آن قرن می‌شناسیم. قرن هیجدهم عصر سیستمهای بزرگ مکانیکی و انقلاب صنعتی بود، و قرن نوزدهم عصر بخار. تکنولوژی کلیدی قرن بیستم نیز جمع‌آوری، پردازش و توزیع اطلاعات بود. شبکه‌های گسترده و بین‌المللی تلفن، اختراع رادیو و تلویزیون، تولد و گسترش باورنکردنی صنعت کامپیوتر، و پرتاب ماهواره‌های مخابراتی از نمادهای این عصر هستند. با رشد سریع تکنولوژیهای جمع‌آوری، پردازش و توزیع اطلاعات، این زمینه‌ها بسرعت در هم ادغام شده، و تفاوت‌های آنها در حال محو شدن است. شرکتهایی که در اقصی نقاط دنیا شعبه و نمایندگی دارند، می‌توانند فقط با فشار یک دکمه از آخرین وضعیت دفاتر خود (حتی دور افتاده‌ترین آنها) مطلع شوند. اما جالب اینجاست که رشد تقاضا برای روشهای پیشرفته‌تر پردازش اطلاعات همیشه یک گام از سرعت رشد این تکنولوژیها جلوتر است. با اینکه صنعت کامپیوتر از صنایع دیگر (از جمله صنایع اتومبیل، و حمل و نقل هوایی) نسبتاً جوانتر است، اما در مدتی بس کوتاه به پیشرفتهای چشمگیری دست یافته است. در دو دهه اول، سیستمهای کامپیوتری بسیار متمرکز بودند، و معمولاً در یک اتاق بزرگ جا می‌گرفتند. کم نبودند مراکزی که این اتاقها دیوارهای شیشه‌ای داشتند، و بازدیدکنندگان با حیرت این موجودات عجیب‌الخلقه الکترونیکی را برانداز می‌کردند. دانشگاهها و شرکتهای متوسط معمولاً یکی دو کامپیوتر بیشتر نداشتند، و تعداد شرکتهایی که استطاعت خرید بیش از یک دوجین از آنها را داشته باشند، چندان زیاد نبود. هیچکس (شاید غیر از نویسندگان داستانهای علمی-تخیلی) حتی نمی‌توانست تصور کند که تا قبل از پایان قرن بیستم کامپیوترهایی با همان قدرت را بتوان روی یک تمبر پستی جای داد، و آنها را بصورت انبوه تولید کرد.

پیوند فرخنده کامپیوتر و مخابرات اتفاقی بود که هر دو صنعت را دچار تحولات عظیم کرد. اکنون دیگر مفهوم اتاقی با یک کامپیوتر بزرگ بنام «مرکز کامپیوتر»، که افراد کارهایشان را به آنجا می‌آورند، بکلی منسوخ شده است. مدل قدیمی کامپیوتر بزرگی که تمام کارهای محاسباتی سازمان را انجام می‌دهد، اکنون جای خود را به تعداد زیادی کامپیوتر کوچک متصل به هم داده است. به این سیستمها شبکه‌های کامپیوتری (computer networks) گفته می‌شود؛ موضوع این کتاب نیز طراحی و ساختار این شبکه‌هاست.

در این کتاب هر جا از «شبکه کامپیوتری» سخن می‌گوئیم، منظورمان مجموعه‌ای از کامپیوترهای مستقل است، که با یک تکنولوژی واحد به هم متصل شده‌اند. دو کامپیوتر وقتی «به هم متصلند»، که بتوانند با یکدیگر اطلاعات رد و بدل کنند. الزامی نیست که این اتصال از طریق سیمهای مسی باشد؛ فیبرهای نوری، امواج مایکروویو و مادون قرمز، و ماهواره‌های مخابراتی هم می‌توانند عامل این ارتباط باشند. بعداً خواهید دید که اندازه، شکل و ساختار

شبکه ها می تواند بسیار متفاوت باشد. همچنین بسیاری از افراد وقتی می شنوند که اینترنت یا وب هیچکدام شبکه کامپیوتری نیستند متعجب می شوند؛ اما در پایان این کتاب علت آنرا هم خواهید فهمید. فعلاً همین قدر کافیت بدانید که: اینترنت یک شبکه نیست، بلکه شبکه ایست از شبکه ها، و وب نیز یک سیستم توزیع شده است که بر پایه اینترنت کار می کند.

لازم است همین جا به یک اشتباه رایج بین دو اصطلاح شبکه کامپیوتری و سیستم توزیع شده (distributed system) اشاره کنم. یک سیستم توزیع شده مجموعه ایست از چندین کامپیوتر مستقل، که کاربر آنرا به شکل یک سیستم واحد و متجانس می بیند. در این سیستمها معمولاً یک لایه نرم افزاری (روی سیستم عامل) بنام میان افزار (middleware) است، که مدل مورد نظر را پیاده سازی می کند. وب (World Wide Web) نمونه ای از یک سیستم توزیع شده است، که در آن همه چیز از دیدگاه کاربر یک سند (صفحه وب) بنظر می رسد. در شبکه کامپیوتری این تجانس، مدل و نرم افزار وجود ندارد. کاربران بطور مستقیم با کامپیوترها در تماسند، و هیچ کوششی برای ایجاد تجانس بین آنها صورت نمی گیرد. کاربر بروشنی تفاوت های نرم افزاری و سخت افزاری کامپیوترها را می بیند، و اگر بخواهد برنامه ای را روی یکی از کامپیوترها اجرا کند، باید ابتدا وارد آن شود (log on). در حقیقت، یک سیستم توزیع شده نرم افزاریست که روی شبکه کار می کند، و تجانس و شفافیت آن توسط این نرم افزار تأمین می شود. به همین دلیل تفاوت سیستم توزیع شده با یک شبکه بیشتر در نرم افزار (بویژه سیستم عامل) نهفته است تا سخت افزار.

با این همه، شباهت های زیادی نیز بین این دو وجود دارد. مثلاً، سیستم های توزیع شده و شبکه ها هر دو به انتقال فایل نیاز دارند؛ تفاوت در اینست که این کار را چه کسی انجام می دهد، سیستم یا کاربر. با آنکه این کتاب درباره شبکه های کامپیوتری است، بسیاری از مطالب آن در سیستم های توزیع شده نیز مصداق دارد. برای کسب اطلاعات بیشتر درباره سیستم های توزیع شده به (Tanenbaum and Van Steen, 2002) نگاه کنید.

۱-۱ کاربردهای شبکه های کامپیوتری

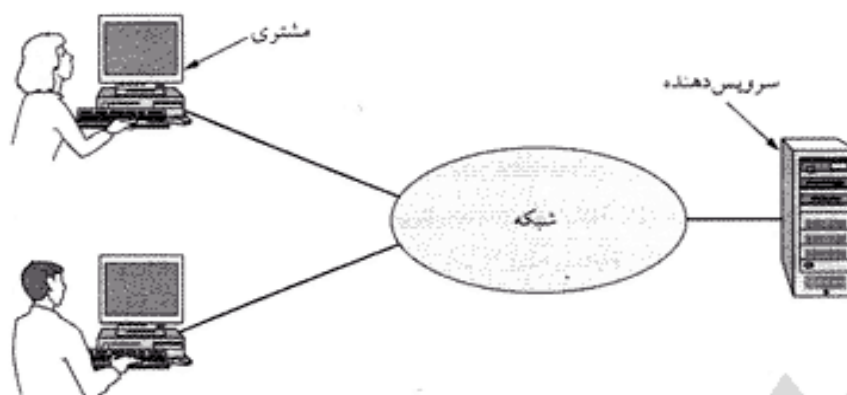
قبل از پرداختن به جزئیات فنی، بهتر است کمی درباره اینکه چرا مردم به شبکه های کامپیوتری اهمیت می دهند و چرا از آنها استفاده می کنند، صحبت کنیم (چرا که اگر کسی به شبکه اهمیت نمی داد، اصلاً شبکه ای ساخته نمی شد). ابتدا از کاربردهای سستی (از قبیل شرکتها و افراد) شروع می کنیم، و سپس به کاربردهای جدیدتر (مانند شبکه های متحرک و خانگی) خواهیم پرداخت.

۱-۱-۱ کاربردهای تجاری

اکثر شرکتها تعداد زیادی کامپیوتر برای کارهای مختلف (تولید، انبارداری، فروش، و حسابداری) دارند. شاید در ابتدا این کامپیوترها از یکدیگر جدا باشند، ولی در مرحله ای از کار برای یکپارچه کردن اطلاعات کل شرکت، مدیریت تصمیم می گیرد تا آنها را به هم متصل کند.

به بیان کلی تر، اشتراک منابع (resource sharing) به ما اجازه می دهد تا برنامه ها، تجهیزات و بخصوص داده ها را (صرف نظر از موقعیت فیزیکی افراد و منابع) در اختیار همه آنهايي که به این شبکه متصلند، قرار دهیم. ساده ترین مثال آن، چاپگر است که برای تمام کارکنان یک دفتر به اشتراک گذاشته شده است. پیداست که تک تک این افراد به یک چاپگر اختصاصی نیاز ندارند، و علاوه بر آن یک چاپگر شبکه اغلب ارزانتر، سریعتر و کم هزینه تر از تعداد زیادی چاپگرهای پراکنده است.

با این حال، اشتراک اطلاعات بسیار مهمتر از اشتراک تجهیزات فیزیکی (مانند چاپگر، اسکنر، و CD نویس) است. امروزه تمام شرکت های بزرگ و متوسط (و بسیاری از شرکت های کوچک) بشدت به اطلاعات کامپیوتری خود



شکل ۱-۱. شبکه‌ای با دو مشتری و یک سرویس دهنده.

(از قبیل اطلاعات مشتریان، انبار، اسنادهای مالی و حسابداری، و اطلاعات مالیاتی) وابسته‌اند. بانکی که تمام کامپیوترهای آن از کار افتاده باشند، پنج دقیقه هم نمی‌تواند دوام بیاورد. حتی شرکت‌های کوچکی مانند آژانس‌های مسافرتی و دفاتر خدمات حقوقی نیز بشدت به اطلاعات کامپیوتری خود متکی هستند.

در یک شرکت کوچک تمام کامپیوترها به احتمال زیاد در یک دفتر (و یا حداکثر یک ساختمان) قرار دارند، در حالیکه کامپیوترهای یک شرکت بزرگ می‌تواند در یک شهر یا کشور (و حتی در قاره‌های مختلف) پراکنده باشد. در این حالت، ممکنست مدیر فروشی که در نیویورک نشسته، به موجودی انبار شرکت در سنگاپور نیاز داشته باشد. بعبارت دیگر، حتی ۱۵۰۰۰ کیلومتر فاصله هم نباید خللی در دسترسی به اطلاعات وارد کند. در واقع می‌توان گفت، ما بدون «از بین بردن فاصله‌ها» هستیم.

در ساده‌ترین شکل، اطلاعات شرکت می‌تواند در یک یا چند پایگاه داده متمرکز باشد، و کارمندان شرکت بایستی بتوانند از راه دور به آنها دسترسی داشته باشند. در این مدل، اطلاعات در کامپیوترهای پُر قدرتی بنام سرویس دهنده (server) - که اغلب در یک مرکز و تحت کنترل سرپرست سیستم قرار دارند - نگهداری می‌شوند. کارمندان نیز، که در اینجا به آنها مشتری (client) گفته می‌شود، از راه دور و از پای کامپیوترهای معمولی خود به این اطلاعات دسترسی پیدا می‌کنند. (گاهی به فردی که از کامپیوتر استفاده می‌کند، نیز «مشتری» گفته می‌شود؛ بهر حال، از فحوای متن باید بتوانید متوجه شوید که منظور کامپیوتر است یا کاربر.) اتصال کامپیوترهای سرویس دهنده و مشتری از طریق شبکه صورت می‌گیرد (شکل ۱-۱ را ببینید). در این شکل شبکه به صورت یک بیضی ساده نشان داده شده است؛ وقتی بخواهیم شبکه را بصورت کلی و انتزاعی (و بدون هیچگونه جزئیاتی) نشان دهیم، از این روش استفاده خواهیم کرد.

به این آرایش مدل مشتری-سرویس دهنده (client-server model) گفته می‌شود، و در بسیاری از شبکه‌های کوچک و بزرگ کاربرد دارد چون مستقل از فاصله است. وب نیز بر مبنای مدل مشتری-سرویس دهنده ساخته شده است؛ وقتی یک صفحه وب را باز می‌کنید، در واقع آنرا از سرویس دهنده وب دریافت کرده، و در کامپیوتر خود (که در اینجا مشتری است) نمایش می‌دهید. در اکثر مواقع یک سرویس دهنده می‌تواند به تعداد زیادی مشتری سرویس بدهد.

در مدل مشتری-سرویس دهنده را دقیقتر بررسی کنیم، متوجه می‌شویم که دو پروسس (process) در آن دخیل هستند: یک پروسس روی کامپیوتر مشتری، و دیگری روی کامپیوتر سرویس دهنده. ارتباط از لحظه‌ای آغاز می‌شود، که پروسس مشتری از طریق شبکه یک پیام به پروسس سرویس دهنده فرستاده، و سپس به انتظار پاسخ

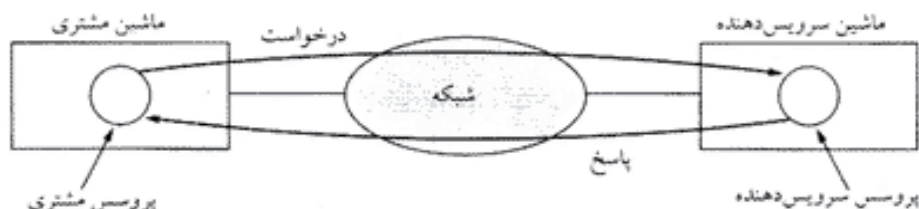
آن می‌ماند. وقتی پروسس سرویس دهنده درخواست مشتری را دریافت کرد، کار خواسته شده را انجام می‌دهد (با اطلاعات خواسته شده را تهیه می‌کند)، و پاسخ را به مشتری پس می‌فرستد. این فرآیند را در شکل ۱-۲ ملاحظه می‌کنید.

گاهی در یک شبکه کامپیوتری رابطه بین افراد اهمیت بیشتری نسبت به تبادل اطلاعات بین کامپیوترها دارد. چنین شبکه‌ای در واقع یک رسانه ارتباطی (communication medium) است. امروزه دیگر تقریباً هیچ شرکتی را نمی‌توان یافت که از سرویس پست الکترونیک (ایمیل: e-mail) استفاده نکند، و در واقع بسیاری از ارتباطات روزمره کارمندان از همین طریق صورت می‌گیرد. این روش آنقدر ساده و کارآمد است که خود باعث بروز مشکلات جدیدی شده است، چون رؤسای شرکتها هم یاد گرفته‌اند چطور فقط با فشار یک دکمه می‌توانند پیامهای (اغلب بی‌محتوای) خود را به این طرف و آن طرف بفرستند!

اما ایمیل تنها شکل از ارتباطات پیشرفته‌ای نیست که به لطف شبکه‌های کامپیوتری ممکن شده است. در یک شبکه، دو نفر که فاصله زیادی هم از یکدیگر دارند، می‌توانند بطور مشترک روی یک گزارش یا مقاله کار کنند. وقتی یکی از آنها تغییری در این گزارش می‌دهد، دیگری بلافاصله آنرا خواهد دید (و دیگر نیازی نیست روزها به انتظار بستگی چشم به در بدوزد). با این روش دیگر نیازی نیست غصه هماهنگ کردن کارمندی که هر کدام ساز خود را می‌زنند، بخورید.

یکی دیگر از امکانات ارتباطی شبکه‌ها، کنفرانس ویدئویی (video conferencing) است. به کمک این تکنولوژی، کارمندی که هزاران کیلومتر از هم فاصله دارند، می‌توانند یکدیگر را ببینند، صدای هم را بشنوند، و یا حتی مطالب خود را روی یک تخته سیاه مجازی بنویسند. کنفرانس ویدئویی جانشین بسیار مناسبی برای کنفرانسهای واقعی (که متضمن تحمل هزینه‌های سفر است) می‌باشد. گاهی گفته می‌شود که صنعت ارتباطات و حمل و نقل با یکدیگر مسابقه مرگ و زندگی گذاشته‌اند، و هر کدام پیروز شود، دیگری را از میدان بدر خواهد کرد. اتفاق دیگری که این روزها شتاب بیشتری گرفته، امکان تجارت الکترونیک بین شرکتها کوچک و بزرگ است. برای مثال، سازندگان کامپیوتر، اتومبیل و هواپیما می‌توانند قطعات مورد نیاز خود را از طریق شبکه‌های کامپیوتری به سازندگان این نوع قطعات سفارش دهند، و سپس آنها را مونتاژ و تبدیل به محصول نهایی کنند. سفارش و خرید قطعات در لحظه نیاز (زمان واقعی) لزوم نگهداری و انبار کردن مقدار زیادی از آنها را مستفی می‌کند.

گرایش تجاری دیگری که حتی اهمیت بیشتری پیدا کرده، فروش محصولات روی اینترنت است. این روزها شرکتها بسیاری (از قبیل خطوط هوایی، کتابفروشیها، و فروشندگان محصولات فرهنگی) به فروش محصولات خود از طریق اینترنت روی آورده‌اند. این شاخه از تجارت (که به تجارت الکترونیک - electronic commerce یا e-commerce - معروف است) در آینده رشد بسیار بیشتری خواهد کرد.



شکل ۱-۲. مدل مشتری سرویس دهنده بر «درخواست و پاسخ» مبتنی است.

۲-۱-۱ کاربردهای خانگی

سال ۱۹۷۷، وقتی از کین اولسین (رئیس شرکت Digital Equipment Corporation - که پس از IBM بزرگترین شرکت کامپیوتری دنیا محسوب می شد) پرسیدند چرا وارد بازار کامپیوترهای شخصی نمی شود، وی پاسخ داد: "هیچ دلیلی ندارد که هر کس توی خانه اش یک کامپیوتر داشته باشد." تاریخ ثابت کرد که اولسین اشتباه می کرد، و اکنون دیگر شرکت DEC وجود خارجی ندارد. اما چرا مردم برای کارهای خانگی خود کامپیوتر می خرند؟ نوشتن نامه، مقاله و حتی کتاب (و تا یادم نرفته، بازی) یکی از مهمترین دلایل آن است؛ اما این وضعیت امروزه در حال تغییر است. شاید مهمترین دلیل خرید کامپیوترهای خانگی در سالهای اخیر اینترنت باشد. کارهای که این قبیل افراد با کامپیوتر خود انجام می دهند، عمدتاً عبارتند از:

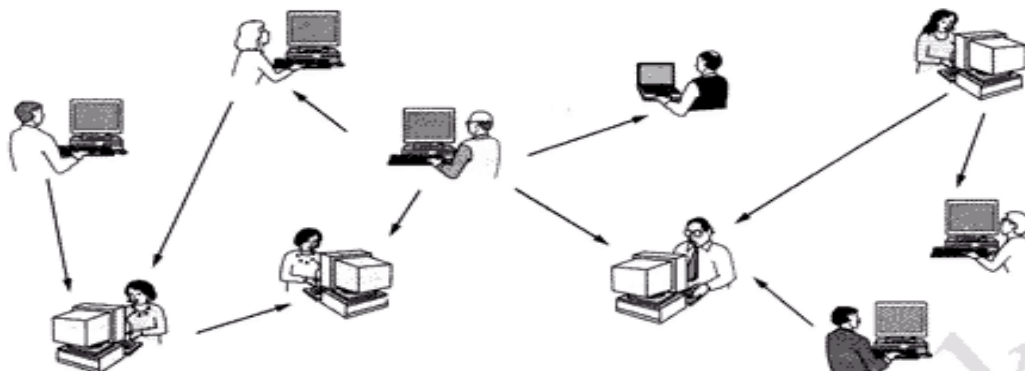
۱. دسترسی به اطلاعات پراکنده در سراسر دنیا
۲. ارتباطات دو جانبه
۳. سرگرمیهای تعاملی
۴. تجارت الکترونیک

امروزه منبع بسیار عظیمی از اطلاعات در تمامی زمینهها (از قبیل هنر، تجارت، آشپزی، بهداشت، تاریخ، سرگرمی، علم، ورزش و تفریحات سالم - و البته گاهی ناسالم!) روی اینترنت وجود دارد، که می توان به آنها دسترسی پیدا کرد. روزنامه های بسیاری روی اینترنت منتشر می شوند، که می توان اخبار را بدلخواه و بصورت گزینشی از آنها بدست آورد. حتی می توانید کاری کنید که مقاله دلخواه شما وقتی خواب هستید، از اینترنت بار شده و سپس چاپ شود، تا موقع صبحانه بتوانید با خیال راحت آنرا بخوانید. (به این ترتیب روزنامه فروشهای بیچاره بیکار خواهند شد، ولی مضبوطات هم هیچ وقت دل خوشی از آنها نداشتند.)

بعد از روزنامه ها و مجلات الکترونیکی نوبت کتابخانه های دیجیتالی است. بسیاری از سازمانهای علمی معتبر مانند ACM (www.acm.org) و IEEE (www.computer.org) مدتهاست که انتشارات و کنفرانسهای متعددی روی اینترنت برگزار می کنند؛ و این روند بسرعت در حال گسترش است. بنظر می رسد که رواج کتابخوانی اینترنتی فقط به قیمت، اندازه و وزن کامپیوترهای کتابی بستگی دارد. (شاید هنوز عده ای به این آینده شک داشته باشند؛ اما بهتر است آنها بلایی را که دستگاه چاپ گوتنبرگ سر کتابهای زیبای خطی آورد، بیاد بیاورند.)

تمام کاربردهایی که در بالا نام بردیم، متضمن ارتباط فرد با یک منبع اطلاعات روی اینترنت بود. اما روش دیگری نیز برای برقراری ارتباط از طریق اینترنت وجود دارد، و آن ارتباط فرد به فرد است (این پاسخ تکنولوژی قرن بیست و یکم است به تلفن قرن نوزدهمی). امروزه میلیونها نفر در سراسر دنیا بطور روزمره از ایمیل استفاده می کنند؛ صوت و تصویر هم مدتهاست به جزیی جدایی ناپذیر از آن تبدیل شده است (و باید بزودی منتظر بوهای اینترنتی هم باشیم!).

این روزها همه نوجوانانی که سری میان سرها در می آورند، به برنامه های پیام رسان قوری (instant messaging) معتاد هستند. این برنامه ها (که از برنامه talk در سیستمهای یونیکس مشتق شده اند) به افراد امکان می دهند پیامهای متنی خود را بلافاصله (و بدون تأخیر زمانی) با هم مبادله کنند. نسخه هایی از این برنامه ها که به چندین نفر اجازه می دهند تا در آن واحد با هم گفتگو کنند، به اتاق گفتگو (chat room) معروفند. گروههای خبری (newgroup) از امکانات قدیمی و پرتعداد اینترنت است، که امکان بحث درباره موضوعات بسیار متنوعی را به شما می دهد. در این سیستم پیامی که می فرستید، به تمام آنهایی که عضو گروه خبری هستند خواهد رسید (خوششان بیاید، یا نیابد). مبادلات گروه خبری (بر خلاف اتاق گفتگو) بصورت بلافاصله و در زمان واقعی نیست، و پیامها در نقطه ای ذخیره می شوند، تا کاربر بتواند هر زمان که خواست آنها را بخواند.



شکل ۱-۳. در یک سیستم همتا-به-همتا مشتری یا سرویس دهنده ثابتی وجود ندارد.

نوع دیگری از امکانات ارتباطی موجود در اینترنت، ارتباط همتا-به-همتا (peer-to-peer) است. این مدل تفاوت اساسی با مدل مشتری-سرویس دهنده دارد (به Parameswaran et al., 2001 نگاه کنید). در این مدل ارتباط افراد در یک گروه غیرثابت و ناپایدار صورت می گیرد (شکل ۱-۳ را ببینید). در واقع، هر فرد می تواند مستقیماً با هر فرد (یا افراد) دیگر تماس برقرار کند، و چیز ثابتی بعنوان سرویس دهنده یا مشتری وجود ندارد. بزرگترین نمونه ارتباط همتا-به-همتا در حوالی سال ۲۰۰۰ با سرویسی بنام Napster شکل گرفت؛ این سرویس در اوج خود امکان ارتباط بیش از ۵۰ میلیون نفر را فراهم می آورد، که بصورت غیر قانونی به رد و بدل کردن موزیک مبادرت می کردند (این بزرگترین نقض حق التألیف در تاریخ موسیقی بود؛ به Lam and Tan, 2001 و Macedonia, 2000 نگاه کنید). ایده کار نسبتاً ساده بود: هر نفر می توانست آهنگهایی را که در هارد دیسک خود داشت، در پایگاه داده مرکزی Napster ثبت کند؛ افرادی هم که بدنبال آهنگ خاصی بودند، این پایگاه داده را جستجو کرده، و بعد مستقیماً به سراغ آن می رفتند. Napster ادعا می کرد که هیچ حق التألیفی را نقض نمی کند، چون اساساً آهنگها در کامپیوترهای آن ذخیره نمی شوند؛ اما دادگاه با این نظر موافق نبود، و حکم به تعطیلی آن داد. سیستمهای همتا-به-همتا جدید با هوشتر شده اند، چون پایگاه داده مرکزی را حذف کرده اند و بجای آن این اطلاعات در کامپیوتر تک تک افراد ذخیره می شود، و آنها لیستی از افراد مجاور خود را هم در اختیار دارند. در این روش جستجو کمی بیشتر طول می کشد (که بار آن هم بر دوش کامپیوترهاست)، ولی در نهایت به همان اندازه مؤثر است.

همه برنامه های همتا-به-همتا هم غیرقانونی نیستند. برای مثال، برنامه هایی هستند که اجازه می دهند تا آهنگها و فیلمهای مجاز و حتی عکسهای خانوادگی خود را روی اینترنت به اشتراک بگذارید، و یا بازیهای دسته جمعی انجام دهید. در حقیقت، پرفرودارترین کاربرد اینترنت، یعنی ایمیل، ذاتاً یک سیستم همتا-به-همتا است، و بنظر می رسد در آینده این سیستمها حتی گسترده تر شوند.

جرایم الکترونیکی به دزدی آثار و نقض حق التألیف محدود نمی شود، و این روزها قمارخانه ها هم پایشان به اینترنت باز شده است. کامپیوترها قادرند هر کاری انجام دهند، پس چرا قمار نکنند؟ البته قمار در بسیاری از کشورها غیرقانونی است، اما مشکل اینجا است که در چند جا هم قانونیست (مانند انگلستان)، و صاحبان کازینوها به امکانات بالقوه اینترنت برای قمار واقف شده اند. اما اگر کازینو و قمارباز در دو کشور متفاوت (که قوانین متفاوتی هم در زمینه قمار دارند) باشند، چطور؟ سؤال خوبیست!

ارتباطات اینترنتی در زمینه تماسهای تلفنی، ویدئویی و رادیو نیز تحولات وسیعی ایجاد کرده اند. آموزش از

راه دور (telelearning) نیز یکی دیگر از امکاناتیست که اینترنت عرضه کرده است. (تصورش را بکنید که ساعت ۸ صبح سر کلاس درس حاضر باشید، بدون اینکه لازم باشد قبل از آن از رختخواب بیرون بیایید!) بنظر می رسد در دراز مدت اینترنت بزرگترین نقش را در بهبود ارتباطات انسانی بازی کند.

سومین دسته از کاربردهای خانگی شبکه های کامپیوتری، صنعت سرگرمی و تفریحات (با رشدی سرسام آور) است. داغترین بحث در این زمینه پخش فیلم برحسب تقاضا (video on demand) است. شاید تا ده سال دیگر براحتی بتوانید فیلم دلخواه خود را انتخاب کرده، و همان لحظه روی صفحه تلویزیون تماشا کنید. فیلمهای جدید تعاملی (interactive) خواهند بود، بدین معنا که بیننده می تواند مسیر سناریو را بدلخواه خود تغییر دهد. تلویزیون زنده (شرکت مستقیم و بلافاصله در مسابقات و شوهای تلویزیونی) نیز یکی دیگر از امکانات آینده است.

بازیهای تعاملی یکی دیگر از امکانات شبکه است که شاید آینده آن حتی از پخش فیلم برحسب تقاضا نیز داغتر باشد. حتی همین حالا هم گروههای بزرگی از جوانان ماجراجو شب و روز مشغول بازی موش و گربه و جنگهای هوایی، زمینی و دریایی در زوایای تاریک و دورافتاده این دنیای مجازی (اینترنت) هستند. اگر آینده بتواند امکانات پخش سه بعدی و کیفیت بالا را عرضه کند، دیگر این دنیای مجازی هیچ چیز از دنیای واقعی کم نخواهد داشت.

چهارمین دسته از کاربردهای شبکه، شاید وسیعترین آنها باشد: خرید از خانه (home shopping). امروزه میلیونها نفر در سراسر جهان هر روز مایحتاج خود را بطور مستقیم از اینترنت تهیه می کنند، و هرگز پا از خانه بیرون نمی گذارند (حداقل برای خرید). هزاران شرکت بزرگ و کوچک کاتالوگ محصولات خود را بصورتی جذاب روی اینترنت گذاشته اند، و برای خرید هر یک از آنها کافیسست روی جنس موردنظر یک کلیک کنید. کالایی را خریدید، و لی نمی دانید چطور کار می کند؟ نگران نباشید، باز هم اینترنت به شما کمک می کند، و هر اطلاعات و راهنمایی که بخواهید در اختیارتان قرار می دهد.

می خواهید صورت حسابهای خود را پرداخت کنید؟ از آخرین وضعیت حسابهای بانکی خود مطلع شوید؟ و یا سرمایه گذاری جدیدی بکنید؟ باز هم اینترنت در خدمت شماست. امروزه میلیونها نفر در سراسر جهان کارهای مالی و بانکی خود را بصورت الکترونیکی انجام می دهند، و با تقویت مسائل امنیتی شبکه این روند حتی گسترش بیشتری نیز خواهد یافت.

یکی از زمینه هایی که شاید هیچکس تصور اینترنتی شدن آنرا نمی کرد، سمساری بود. حراج اینترنتی اشیاء دست دوم اینک به یکی از تجارتهای بزرگ تبدیل شده است. بر خلاف تجارت الکترونیک معمولی که از مدل مشتری - سرویس دهنده استفاده می کند، حراج اینترنتی در واقع یک سیستم همتا-به-همتا یا خریدار-به-خریدار (consumer-to-consumer) است. امروزه با اصطلاحات زیادی از این دست برخورد می کنید، که در آنها بجای "to" از "2" استفاده می شود (چون تلفظ آنها یکسان است). در شکل ۱-۴ تعدادی از رایجترین این اصطلاحات را ملاحظه می کنید.

شکی نیست که کاربردهای شبکه و اینترنت در آینده بسرعت افزایش خواهد یافت، و در زمینه هایی رسوخ خواهد کرد که امروز حتی به تصور کسی نمی آید. چه کسی در سال ۱۹۹۰ می توانست تصور کند که بخش بزرگی از درآمد شرکتهای تلفن از محل پیامهای کوتاهی باشد که دانش آموزان دبیرستانی بطور خستگی ناپذیر و در حالیکه سوار اتوبوس مدرسه هستند، با تلفن همراه خود برای دوستانشان می فرستند؟ این شرکتهای بخوبی می دانند که سرویس پیام کوتاه (Short Message Service - SMS) بسیار سودآور است.

شبکه های کامپیوتری برای افرادی که دور از شهرها زندگی می کنند، نیز مفید است. اینان می توانند راحت در

اصطلاح	نام کامل	مثال
B2C	فروشنده-به-خریدار	خرید کتاب روی اینترنت
B2B	فروشنده-به-فروشنده	خرید قطعات یدکی توسط تولیدکننده
G2C	دولت-به-خریدار	توزیع فرمهای مالیاتی از طریق اینترنت
C2C	خریدار-به-خریدار	حراج اشیاء دست دوم
P2P	همتا-به-همتا	اشتراک فایل

شکل ۱-۴. برخی از انواع تجارت الکترونیک.

روستاهای خود زندگی کنند، و در عین حال به تمام امکانات شهرهای بزرگ هم دسترسی داشته باشند. دانشگاههای آینده به احتمال زیاد حالت ملی و محلی خود را از دست داده، و بصورت بین المللی درخواهند آمد. درمان از راه دور (telemedicine) امروزه به کنترل بیماران محدود می شود، ولی چه کسی می تواند امکانات بالقوه آنرا پیش بینی کند. (یا مثلاً، چقدر خوب می شد اگر می توانستیم یک دوربین دیجیتالی در یخچال خود نصب کنیم، تا هر وقت شیر تمام شد بتوانیم سر راه خانه شیر بخریم!)

۳-۱-۱ کاربردان سیار

کامپیوترهای سیار، مانند کامپیوترهای کتابی و دستیاران دیجیتالی (PDA)، یکی از سریعترین رشدها را در صنعت کامپیوتر تجربه می کنند. اغلب دازندگان این وسایل میل دارند حتی وقتی از خانه دور و یا در سفر هستند، با کامپیوتر خانگی یا دفتری خود ارتباط داشته باشند. در این قبیل موارد دیگر شبکه های کابلی محلی از اعراب ندارد، و باید به فکر شبکه های بیسیم (wireless network) باشیم. در این قسمت نگاهی به کاربردهای شبکه های بیسیم خواهیم داشت.

جالبترین کاربرد شبکه های بیسیم در ایجاد دفاتر سیار است. اغلب افراد میل دارند در سفر همان کارهایی را انجام دهند که در دفتر کار خود انجام می دهند (ایمیل و فکس بفرستند، تلفن راه دور بزنند، فایل های خود را باز کنند، و یا در وب گشت بزنند)، و اصلاً هم کاری به این ندارند که کجا هستند! برای مثال، امروزه در اغلب کنفرانسهای کامپیوتری گردانندگان کنفرانس یک شبکه بیسیم در محوطه کنفرانس راه می اندازند، و هر کسی می تواند با استفاده از یک مودم بیسیم به اینترنت دسترسی پیدا کند. بسیاری از دانشگاهها هم در محوطه خوابگاهی خود شبکه های بیسیم دارند، و به دانشجویان امکان می دهند تا زیر درختان محوطه نشسته و ایمیل های خود را چک کنند، و یا در کتابخانه دانشگاه دنبال مقاله بگردند.

شبکه های بیسیم در امور حمل و نقل (کشتیها، کامیونها و تاکسیها) تحولی بزرگ ایجاد کرده اند. برای مثال، در بسیاری از شهرهای بزرگ رانندگان تاکسی مستقل بوده و عضو هیچ شرکت یا اتحادیه ای نیستند. وقتی کسی به تاکسی نیاز دارد، به یک سرویس مرکزی تلفن می کند، و مشخصات وی (از قبیل مبدأ و مقصد) توسط این سرویس به تمام تاکسیها ارسال می شود. اولین تاکسی که مایل به انجام این سرویس باشد، با فشار یک دکمه اعلام آمادگی کرده و به سراغ مسافر می رود.

شبکه های بیسیم از نظر نظامی نیز اهمیت فوق العاده ای دارند. هیچ ارتشی نمی تواند در جنگهای بزرگ به شبکه های عمومی تکیه کند، و بهتر است شبکه ای خاص خود بر پا کند؛ و چه چیزی بهتر از یک شبکه بیسیم. با وجود شباهتهای بسیار بین شبکه های بیسیم و کامپیوترهای سیار، آنها یکی نیستند (شکل ۱-۵ را ببینید). به تفاوت شبکه های بیسیم ثابت (fixed wireless) و شبکه های بیسیم سیار (mobile wireless) توجه کنید. در بسیاری از دفاتر، کامپیوترهای کتابی سیار بصورت ثابت به شبکه محلی متصل شده اند؛ از طرف دیگر کامپیوتری که با استفاده از مودم به شبکه وصل می شود، سیار است - ولی مسلماً به آن بیسیم نمی توان گفت.

کاربردها	سیار	بسیم
کامپیوترهای رومیزی در دفتر کار	No	No
کامپیوتری که با مودم به شبکه وصل شده	Yes	No
شبکه بیسیم مستقر در یک ساختمان	No	Yes
دفاتر سیار؛ PDAها	Yes	Yes

شکل ۱-۵. ترکیب شبکه‌های بیسیم و کامپیوترهای سیار.

از طرف دیگر، هر شبکه بیسیمی الزاماً سیار نیست. ساختمانهای بسیاری وجود دارند، که بدلیل مشکلات کابل کشی از شبکه‌های بیسیم استفاده می‌کنند. امروزه نصب شبکه‌های بیسیم بسیار ساده شده است، و دردسرهای کابل کشی را هم ندارد.

البته ترکیب بیسیم با کامپیوترهای سیار نیز عملیست، و امروزه کاربردهای مهمی دارد. کسانی که در انبارهای بزرگ یا فروودگاههای شلوغ کار می‌کنند، تجربه استفاده از این سیستمهای ترکیبی را دارند. این کامپیوترها با اطلاعات ورودی کمی که می‌گیرند، و با اتصال بیسیم به پایگاه داده مرکزی، کار خود را با سرعت و دقت انجام می‌دهند. با رشد تکنولوژی بیسیم، مسلماً کاربردهای آن نیز گسترش خواهد یافت. اجازه دهید نگاهی به این احتمالات بیندازیم.

پارکومترهای بیسیم کار دولت و افراد را راحتتر خواهند کرد. این پارکومترها می‌توانند کارت اعتباری نیز قبول کرده، و با سرعت اعتبار آنرا چک کنند؛ و وقتی مدت پارک تمام شد، اگر هنوز اتومبیل شما آنجا بود، نزدیکترین پلیس را خبر می‌کنند تا آنرا جریمه کند! تخمین زده شده که فقط پلیس ایالات متحده می‌تواند از این طریق ۱۰ میلیارد دلار بر درآمد خود بیفزاید (Harte et al., 2000). این روش اثرات مثبت زیست محیطی نیز دارد، چون رانندگان خودروها مطمئنند که راهی برای فرار از دست پلیس ندارند، و به ناچار به وسایل نقلیه عمومی روی می‌آورند.

امروزه ماشینهای خودکار فروش غذا، نوشابه و چیزهای دیگر در همه جا یافت می‌شوند. اما غذا و نوشابه که از آسمان وارد این ماشینها نمی‌شود؛ یک مأمور سوار بر کامیون هر از چند گاهی به این ماشینها سرکشی می‌کند، تا در صورت نیاز آنها را پر کند. اگر این ماشینها هر روز موجودی خود را از طریق شبکه بیسیم به مرکز اطلاع دهند، مأمور ما می‌داند سراغ کدام ماشینها باید برود، و چه مقدار کالا احتیاج دارد (و حتی می‌تواند مسیر حرکت خود را به بهترین نحو برنامه ریزی کند). البته این اطلاعات را از طریق خطوط تلفن هم می‌توان به مرکز منتقل کرد؛ اما کشیدن یک خط تلفن برای هر ماشین (آن هم برای یک تماس در روز) اصلاً مقرون بصرفه نیست.

زمینه دیگری که شبکه‌های بیسیم می‌توانند باعث صرفه جویی شوند، قرائت کتورهای مختلف خانگی است. اگر هر کتور آب، برق و گاز اطلاعات خود را از طریق شبکه بیسیم به شرکت مربوطه منتقل کند، دیگر نیازی به مراجعه کتورخوانها به درب منازل نیست. به همین ترتیب، اگر آشکارسازهای دود و حرارت آلارم خود را بجای راه انداختن آژیر و سر و صدا) مستقیماً به مراکز آتش نشانی بفرستند، بسیار مؤثرتر خواهد بود. با کاهش قیمت دستگاههای رادیویی (که مبنای شبکه‌های بیسیم هستند)، وسایل اندازه گیری و گزارش دهی بیشتری به استفاده از آنها روی خواهند آورد.

زمینه دیگری از کاربردهای شبکه‌های بیسیم (که از مدتها قبل نیز انتظار آن می‌رفت)، ادغام تلفنهای همراه و PDA ها با کامپیوترهای بیسیم است. اولین PDA های بیسیم قادر بودند صفحات ساده وب را روی صفحات کوچک خود نمایش دهند. این سیستم که WAP 1.0 (Wireless Application Protocol) نام داشت، بدلیل خوانا نبودن صفحات، پهنای باند کم، و سرویس ضعیف با شکست مواجه شد. ولی با سرویسها و وسایل جدید WAP 2.0 اوضاع مسلماً بهتر خواهد شد.

ترکیب این تکنولوژیها می تواند به سرویس جدیدی منجر شود که می توان آنرا تجارت سیار (mobile commerce) نامید (Senn, 2000). این پدیده در واقع عامل ترکیب کننده تکنولوژی PDA های بیسیم با تجارت الکترونیک است، که این روزها همه از آن سهم می خواهند، و این امیدواری وجود دارد که افراد به خرید و انجام کارهای بانکی با آن روی آورند. PDA بیسیم می تواند در فروشگاهها و مراکز خرید بعنوان عامل انتقال پول و یا کارت اعتباری عمل کند (بدین ترتیب که این هزینه ها بعداً با صورت حساب تلفن پرداخت شود). نکته مثبت این روش برای فروشندگان آن است که هزینه های کار با شرکتهای اعتباری را به مقدار زیادی پائین می آورد. البته عیب بزرگی نیز دارد: خریداران می توانند قبل از خرید با PDA بیسیم خود قیمتها را با فروشندگان دیگر مقایسه کنند. اگر شرکتهای تلفن سرویس فرائت بارکد را هم به این PDA ها اضافه کنند، که دیگر اوضاع خرابتر می شود! چون دیگر حتی با مخفی کردن قیمتها هم نمی توان جلوی مقایسه قیمتها توسط خریدار را گرفت (کاری که خیلی از فروشندگان به آن امید بسته اند)!

از آنجائیکه محل این دستگاهها همیشه برای اپراتور سیستم مشخص است، می توان سرویسهای خاصی را در اختیار کاربران آنها گذاشت؛ مثلاً، می توان آدرس نزدیکترین کتابفروشی یا رستوران چینی را در اختیار وی گذاشت، و یا آخرین پیش بینی وضعیت هوا را به وی اعلام کرد.

امکانات و کاربردهای این سرویس جدید می تواند بسیار فراتر از مثالهای ساده فوق باشد؛ و خوبی قضیه اینست که کاربران تلفنهای موبایل عادت دارند برای هر چیزی پول بدهند (درست بر خلاف کاربران اینترنت که همه چیز را مجانی می خواهند)! اگر یک سایت اینترنتی برای قبول کارت اعتباری از شما درخواست پول کند، فوراً فریادتان بلند خواهد شد، ولی اگر همین اتفاق روی سرویسهای موبایل بیفتد، بدون هیچ اعتراضی قبول می کنید (البته فعلاً).

بگذارید کمی هم به آینده نگاه کنیم: شبکه های شخصی (Personal Area Network)، و کامپیوترهای پوشیدنی (wearable computer). به تازگی IBM ساعتی ساخته که سیستم عامل لینوکس (Linux) روی آن اجرا می شود، و می تواند به اینترنت وصل شده و ایمیل رد و بدل کند (Narayanaswami et al., 2002). در آینده دیگر چیزی بعنوان کارت ویزیت بین افراد رد و بدل نخواهد شد، و آنها می توانند با یک تماس ساعت مچی تمام اطلاعات طرف مقابل را دریافت کنند. به احتمال زیاد کامپیوترهای پوشیدنی (که اطلاعات زیستی فرد را در خود دارند) جای کارتهای مغناطیسی را برای ورود به مکانهای حساس خواهند گرفت، و یا می توانند در هر لحظه مکان فرد را اعلام کنند. امکانات این سیستمها تقریباً بی شمار است.

ساعتها و رادیوهای هوشمند را همه ما سالهاست از طریق فیلمهای جیمز باند می شناسیم، ولی آیا تا بحال گرد و غبار هوشمند به گوشتان خورده است؟ محققان دانشگاه برکلی اخیراً یک کامپیوتر بیسیم ساخته اند، که در مکعبی با ابعاد 1 mm جای می گیرد (Warneke et al., 2001). با این کامپیوترها می توان مسیر حرکت چمدان در فرودگاهها، و یا جانوران (مثلاً، پرندگان مهاجر) را با دقت کنترل کرد.

۱-۱-۱ تبعات اجتماعی

گسترش روزافزون شبکه های کامپیوتری باعث ایجاد مسائل اجتماعی، اخلاقی و سیاسی خاص خود شده است که در این قسمت برخی از آنها را بررسی می کنیم. یکی از امکانات شبکه های کامپیوتری تبادل آزاد و سریع اطلاعات و اخبار است. البته تا وقتی این پیامها در محدوده های فنی باقی بماند، مشکل چندانی وجود نخواهد داشت؛ در دسر وقتی شروع می شود که صحبت به مسایل حساس (از قبیل سیاست، مذهب یا سکس) کشیده شود.

نظرها و دیدگاههایی که توسط اعضای یک گروه خبری پست می شود، ممکنست برای افراد دیگر بسیار ناهنجار و موهن باشد؛ و وقتی پیامها به متن محدود نشود، کار بدتر هم خواهد شد. امروزه پراحتی می توان

عکسهای بسیار واضح و باکیفیت عالی (و یا حتی کلیپهای ویدئویی کوچک) را از طریق اینترنت منتشر کرد. برخی از افراد در زندگی به فلسفه «زندگی کن، و بگذار زندگی کنند» معتقدند، اما عده زیادی هم هستند که احساس می‌کنند به برخی مطالب (مانند حمله به کشورها یا مذاهب دیگر، صور قبیحه و غیره) نباید اجازه انتشار داد. کشورهای مختلف هم دارای قوانین متفاوتی در این زمینه‌ها هستند؛ و جدال از همین جا شروع می‌شود. در این میان بسیاری از افراد اپراتورهای شبکه را (مانند روزنامه‌ها و مجلات) مسئول محتویات شبکه می‌دانند، اما واقعیت اینست که یک شبکه بیشتر شبیه اداره تلفن و پست است تا روزنامه یا مجله (و نمی‌توان آنرا مسئول چیزهایی که از این طریق مبادله می‌شود، دانست). از طرف دیگر، اگر اپراتور شبکه اجازه داشته باشد مطالب را سانسور کند، به احتمال زیاد (برای فرار از متهم شدن) روی کوچکترین چیزها هم انگشت خواهد گذاشت، و بدین ترتیب حقوق افراد در زمینه آزادی بیان از بین خواهد رفت. بحث موافق و مخالف همچنان ادامه دارد (و براحتی می‌توان حدس زد که به این زودی‌ها هم به نتیجه نخواهد رسید).

بحث جالب دیگر حقوق و رابطه کارگر و کارفرما است. بسیاری از افراد در محل کار خود ایمیل می‌فرستند، و یا ایمیل‌های رسیده را می‌خوانند. برخی از کارفرمایان ادعا می‌کنند که آنها حق دارند ایمیل‌های کارمندان خود را بخوانند و یا آنها را سانسور کنند؛ و صد البته کارمندان با این حرفها موافق نیستند!

حتی اگر بپذیریم کارفرمایان در ادعای خود محق هستند، آیا می‌توان این رابطه را به دانشجو و دانشگاه (و یا دانش آموز و مدرسه) تعمیم داد؟ در سال ۱۹۹۴ دانشگاه کارنگی سلون تصمیم گرفت برخی از پیامهای رسیده را که در آنها به موضوعات سکسی پرداخته شده بود، سانسور کند (با این استدلال که این مطالب برای افراد زیر ۱۸ سال مناسب نیست). سالها طول کشید تا پس لرزه‌های این اقدام فروکش کند.

و از همه مهمتر رابطه دولت با شهروندان است. اداره آگاهی فدرال ایالات متحده (FBI) سالهاست سیستمی را در مراکز ارائه سرویس اینترنت (ISP) نصب کرده، که به آن اجازه می‌دهد تا ایمیل‌های ورودی و خروجی را تجسس کند (Blaze and Bellovin, 2000; Sobel, 2001; Zacks, 2001). نام این سیستم Carnivore (گل گوشتخوار) بود، که بدلیل حساسیتهای ایجاد شده در جامعه به نام کم ضررتر DCS1000 تغییر داده شد (با این حال کار آن همچنان جاسوسی در ایمیل‌های مردم بود). طبق اصلاحیه چهارم قانون اساسی ایالت متحده آمریکا، دولت بدون مجوز قانونی حق تجسس در احوال شخصی افراد را ندارد. اینکه این قانون نوشته شده در قرن هیجدهم هنوز در قرن بیست و یک اعتبار دارد یا خیر، را آینده روشن خواهد کرد.

فضولی در کار مردم به دولت محدود نمی‌شود؛ بخش خصوصی هم از این گناه در امان نیست. برای مثال، مرورگرهای وب از فایل‌های کوچکی بنام کوکی (cookie) استفاده می‌کنند که اطلاعات شخصی افراد را در اختیار شرکتها می‌گذارد، و حتی می‌تواند منجر به افشای شماره کارتهای اعتباری و اطلاعات محرمانه دیگر روی اینترنت شود (Berghel, 2001).

در شبکه‌های کامپیوتری می‌توان پیامهای بدون نام و نشانی فرستاد، که در جای خود می‌تواند مفید باشد. مثلاً، کارمندان، دانشجویان و یا مردم عادی می‌توانند بدین طریق اعمال خلاف رؤسای شرکتها، استادان و سیاستمداران را به اطلاع عموم برسانند بدون آنکه از اقدامات تلافی‌جویانه آنها ترسی به دل راه دهند. (البته در بسیاری از کشورها از این قبیل اطلاعات بدون منبع نمی‌توان در دادگاه بعنوان مدرک جرم استفاده کرد.)

وضعیت فعلی شبکه‌های کامپیوتری شبیه موقعیت کتابهای چاپی در ابتدای اختراع این صنعت بود: افراد عادی وسیله‌ای بدست آورده بودند تا با آن صدای خود را به گوش دیگران برسانند. اما این آزادی با خود تبعات اجتماعی، سیاسی و اخلاقی خاصی بدنال داشت، که همچنان لاینحل باقی مانده است. زندگی همیشه بدین منوال است: هر سکه‌ای دو رو دارد. اینترنت هم از این قاعده مستثنی نیست. امکان

دسترسی سریع و آسان به اطلاعات از مغان اینترنت است، ولی خروارها اطلاعات منفی، غلط و گمراه‌کننده وجه دیگر آن است. راهنمایی بهداشتی که تازگی در اینترنت خوانده‌اید (و احتمالاً می‌خواهید به آن عمل کنید)، ممکنست از یک برنده جایزه نوبل آمده باشد، یا یک دانش‌آموز بازیگوش دبیرستانی.

شبکه‌های کامپیوتری انواع جدیدی از جرم و رفتارهای ضداجتماعی را نیز با خود آورده‌اند. هر روز که صندوق پستی خود را باز می‌کنید، دهها و صدها پیام مزخرف و بدردنخور در آن می‌بینید (و حتی کم‌کم به آن عادت کرده‌اید). این نتیجه کار افراد پست که میلیونها آدرس ایمیل را روی یک CD جمع کرده، و (بدون رضایت صاحبان این آدرسها) به این و آن می‌فروشند. و تازه اینها دسته‌بی‌آزارها هستند؛ این روزها کسی پیدا نمی‌شود که صابون و ویروسهایی که از طریق ایمیل منتشر می‌شوند، به تنش نخورده باشد.

دزدی هویت یکی دیگر از خطرات سرعت اطلاعات از طریق اینترنت است. (درباره سرقتهای ادبی و نقض گسترده و وسیع قانون حق‌التألیف در اینترنت قبلاً هم صحبت کردیم.)

بسیاری از این مشکلات حاصل ضعف (و یا عدم رعایت) مسائل امنیتی در اینترنت است. اگر تمام پیامهای ایمیل بصورت رمز در آیند، سرعت اطلاعات بسیار مشکلتر خواهد شد (این تکنولوژی توسعه زیادی یافته، که در فصل ۸ مفصلاً به آن خواهیم پرداخت). مشکل اینجاست که بالا بردن سطح ایمنی مترادف است با بالا رفتن هزینه، و این چیزی نیست که به آسانی پذیرفته شود. تعداد زیادی از این مسایل نیز به مشکلات و باگهای موجود در نرم‌افزارها مربوط می‌شود، که خود حاصل بزرگتر و پیچیده‌تر شدن آنهاست. اگر روی برنامه‌ها بر حسب بزرگی و پیچیدگی آنها مالیات بسته شود، شاید این مشکل تا حدی حل شود؛ البته خیلی‌ها این راه حل را نمی‌پسندند! پس دادن پول برنامه‌های معیوب نیز می‌تواند راه حل خوبی باشد، فقط مشکل اینجاست که چنین قانونی ظرف یک سال کل صنعت نرم‌افزار را ورشکست خواهد کرد!

۲-۱ سخت‌افزار شبکه

اکنون وقت آنست که توجه خود را از مسایل متفرقه به موضوع اصلی (یعنی همان شبکه‌های کامپیوتری) معطوف کنیم. هیچ طبقه‌بندی پذیرفته شده‌ای که در برگیرنده تمام انواع شبکه‌های کامپیوتری باشد، وجود ندارد، ولی در این میان می‌توان به دو عامل مهم توجه کرد: تکنولوژی انتقال و اندازه شبکه. اجازه دهید این دو را جداگانه بررسی کنیم. امروزه دو تکنولوژی انتقال بیش از همه گسترش یافته و فراگیر هستند:

۱. ارتباطات پخش (broadcast)

۲. ارتباطات همتا به همتا (peer-to-peer)

شبکه‌های پخش (broadcast network) دارای یک کانال مخابراتی هستند که بین همه کامپیوترهای شبکه به اشتراک گذاشته شده است. هر یک از کامپیوترها می‌توانند پیامهای خود را در بسته (packet) های کوچک مخابره کنند، و تمام کامپیوترهای دیگر این پیامها را دریافت خواهند کرد. آدرس کامپیوتری که این بسته در حقیقت برای وی ارسال شده، در بخشی از پیام نوشته می‌شود. هر کامپیوتری به محض دریافت بسته، آدرس گیرنده را چک می‌کند: اگر پیام برای او باشد، آنرا پردازش می‌کند؛ ولی اگر پیام متعلق به دیگری باشد، بسادگی آنرا نادیده می‌گیرد.

بعنوان مقایسه، فرض کنید کسی در انتهای راهرویی که در دو طرف آن پُر از اتاقهای متعدد است، فریاد بزند «آقای واتسون، بیایید. با شما کار دارم.» با اینکه این پیام به گوش همه افراد می‌رسد، فقط آقای واتسون به آن پاسخ می‌دهد و دیگران توجهی به آن نخواهد کرد. یا وقتی در سالن انتظار فرودگاه اعلام می‌شود که «مسافران پرواز ۶۴۴ به خروجی ۱۲ مراجعه کنند»، فقط آنهایی که بلیط این پرواز را دارند، عکس‌العمل نشان می‌دهند.

در شبکه‌های پخش با تعبیه یک کُد خاص در فیلد آدرس (address field) می‌توان یک پیام را به تمام

کامپیوترها ارسال کرد. چنین پیامی را همه کامپیوترها متعلق به خود تلقی کرده، و آنرا می خوانند. به این تکنیک پخش (broadcasting) گفته می شود. در برخی از سیستمهای پخش امکان ارسال پیام به دسته ای از کامپیوترها نیز وجود دارد، که به آن پخش گروهی (multicasting) می گویند. بدین منظور، معمولاً از یک بیت خاص در فیلد آدرس استفاده می شود، و همه آنهايي که این بیت در آنها وجود دارد عضو گروه محسوب شده و پیام را می گیرند. در شبکه های همتا به همتا (peer-to-peer network) بین تک تک کامپیوترها مسیر ارتباطی مستقل وجود دارد. البته وقتی یک بسته بخواهد از کامپیوتری به کامپیوتر دیگر برود، احتمالاً سر راه خود از چند ماشین بینابینی نیز عبور خواهد کرد. معمولاً در این قبیل شبکه ها مسیرهای متعددی بین دو کامپیوتر خاص می توان برقرار کرد، که از نظر طول مسیر با هم تفاوت دارند، و یافتن کوتاهترین مسیر یکی از مسایل مهم در این گونه شبکه ها است. بعنوان یک قاعده کلی (البته با استثنای متعدد)، شبکه های کوچک، متمرکز و محلی از نوع پخش هستند، و شبکه های بزرگ و گسترده از نوع همتا به همتا. به ارتباط همتا به همتا گاهی پخش تکي (unicasting) نیز گفته می شود. روش دیگر طبقه بندی شبکه ها اندازه شبکه است. در شکل ۱-۶ نوعی طبقه بندی بر اساس اندازه را مشاهده می کنید.

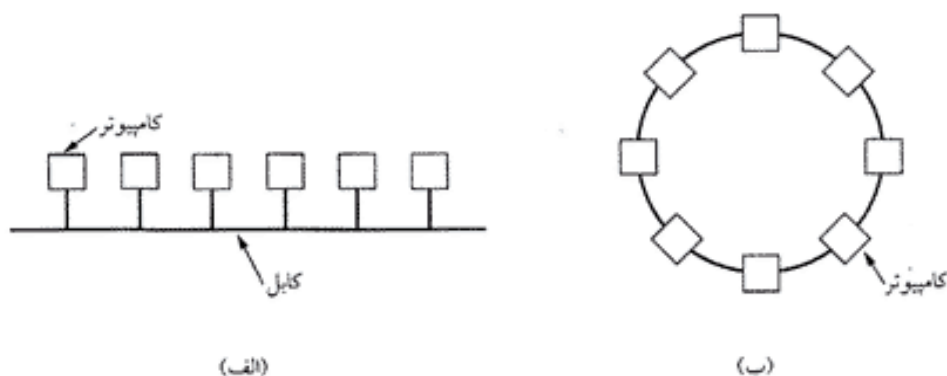
در بالا شبکه های شخصی (Personal Area Network) را می بینید (شبکه هایی که متعلق به یک فرد خاص هستند). ارتباط بیسیم بین ماوس، کی بورد، چاپگر، PDA و کامپیوتر از این نوع است. بعد از آن شبکه های محلی (LAN)، شهری (MAN) و گسترده (WAN) می آیند. در آخر هم شبکه های (شبکه ای که هر نقطه از آن خود یک شبکه کامل است) - و اینترنت معروفترین نمونه آن است - می آید. در این طبقه بندی فاصله کامپیوترها اهمیت زیادی دارد، چون تکنولوژی ارتباطی به شدت به آن وابسته است. در این کتاب درباره تمام این شبکه ها صحبت خواهیم کرد. در زیر هر یک از این شبکه ها را مختصراً معرفی می کنیم.

۱-۲-۱ شبکه های محلی (Local Area Network)

شبکه محلی، یا LAN، شبکه ایست خصوصی واقع در یک ساختمان یا مجتمع، که حداکثر ابعاد آن یکی دو کیلومتر باشد. از این نوع شبکه معمولاً برای متصل کردن کامپیوترهای یک شرکت و به اشتراک گذاشتن منابع (مانند چاپگر) یا مبادله اطلاعات استفاده می شود. یک شبکه LAN سه مشخصه اصلی دارد، که آنرا از سایر انواع شبکه متمایز می کند: (۱) اندازه، (۲) تکنولوژی انتقال اطلاعات، و (۳) توپولوژی (topology). اندازه LAN بسیار محدود است، بگونه ای که زمان انتقال سیگنالها در آن (حتی در بدترین شرایط) بسیار کم و از قبل قابل پیش بینی است. دانستن این محدودیتها برای طراحی شبکه بسیار مهم و اساسی است، و باعث ساده تر شدن مدیریت شبکه نیز می شود.

فاصله پردازنده ها	محل نسبی پردازنده ها	نمونه
1 m	روی یک میز	شبکه شخصی
10 m	یک اتاق	شبکه محلی
100 m	یک ساختمان	.
1 km	یک مجتمع	.
10 km	یک شهر	شبکه شهری
100 km	یک کشور	شبکه گسترده
1000 km	یک قاره	.
10,000 km	کره زمین	اینترنت

شکل ۱-۶. طبقه بندی شبکه ها بر اساس اندازه و فاصله پردازنده ها.



(الف)

(ب)

شکل ۷-۱. دو شبکه پخش. (الف) باس. (ب) حلقوی.

تکنولوژی انتقال اطلاعات در LAN معمولاً به کابل متکیست (و از این نظر بسیار شبیه شبکه های تلفن است). سرعت انتقال اطلاعات در LAN بین ۱۰ تا ۱۰۰ میلیون بیت در ثانیه (که با Mbps مشخص می شود)، تأخیر انتشار در آن کم (در حد میکرو یا نانو ثانیه)، و خطا در آن بسیار اندک است. LAN های جدیدتر به سرعت 10 Gbps نیز دست یافته اند. سرعت انتقال در شبکه معمولاً با واحد مگابیت بر ثانیه (1,000,000 bits/sec) یا گیگابیت بر ثانیه (1,000,000,000 bits/sec) اندازه گیری می شود.

توپولوژی های مختلفی برای شبکه های محلی پخش وجود دارد، که در شکل ۷-۱ دو تا از آنها را می بینید. در یک شبکه باس (bus network - شبکه ای با کابل کشی خطی) در هر لحظه فقط یکی از کامپیوترها مجاز به استفاده از خط و ارسال اطلاعات است، و تمام ماشینهای دیگر بایستی در این مدت از ارسال هر گونه اطلاعات خودداری کنند. در این قبیل شبکه ها بایستی مکانیزمی برای حل اختلاف (در مواقعی که دو کامپیوتر همزمان با هم شروع به ارسال می کنند) وجود داشته باشد. این مکانیزم می تواند متمرکز (centralized) یا توزیع شده (distributed) باشد. یکی از مکانیزمهای حل اختلاف در شبکه های باس پخش IEEE 802.3 نام دارد (که به اترنت - Ethernet - نیز معروف است)، و با کنترل غیرمتمرکز در سرعت های 10 Mbps تا 10 Gbps کار می کند. کامپیوترهای یک شبکه اترنت در هر زمانی می توانند اقدام به ارسال کنند، ولی اگر تصادفی بین آنها پیش آمد، هر یک از آنها مدتی (که بصورت تصادفی تعیین می شود) صبر کرده و دوباره سعی خواهد کرد.

نوع دیگری از شبکه های پخش، شبکه حلقوی (ring network) است. در یک شبکه حلقوی، هر بیت اطلاعات بصورت مستقل (و بدون اینکه بخواهد منتظر سایر بیت های بسته ای که به آن تعلق دارد، شود) در شبکه منتشر می شود. با توجه به سرعت بالای انتشار الکترونها در محیطهای رسانا، هر بیت حتی قبل از انتشار بیت های بعدی، می تواند بارها محیط شبکه را دور بزند. در این نوع شبکه هم بایستی مکانیزمی برای حل اختلاف بین کامپیوترهای متخاصم وجود داشته باشد. اغلب این مکانیزمها به نوعی نوبت بندی متکی هستند. یکی از این مکانیزمها IEEE 802.5 (یا IBM Token Ring) است، که در سرعت های 4 Mbps و 16 Mbps کار می کند. FDDI یکی دیگر از شبکه های پخش حلقوی است.

نوع دیگری از تقسیم بندی شبکه های پخش بر حسب نحوه اختصاص کانال است، و به استاتیک و دینامیک تقسیم می شود. در اختصاص کانال استاتیک هر کامپیوتر برای مدت زمانی محدود و مشخص کانال را در دست می گیرد، و فقط در این پُرش زمانیست که می تواند اطلاعات ارسال کند. در این روش پهنای باند کانال بشدت هدر می رود، چون بسیار پیش می آید که وقتی نوبت به یک کامپیوتر می رسد، چیزی برای گفتن ندارد. به همین دلیل، سیستمهای امروزی اغلب پهنای باند را بصورت دینامیک (بر حسب نیاز) تخصیص می دهند.

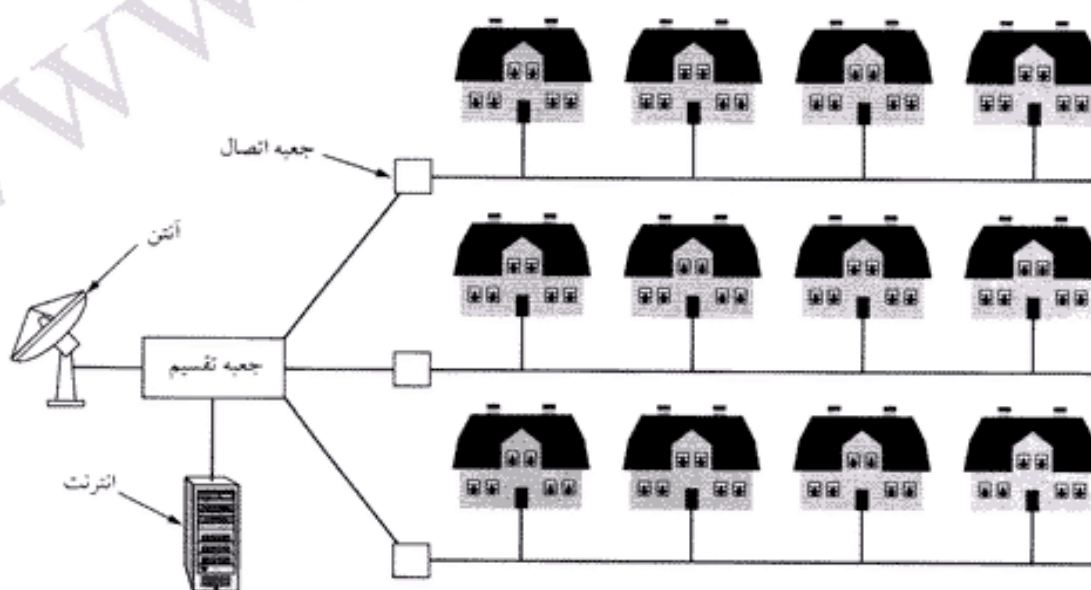
تخصیص دینامیک کانال خود بر دو نوع است: متمرکز و غیرمتمرکز. در نوع متمرکز یک موجودیت مشخص (بنام واحد تصمیم گیرنده - Arbitration Unit) وجود دارد، که درخواستها را دریافت کرده، و بر اساس نوعی الگوریتم داخلی نوبتها را تعیین می کند. در تخصیص کانال غیرمتمرکز این موجودیت تصمیم گیرنده وجود ندارد، و تصمیم گیری بر عهده تک تک کامپیوترهاست. شاید فکر کنید این روش جز هرج و مرج نتیجه ای ندارد، ولی چنین نیست (در آینده خواهید دید که الگوریتمهایی وجود دارند که می توانند به اوضاع سر و سامان بدهند).

۲-۲-۱ شبکه های شهری (Metropolitan Area Network)

شبکه شهری، یا MAN، شبکه ایست که یک شهر را پوشش می دهد. شبکه های تلویزیون کابلی بهترین نمونه MAN هستند. اولین شبکه های تلویزیون کابلی در نقاط کور شهرها راه اندازی شدند، بدین ترتیب که یک آنتن مرکزی و بزرگ در محلی که فرستنده اصلی را می دید نصب، و از این آنتن کابلهایی به مشترکان محروم از برنامه های تلویزیونی کشیده می شد.

در ابتدا این سیستمها بطور اختصاصی برای هر محل ساخته می شد، ولی بزودی شرکتهای بزرگ بوی پول را از آن استشمام کردند، و با کسب اجازه دولت تمام شهر را زیر پوشش کابلهای خود بردند. این شبکه ها برای پخش برنامه هم برنامه ریزی خاصی دارند، مثلاً یک شبکه فقط اخبار پخش می کند، دیگری فقط برنامه های ورزشی دارد، و آن یکی فقط آشپزی. این شبکه ها بسیار تخصصی بودند، و تا اواخر دهه ۱۹۹۰ فقط برنامه های تلویزیونی پخش می کردند.

با شروع گرایش عمومی به اینترنت، گردانندگان این شبکه ها بزودی دریافتند که با تغییری مختصر در سیستمهای خود می توانند از قسمتهای بالاستفاده پهنای باند برای ارائه سرویسهای دوطرفه اینترنت بهره ببرند. این لحظه بود که شبکه های تلویزیون کابلی تبدیل به شبکه های شهری (MAN) شدند. در شکل ۸-۱-۸ نمای تقریبی از یک شبکه شهری را ملاحظه می کنید. در این شکل می بینید که ابتدا سیگنالهای تلویزیونی و اینترنتی ترکیب شده و به یک مرکز فوق-توزیع (head end) می روند، تا از آنجا در خانه های مشترکان توزیع شوند. (در فصل ۲ باز هم به این مبحث خواهیم پرداخت).



شکل ۸-۱-۸ یک شبکه شهری مبتنی بر تلویزیون کابلی.

تلویزیون کابلی تنها مثال زنده MAN نیست. اخیراً تحقیقاتی بر روی اینترنت بیسیم پرسرعت (high-speed wireless Internet) انجام شده، که نتیجه آن نوع دیگری از MAN خواهد بود. با این استاندارد که IEEE 802.16 نام دارد، در فصل ۲ بیشتر آشنا خواهید شد.

۳-۲-۱ شبکه‌های گسترده (Wide Area Network)

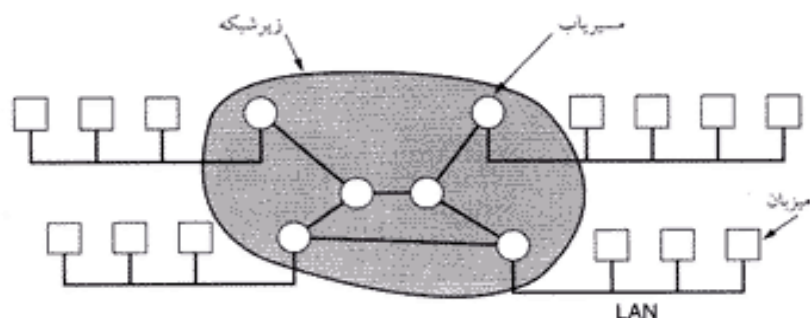
شبکه گسترده، یا WAN، گستره جغرافیایی بزرگی (مانند یک کشور یا قاره) دارد. در این نوع شبکه کامپیوترهایی هستند که برنامه‌های کاربردی روی آنها اجرا می‌شود، و معمولاً به آنها میزبان (host) می‌گویند. این کامپیوترها توسط زیرشبکه‌های مخابراتی (communication subnet) - یا بطور مختصر، زیرشبکه - به هم متصل می‌شوند. میزبانها متعلق به افراد هستند، در حالیکه زیرشبکه اغلب به شرکتهای مخابرات تعلق دارد. وظیفه زیرشبکه انتقال پیام از یک میزبان به میزبان دیگر است. جدا کردن این دو بخش (میزبانها و زیرشبکه) طراحی شبکه‌های WAN را تا حد زیادی ساده می‌کند.

در اغلب شبکه‌های گسترده، زیرشبکه از دو بخش مجزا تشکیل می‌شود: خطوط انتقال (transmission lines) و تجهیزات سوئیچینگ (switching elements). خطوط انتقال وظیفه رد و بدل کردن اطلاعات را بر عهده دارند، و می‌توان برای ایجاد آنها از سیم مسی، فیبر نوری یا حتی امواج رادیویی استفاده کرد. تجهیزات سوئیچینگ کامپیوترهای خاصی هستند که ارتباط بین خطوط انتقال را برقرار می‌کنند. وقتی داده‌ها از یک خط وارد می‌شود، این کامپیوتر باید مسیر خروجی آنها را مشخص کند. این کامپیوترهای سوئیچینگ به نامهای مختلفی خوانده می‌شوند، که می‌توان از معروفترین آنها به مسیریاب (router) اشاره کرد.

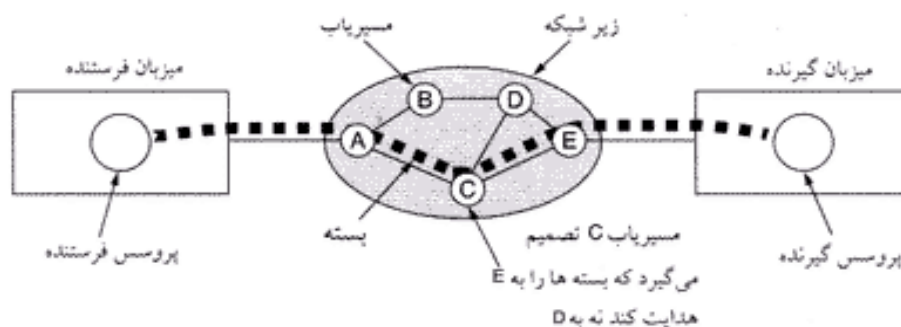
در این مدل (شکل ۱-۹) معمولاً هر کامپیوتر میزبان در یک شبکه محلی قرار دارد که از طریق یک مسیریاب به قسمتهای دیگر متصل می‌شود (البته در مواردی میزبان می‌تواند مستقیماً نیز به مسیریاب وصل باشد). به مجموعه خطوط مخابراتی و مسیریاب‌ها (منهای کامپیوترهای میزبان) زیرشبکه گفته می‌شود.

معنای اولیه زیرشبکه همان است که در بالا گفته شد، یعنی مجموعه خطوط مخابراتی و مسیریاب‌ها که وظیفه آنها انتقال اطلاعات از یک میزبان به میزبان دیگر است. اما سالها بعد از این اصطلاح در ارتباط با آدرس دهی شبکه‌ها نیز استفاده شد (فصل ۵ را ببینید). متأسفانه هنوز اصطلاح مناسبی برای کاربرد اولیه آن پیدا نشده است، و ما هم (با کمی شک و تردید) آنرا در هر دو مورد بکار خواهیم برد، که با توجه به موضوع بحث می‌توان معنای موردنظر را استنباط کرد.

در بسیاری از WANها تعداد زیادی خطوط انتقال وجود دارد، که هر کدام یک جفت مسیریاب را به هم وصل می‌کنند. اگر دو مسیریاب که اتصال فیزیکی مستقیم ندارند، بخواهند با یکدیگر ارتباط برقرار کنند، باید این کار را بصورت غیرمستقیم (از طریق مسیریاب‌های دیگر) انجام دهند. وقتی یک بسته داده در مسیر خود (از مسیریاب



شکل ۱-۹. ارتباط بین کامپیوترهای میزبان و LANها در یک زیرشبکه.



شکل ۱-۱۰. استریم (جریان) بسته‌ها از مبدأ به مقصد.

مبدأ به مسیریاب مقصد) از چند مسیریاب بینابینی عبور می‌کند، ابتدا بصورت کامل دریافت و ذخیره شده، و پس از آزاد شدن خط خروجی به سمت مقصد فرستاده می‌شود. زیر شبکه‌هایی که بر اساس این قاعده عمل می‌کنند، به زیر شبکه ذخیره-ارسال (store-and-forward) یا سوئیچ بسته (packet-switched) معروفند. تقریباً تمامی شبکه‌های WAN (بجز شبکه‌های ماهواره‌ای) از این نوع هستند. اگر اندازه بسته‌ها کوچک و یکسان باشد، به آنها سلول (cell) نیز گفته می‌شود.

به دلیل اهمیت مفهوم زیر شبکه سوئیچ بسته، لازمست کمی بیشتر درباره آن توضیح دهیم. وقتی پروسی در یک میزبان می‌خواهد پیامی به میزبان دیگر بفرستد، ابتدا آنرا به بسته‌های کوچکتر (که پشت سر هم شماره‌گذاری می‌شوند) تقسیم می‌کند. این بسته‌ها بصورت مستقل به طرف مقابل ارسال می‌شوند، و بعد از رسیدن تمامی آنها به مقصد، در آنجا دوباره به یکدیگر مونتاژ شده و پیام اصلی را می‌سازند (شکل ۱-۱۰ را ببینید).

در این شکل تمام بسته‌ها از طریق مسیر ACE به مقصد رسیده‌اند (در حالیکه مسیرهای ABDE و ACDE نیز وجود داشت). در برخی از شبکه‌ها این یک الزام است، یعنی تمام بسته‌های یک پیام باید از یک مسیر عبور کنند، در حالیکه در شبکه‌های دیگر این بسته‌ها می‌توانند از مسیرهای مختلف عبور کنند. البته، اگر مسیری بهترین مسیر ممکن باشد (مانند ACE در اینجا)، همه بسته‌ها از آن مسیر عبور خواهند کرد، حتی اگر شبکه چنین الزامی را تحمیل نکرده باشد.

تصمیم‌گیری درباره مسیر ارسال بسته‌ها امری داخلی است، یعنی هر مسیریاب خود درباره آن تصمیم می‌گیرد. وقتی یک بسته به مسیریاب A می‌رسد، این مسیریاب A است که تصمیم می‌گیرد آنرا از طریق خط متصل به B بفرستد یا از خط متصل به C. مسیریاب‌ها برای تصمیم‌گیری درباره مسیر بسته‌ها از الگوریتمهای مسیریابی (routing algorithm) استفاده می‌کنند، که درباره آنها در فصل ۵ صحبت خواهیم کرد.

تمام شبکه‌های WAN از نوع سوئیچ بسته نیستند، مانند سیستمهای ماهواره‌ای. در این سیستمها هر روتر آنتنی دارد که از طریق آن اطلاعات را به ماهواره می‌فرستد، یا اطلاعات ارسالی آن را دریافت می‌کند. تمام مسیریاب‌های این مجموعه می‌توانند به ماهواره گوش کنند (و حتی برخی از آنها به اطلاعات ارسالی از مسیریاب‌های همسایه نیز گوش می‌کنند). البته شبکه‌هایی هم وجود دارد که فقط برخی از مسیریاب‌های آن (و نه همه آنها) ارتباط ماهواره‌ای دارند. شبکه‌های ماهواره‌ای ذاتاً از نوع پخش هستند، و اغلب در جاهایی بکار می‌روند که این طریقه پخش اهمیت داشته باشد.

۴-۲-۱ شبکه‌های بیسیم (Wireless Network)

مخابرات دیجیتال بیسیم ایده جدیدی نیست. گد موری که فیزیکدان ایتالیایی گالیلیمو مارکونی در سال ۱۹۰۱ از یک کشتی به ساحل مخابره کرد، را می‌توان اولین پیام دیجیتال بیسیم محسوب کرد. سیستمهای جدید مخابرات

بیسیم فقط کارایی بهتری دارند، اما ایده اصلی در واقع همان است. در ساده ترین صورت، شبکه های بیسیم را می توان به سه دسته بزرگ تقسیم کرد:

۱. ارتباطات بین سیستمی
۲. LAN های بیسیم
۳. WAN های بیسیم

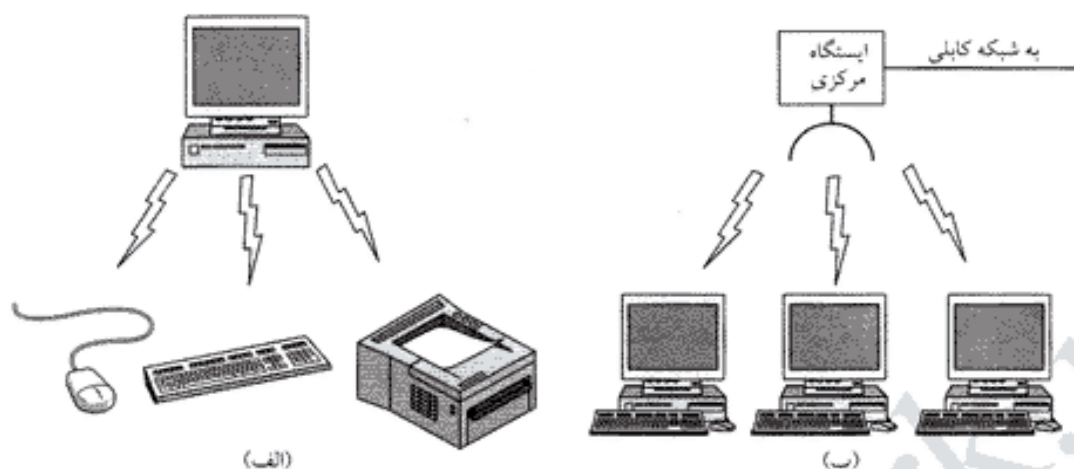
ارتباطات بین سیستمی (system interconnection) یعنی برقراری ارتباط بین قطعات داخلی یک کامپیوتر با استفاده از امواج رادیویی کوتاه بُرد. تقریباً هر کامپیوتری یک مانیاتور، صفحه کلید یا ماوس دارد که معمولاً با کابل به آن متصل می شوند. برای بسیاری از کاربران خانگی (و حتی اداری) وصل کردن این کابلها (با اینکه آنها طوری طراحی شده اند که نتوان هیچکدام را به دیگری وصل کرد) یک کار شاق است، و برای این کار دست به دامان تکنسین های کامپیوتر می شوند. به همین علت، برخی از شرکتهای سازنده کامپیوتر دور هم جمع شدند، و یک شبکه بیسیم با بُرد کوتاه بنام بلوتوث (Bluetooth) اختراع کردند که این قطعات را بدون استفاده از سیم به کامپیوتر متصل می کند. تکنولوژی بلوتوث اجازه می دهد تا دستگاههایی مانند چاپگر، دوربین دیجیتال، گوشی، و اسکتر نیز (با قرار گرفتن در بُرد امواج آن) به کامپیوتر متصل شوند. برای این کار به هیچ اتصال فیزیکی یا حتی نصب درایور نیاز نیست، و فقط کافیست دستگاه را روشن کرده و در بُرد کامپیوتر قرار دهید، تا کار کند. برای بسیاری از کاربران این یک مزیت خارق العاده است.

ارتباطات بین سیستمی اساساً بر الگوی اصلی-پیرو (master-slave) مبتنی است (شکل ۱-۱۱ الف). در این سیستم، کامپیوتر اصلی است و با وسایل جانبی بعنوان رعایای خود صحبت می کند. این کامپیوتر اصلی است که به رعایا می گوید از چه آدرسی استفاده کنند، کی حرف بزنند، چه مدت حرف بزنند، روی چه فرکانسی صحبت کنند، و مانند آن. در باره تکنولوژی بلوتوث در فصل ۴ مفصلاً صحبت خواهیم کرد.

نوع دیگر ارتباطات بیسیم، شبکه محلی بیسیم (یا LAN بیسیم) است. در این سیستم هر کامپیوتر یک مودم رادیویی و یک آنتن دارد، که به وسیله آن با کامپیوترهای دیگر ارتباط برقرار می کند. در اغلب این سیستمها یک آنتن مرکزی روی پشت بام وجود دارد (شکل ۱-۱۱ ب)، که ارتباط بین کامپیوترها را تسهیل می کند، اما اگر شبکه باندازه کافی کوچک باشد، آنها می توانند مستقیماً با هم حرف بزنند. این نوع شبکه در دفاتر کوچک، خانه ها و جاهایی که کابل کشی مشکل است، بسرعت در حال گسترش است. مهمترین استاندارد LAN های بیسیم IEEE 802.11 نام دارد، که در اغلب سیستمها از آن استفاده می شود.

نوع سوم ارتباطات بیسیم، سیستمهای WAN بیسیم است. شبکه رادیویی بکار رفته در سیستمهای تلفن همراه از این نوع است. این سیستمها اکنون نسل سوم خود را پشت سر می گذارند. نسل اول آنالوگ بود و فقط برای صدا از آن استفاده می شد. نسل دوم با اینکه دیجیتال شده بود، ولی باز هم فقط از صدا پشتیبانی می کرد. نسل سوم نیز دیجیتال است، و اینک همزمان از صدا و دیتا پشتیبانی می کند. WAN های بیسیم اساساً تفاوتی با LAN بیسیم ندارند، و فقط بُرد آنها بیشتر و البته نرخ انتقال داده ها کمتر است. LAN های بیسیم می توانند داده ها را با سرعتهایی در حد 50 Mbps (در محدوده چند ده متر) منتقل کنند. نرخ انتقال داده ها در WAN های بیسیم بزحمت به 1 Mbps می رسد، ولی بُرد آنها به جای متر با کیلومتر سنجیده می شود. در فصل ۲ درباره این سیستمها بسیار خواهیم گفت.

علاوه بر این شبکه های کم سرعت، اکنون WAN های بیسیم پُر ظرفیت نیز در دست توسعه است. دسترسی پُر سرعت به اینترنت از منزل و دفتر کار بدون استفاده از خطوط تلفن، از اولین کاربردهای این شبکه هاست. استاندارد این سیستم (که به آن سرویس توزیع چند نقطه ای محلی گفته می شود) IEEE 802.16 نام دارد، که درباره آن در فصل ۴ بیشتر صحبت خواهیم کرد.



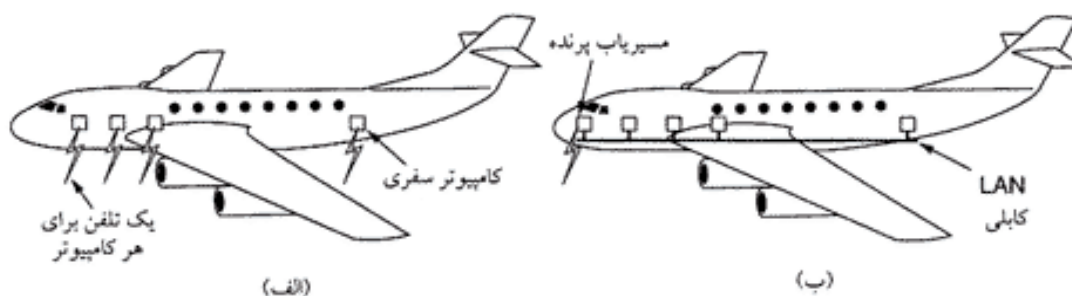
شکل ۱-۱۱. (الف) پیکربندی بلوتوث. (ب) LAN بیسیم.

تقریباً تمام شبکه‌های بیسیم باید در جایی به یک شبکه معمولی متصل شوند، تا بتوانند وظایف خود را انجام دهند؛ این کار را به روشهای مختلف می‌توان انجام داد. برای مثال، در شکل ۱-۱۲ الف هواپیمایی را می‌بینید، که در آن تعدادی از مسافران مودمهای خود را به تلفنهای تعبیه شده در صندلی‌های هواپیما وصل کرده‌اند؛ این تماسها بکلی از هم مستقلند. اما در شکل ۱-۱۲ ب روش بهینه‌تری را ملاحظه می‌کنید، که در آن هر صندلی یک اتصال اینترنت دارد، و مجموعه آنها تشکیل یک LAN معمولی می‌دهند؛ و این LAN به یک مسیر یاب بیسیم وصل است که ارتباط آنها با دنیای خارج برقرار می‌کند.

بسیاری افراد فکر می‌کنند که بیسیم موج آینده است (برای مثال، Bi et al., 2001؛ Leeper, 2001؛ Varshey and Vetter, 2000)، ولی حداقل یک صدای مخالف هم وجود دارد. باب متکالف، مخترع اینترنت، می‌گوید: «کامپیوترهای موبایل بیسیم مثل دستشویی‌های متحرک هستند - آنها فقط بدرد بیک نیک، کارگاه‌های ساختمانی و اردوهای کوتاه مدت می‌خورند. نصیحت من اینست که در خانه خود کابل اینترنت بکشید، و منتظر آینده بمانید» (Metcalf, 1995). شاید تاریخ بعدها این اظهار نظر را در کنار جمله تاریخی تی جی واتسون رئیس IBM در سال ۱۹۴۵ بگذارد، که در پاسخ اینکه چرا IBM وارد بازار کامپیوتر نمی‌شود، گفته بود: «دنیا تا سال ۲۰۰۰ به چهار یا پنج کامپیوتر بیشتر نیاز نخواهد داشت.»

۵-۲-۱ شبکه‌های خانگی (Home Network)

از هم اکنون می‌توان شبکه‌های خانگی را در افق آینده دید. ایده اصلی آن است که تمام یک وسایل یک خانه بتوانند



شکل ۱-۱۲. (الف) کامپیوترهای سفری منفرد. (ب) یک LAN پرنده.

با یکدیگر ارتباط برقرار کنند، و بتوان آنها را از طریق اینترنت کنترل کرد. این یکی از آن چیزهایی است که هیچکس منتظر آن نبوده (مانند، کنترل از راه دور تلویزیون و تلفن همراه)، ولی وقتی آمد دیگر هیچکس نمی تواند زندگی بدون آن را تصور کند.

وسایل زیادی را می توان در یک شبکه به هم متصل کرد، که از واضحترین آنها می توان به موارد زیر اشاره کرد:

۱. کامپیوترها (رومیزی، سفری، PDA، وسایل جانبی)
۲. وسایل سرگرمی (تلویزیون، DVD، ویدئو، دوربین دیجیتال، استریو، MP3)
۳. وسایل مخابراتی (تلفن معمولی و همراه، فکس، دستگاههای ارتباط داخلی)
۴. لوازم خانگی (میکروویو، یخچال، ساعت، بخاری، تهویه مطبوع، چراغ)
۵. وسایل اندازه گیری از راه دور (آلارم دود یا دزدی، قرائت کنتور، ترموستات، دوربین اتاق بچه)

شبکه های خانگی به بسیاری از خانه ها راه یافته است؛ در این خانه ها وسایلی وجود دارد که یک ارتباط پرسرعت اینترنت را بین چند کامپیوتر به اشتراک می گذارند. با گسترش سرگرمی ها روی اینترنت، بزودی شاهد تلویزیونها و استریوهایی خواهیم بود که مستقیماً به اینترنت متصلند (و این ارتباط دو جانبه خواهد بود، چرا که شاید شما هم مایل باشید فیلمها و موزیکهای خود را با دوستان و آشنایان به اشتراک بگذارید). مخابرات بین المللی از هم اکنون یک کالای در دسترس است، ولی بزودی این سرویسها بصورت دیجیتالی و از طریق اینترنت ارائه خواهند شد. امروزه کمتر خانه ای را می توان یافت که یک دوچین ساعت نداشته باشد، و با آمدن بهار و پائیز صاحبخانه مجبور است آنها را دستی جلو یا عقب بکشد؛ اگر تمام این ساعتها به اینترنت متصل باشند، می توان آنها را بصورت خودکار تنظیم کرد. کنترل خانه و مشاهده اتفاقاتی که در غیبت ما در آن می گذرد، یکی از آرزوهای دیرینه ماست، که اینک به واقعیت تبدیل شده است. (دیگر می توانید با خیال راحت به سینما بروید، و بچه ها در خانه تنها بگذارید!) شاید فکر کنید هر یک از این کاربردها به شبکه ای مجزا نیاز دارد، اما یکپارچه کردن آنها احتمالاً ایده بهتری است.

شبکه های خانگی تفاوت های ذاتی با سایر انواع شبکه دارد. اول اینکه نصب آن نباید پیچیده باشد. آنهایی که در این سالها درگیر کار نصب شبکه بوده اند، با جوابهای زیر (وقتی یک مشتری با مشکلی مواجه شده و به شما تلفن می زند) کاملاً آشنا هستند: (۱) دفترچه راهنما را با دقت بخوانید، (۲) کامپیوتر را دوباره بوت کنید، (۳) تمام سخت افزارها و نرم افزارهای اضافی - آنهایی که مال شرکت ما نیست! - را حذف کنید، (۴) جدیدترین درایور را از سایت وب ما بردارید و نصب کنید، و بالاخره وقتی هیچکدام از این کارها فایده ای نبخشد، (۵) کامپیوتر را فرمت و ویندوز را از نو نصب کنید. اگر به کسی که یک یخچال اینترنتی خریده، بگوئید آخرین ویرایش سیستم عامل یخچال اینترنتی را بار کرده و نصب کند، مسلماً باعث خوشحالی وی نخواهد شد! آنهایی که کامپیوتر می خرند، عادت دارند با سیستمهایی سر و کله بزنند که کار نمی کند، ولی خریداران اتومبیل، تلویزیون و یخچال این حرفها سرشان نمی شود؛ آنها انتظار دارند وسیله ای که خریده اند از اول بسماله بدون هیچ مشکلی کار کند.

دوم اینکه شبکه های خانگی باید بتوانند تحت هر شرایطی کار کنند. حتی همین حالا هم یک دستگاه کولر گازی با فقط چهار دکمه OFF، LOW، MEDIUM و HIGH، یک دفترچه راهنمای ۳۰ صفحه ای دارد. تصورش را بکنید وقتی این کولر اینترنتی شود، فقط قسمت امنیت آن دستکم ۳۰ صفحه خواهد بود! و این از حد تحمل قسمت اعظم کاربران این قبیل دستگاهها خارج است.

قیمت پائین سومین عامل برای موفقیت شبکه های خانگی است. خیلی از افراد حاضر نیستند ۵۰ دلار پول اضافه برای یک ترموستات اینترنتی بدهند، که چي، فقط درجه حرارت خانه خود را از اداره چک کنند! حتی ۵ دلار را هم شاید خیلی ها بزور بدهند.

از آنجائیکه دنیای آینده دنیای چندرسانه‌ایست، عامل چهارم در مقبولیت شبکه‌های خانگی بالا بودن پهنای باند آنهاست. مطمئن باشید هیچکس یک تلویزیون اینترنتی که فیلمها را با وضوح ۳۲۰×۲۴۰ پیکسل و ۱۰ فریم در ثانیه نشان می‌دهد، نخواهد خرید. حتی اینترنت (که در اغلب ادارات از آن استفاده می‌شود) نیز مناسب چندرسانه‌ای نیست. شبکه‌های خانگی برای آن که بتوانند فراگیر شوند، باید پهنای باند بیشتر را با قیمتی کمتر ارائه کنند.

پنجم، امکان گسترش شبکه‌های خانگی است. این به معنای آن است که جنگ استانداردها باید یک بار و برای همیشه حل شود. اگر امروز به مشتریان خود توصیه کنید وسایل IEEE 1394 (که به فایر وایر - FireWire - معروف است) را بخرند، و سال آینده بگویند امسال مد USB 2.0 است، مطمئن باشید همه آنها را پرانده‌اید. استانداردهای ارتباطی باید سالها (و کابل کشی، دهها سال) ثابت و بدون تغییر باقی بماند.

امنیت و قابلیت اعتماد ششمین عامل موفقیت شبکه‌های خانگی است. اینکه چند تا ایمیل و پروسی از دست بدهید یک چیز است، و به باد رفتن تمام خانه در اثر لو رفتن کدهای امنیتی آن، یک چیز دیگر.

در این میان یکی از سوالات جالب اینست که شبکه‌های خانگی باید بیسیم باشند یا باسیم. همین حالا هم در اغلب خانه‌ها شش شبکه وجود دارد: شبکه برق، تلفن، تلویزیون کابلی، آب، گاز و فاضلاب. اضافه کردن شبکه هفتم به خانه‌های نوساز هزینه چندانی ندارد، ولی در خانه‌های موجود کاری پُرخرج خواهد بود. شبکه‌های بیسیم کم خرج تر هستند، ولی شبکه‌های کابلی از نظر ایمنی بسیار قابل اعتمادترند. امواج رادیویی بسادگی از محدوده خانه خارج می‌شوند، و شاید یک همسایه فضول بتواند ایمیل‌های شما را بخواند، و یا دزدکی از اشتراک اینترنت شما استفاده کند. در فصل ۸ درباره روشهای رمزنگاری برای مقابله با این مشکلات صحبت خواهیم کرد، ولی در شبکه‌های خانگی موضوع امنیت از اهمیت چندگانه‌ای برخوردار است.

بطور خلاصه، شبکه‌های خانگی علاوه بر امکاناتی که ارائه می‌کنند، چالشهایی را نیز با خود به همراه دارند. اغلب این چالشها به مدیریت ساده، قابلیت اعتماد، ایمنی بالا (بویژه در مورد کاربران غیرحرفه‌ای)، کارایی بالا و قیمت پائین مربوط می‌شوند.

۶-۲-۱ شبکه شبکه‌ها (Internetwork)

شبکه‌های متعددی با نرم‌افزارها و سخت‌افزارهای بسیار مختلف در سراسر دنیا وجود دارد، و بسیار پیش می‌آید که کاربری از یک شبکه بخواهد با کاربران شبکه‌های دیگر ارتباط برقرار کند. برای انجام این خواسته بایستی شبکه‌های مختلف (که بعضاً با هم ناسازگار هم هستند) با وسایلی بنام دروازه gateway - که می‌تواند سخت‌افزاری یا نرم‌افزاری باشد - به هم متصل شده، و داده‌ها از فرمتی به فرمت دیگر تبدیل شود. به مجموعه‌ای از این شبکه‌های بهم پیوسته شبکه شبکه‌ها (internetwork یا internet) گفته می‌شود. کلمه internet وقتی با i نوشته می‌شود، معنای عام می‌دهد، ولی با I همان شبکه جهانی اینترنت از آن مستفاد می‌شود.

متداولترین شکل شبکه شبکه‌ها عبارتست از تعدادی LAN که با ارتباطات WAN به هم متصل شده‌اند. در حقیقت، اگر در شکل ۹-۱ WAN را به جای subnet (زیرشبکه) قرار دهیم، هیچ چیز تغییر نخواهد کرد. در این مورد تنها تفاوت تکنیکی بین WAN و subnet وجود یا عدم وجود کامپیوترهای میزبان (host) است: اگر سیستم تاحیه خاکستری فقط از مسیر یاب تشکیل شده باشد، این یک زیرشبکه است؛ اگر در آن میزبان هم وجود داشته باشد، یک WAN است. تفاوت واقعی در مالکیت و طرز استفاده است.

اغلب افراد مفاهیم زیرشبکه (subnet)، شبکه (network) و شبکه شبکه‌ها (internetwork) را با هم اشتباه می‌کنند. زیرشبکه (که بیشتر در شبکه‌های گسترده مفهوم پیدا می‌کند) مجموعه‌ایست از مسیر یاب‌ها و خطوط مخابراتی متعلق به راهبر شبکه. برای مقایسه، شبکه تلفن شهری از یک سری مراکز سونیچینگ، خطوط مخابراتی

پُرسرعت، و خطوط کم سرعت که مشترکان را به مرکز وصل می کند، تشکیل می شود. خطوط پُرسرعت و مراکز سونیچینگ (که متعلق به شرکت مخابرات هستند) همان زیرشبکه را می سازند. تلفنهای مشترکان که متعلق به افراد (و معادل میزبان) است، جزئی از زیرشبکه نیستند. مجموعه این دو (زیرشبکه و میزبانها) شبکه را می سازد. در شبکه های LAN (که فقط کامپیوتر و کابل است) زیرشبکه وجود ندارد.

وقتی چند تا از این شبکه ها به هم متصل می شوند، یک شبکه شبکه ها شکل می گیرد. با این تعریف، مجموعه دو LAN، یا یک LAN و یک WAN، را هم می توان شبکه شبکه ها نامید، ولی در صنعت کامپیوتر توافق یکپارچه ای درباره این اصطلاحات وجود ندارد. بعنوان یک قاعده کلی، اگر چند شرکت یا سازمان مختلف هر یک (به هزینه خود) قسمتی از یک شبکه را بنا کنند، با یک شبکه شبکه ها سروکار داریم. همچنین اگر تکنولوژی زیربنایی در قسمتهای مختلف متفاوت باشد، به احتمال زیاد دو شبکه داریم.

۳-۱ نرم افزار شبکه

در اولین شبکه های کامپیوتری سخت افزار از اهمیت ویژه ای برخوردار بود، و به نرم افزار فقط بعنوان چیزی که باید بعداً به آن فکر می شد، نگاه می کردند. اما این استراتژی دیگر کارایی ندارد. امروزه نرم افزار شبکه بسیار ساخت یافته است، که در این قسمت آنرا بررسی می کنیم. روشهای مورد بحث در این قسمت سنگ بنای کتاب را تشکیل می دهند، و در آینده بسیار به آنها مراجعه خواهیم کرد.

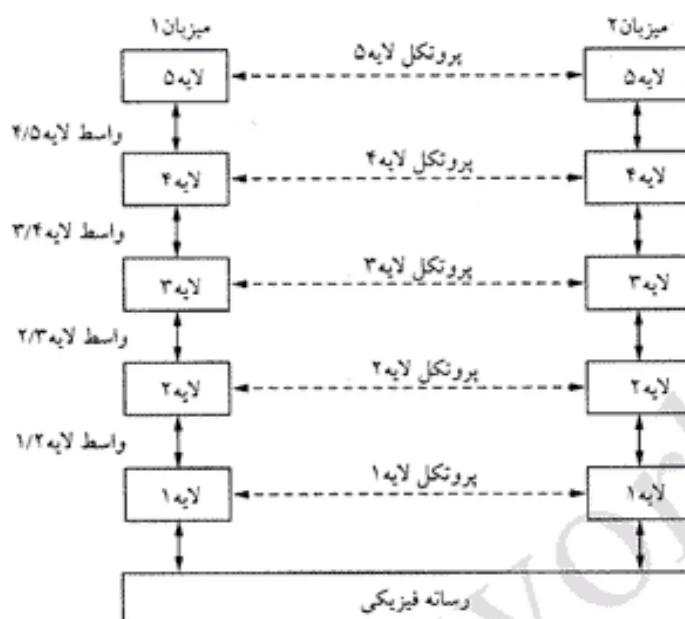
۱-۳-۱ سلسله مراتب پروتکل ها

برای کاهش پیچیدگیهای طراحی، اغلب شبکه ها بصورت مجموعه ای از چند لایه (layer) یا سطح (level) - که هر کدام روی دیگری قرار می گیرند - طراحی می شوند. تعداد لایه ها، نام هر لایه، محتوای آن، و کاری که هر لایه انجام می دهد، از شبکه ای به شبکه دیگر متفاوت است. وظیفه هر لایه ارائه سرویسهای خاص به لایه های بالاتر، و پنهان کردن جزئیات کار از دید آنهاست. در این مفهوم، هر لایه یک ماشین مجازی (virtual machine) است که سرویسهای خاصی را در اختیار لایه های بالاتر می گذارد.

این یکی از مفاهیم آشنا و کلیدی در کلیه علوم کامپیوتری است، که با نامهایی از قبیل پنهان کردن اطلاعات (information hiding)، انسواع داده مجرد (abstract data types)، کسپولی کردن داده ها (data encapsulation) و برنامه نویسی شیءگرا (object-oriented programming) شناخته می شود. ایده اصلی این است که یک قطعه نرم افزار (یا سخت افزار) سرویسی را به کاربران خود عرضه کند، ولی جزئیات کار (از قبیل حالت داخلی خود و الگوریتمهای بکار رفته) را از آنها مخفی نگه دارد.

لایه n یک ماشین همیشه با لایه n ماشین دیگر حرف می زند. قواعد و قراردادهای این ارتباط را پروتکل لایه n (layer n protocol) می نامند. در ساده ترین حالت، پروتکل (protocol) عبارتست از قراردادهای توافق شده بین دو طرف برای برقراری و پیشبرد یک ارتباط. بعنوان مقایسه، وقتی یک خانم به یک آقا معرفی می شود، آن خانم می تواند دستش را پیش بیاورد؛ و بنوبه خود آن آقا هم می تواند دست خانم را (اگر یک شاهزاده اروپایی در یک میهمانی رسمی باشد) ببوسد، یا فقط با او دست بدهد (اگر آن خانم یک وکیل امریکایی در یک جلسه کاری باشد). سرپیچی از پروتکل ها برقراری ارتباط را بسیار دشوار (اگر نگوئیم، غیرممکن) خواهد کرد.

در شکل ۱-۱۳ یک شبکه پنج لایه به تصویر کشیده شده است. به اجزایی که در یک لایه هستند، همتا (peer) گفته می شود. این همتاها می توانند پروسس های نرم افزاری، وسایل سخت افزاری، و یا حتی دو انسان باشند. عبارت دیگر، این همتاها هستند که با استفاده از پروتکل با هم رابطه برقرار می کنند.



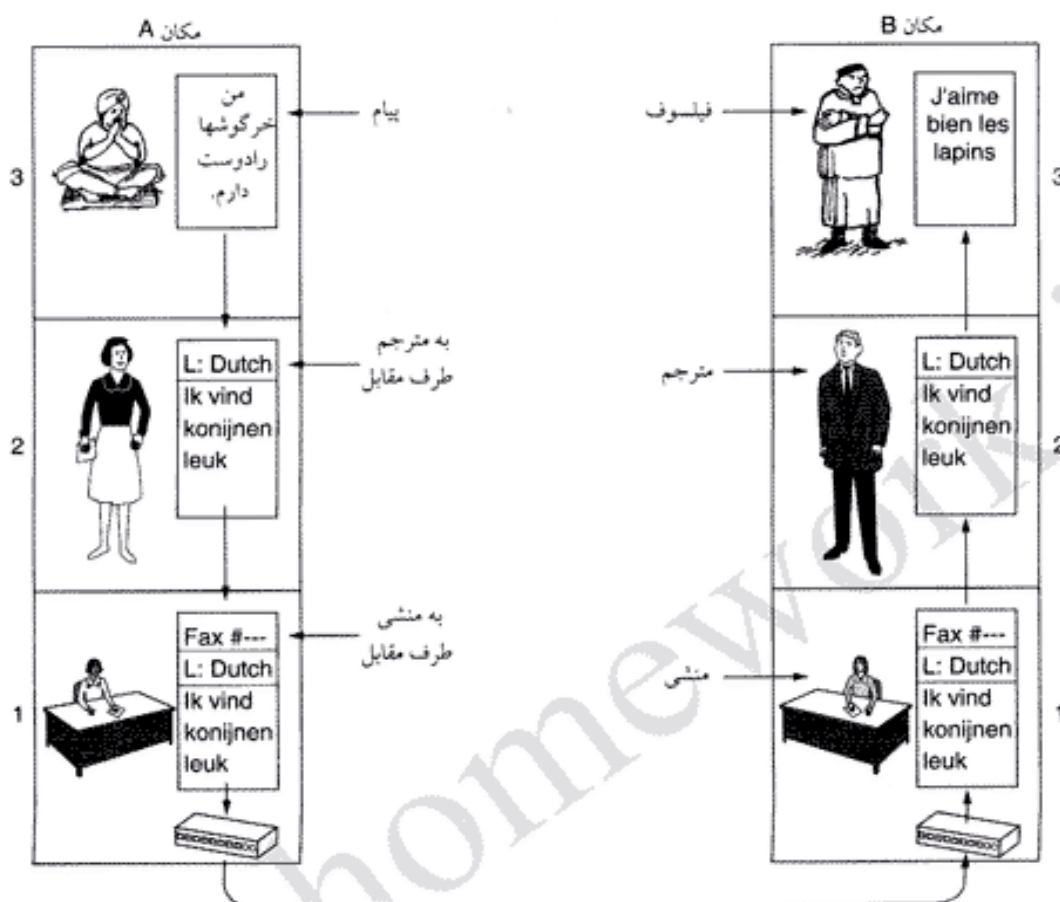
شکل ۱-۱۳. لایه‌ها، پروتکل‌ها، واسط‌ها.

در حقیقت، داده‌ها هرگز مستقیماً از لایه n یک ماشین به لایه n ماشین دیگر منتقل نمی‌شوند. بلکه، هر لایه داده‌ها (و اطلاعات کنترلی) را به لایه زیرین خود می‌دهد، تا به پائین‌ترین لایه برسد. در زیر پائین‌ترین لایه (لایه 1) رسانه فیزیکی (physical medium) قرار دارد، که داده‌ها را جابجا می‌کند. در شکل ۱-۱۳ ارتباط مجازی لایه‌ها با خط چین و ارتباط واقعی و فیزیکی با خط ممتد نشان داده شده است.

بین هر زوج از لایه‌های مجاور واسط (interface) قرار دارد. واسط مشخص می‌کند که هر لایه چه سرویسها و عملکردهای پایه‌ای در اختیار لایه بالاتر می‌گذارد. تعریف واسط‌های مناسب از مهمترین وظایف طراحان شبکه است. لازمه این امر آنست که وظایف هر لایه دقیقاً مشخص و شناخته شده باشد. علاوه بر به حداقل رساندن اطلاعات رد و بدل شده بین لایه‌ها، یک واسط شسته و رفته کار عوض کردن پیاده‌سازی لایه‌ها را نیز آسان می‌کند، چون تنها کاری که باید کرد این است که پیاده‌سازی جدید دقیقاً همان سرویسهای پیاده‌سازی قدیمی را به همسایگان خود ارائه کند. در حقیقت، پیاده‌سازیهای متعددی در شبکه‌های مختلف وجود دارد، که هیچ خللی در ارتباط لایه‌ها ایجاد نمی‌کنند.

به مجموعه لایه‌ها و پروتکل‌ها معماری شبکه (network architecture) می‌گویند. مشخصه‌های یک معماری باید آنچنان دقیق و جامع باشد که طراح شبکه بتواند نرم‌افزارها و سخت‌افزارهای لازم برای کارکرد صحیح آنرا فراهم آورد. جزئیات پیاده‌سازی و مشخصات واسط‌ها هرگز جزء معماری شبکه نیست، چون آنها باید در دل ماشین مخفی باشند (و از خارج دیده نشوند). حتی لازم نیست واسط‌ها در تمام ماشینهای یک شبکه یکسان باشند، مشروط باینکه تمام این ماشینها بتوانند از تمام پروتکل‌ها استفاده کنند. به مجموعه پروتکل‌هایی که در یک سیستم خاص بکار می‌روند (یک پروتکل در هر لایه)، پشته پروتکل (protocol stack) گفته می‌شود. پروتکل، پشته پروتکل و معماری شبکه از مهمترین موضوعات مورد بحث این کتاب هستند.

شاید یک مثال بتواند در روشن شدن مفهوم ارتباط چند لایه کمک کند. دو فیلسوف را، که اولی فقط زبانهای اردو و انگلیسی می‌داند و دیگری فقط چینی و فرانسه، در نظر بگیرید (آنها معادل پروسهای همتا در لایه 3 هستند). از آنجائیکه این دو فیلسوف نمی‌توانند مستقیماً با هم حرف بزنند، دو مترجم استخدام می‌کنند



شکل ۱-۱۴. معماری فیلسوف-مترجم-منشی.

(پروسسهای همتا در لایه 2)، که آنها هم بنویسند خود هر کدام یک منشی دارند (پروسسهای همتا در لایه 1). فیلسوف 1 میل دارد علاقه خود به *oryctolagus cuniculus* را به فیلسوف 2 (همتای خود) ابلاغ کند. برای اینکار، از طریق واسط 2/3 یک پیام بزبان انگلیسی با مضمون "I like rabbits" به مترجم خود می فرستد (شکل ۱-۱۴ را ببینید). مترجمها بین خود توافق کرده اند که به زبان هلندی حرف بزنند، پس پیام فوق به "Ik vind konijnen leuk" تبدیل می شود. انتخاب زبانی که همتاهای لایه 2 با آن صحبت کنند، بر عهده خود آنان (پروسسهای این لایه) است.

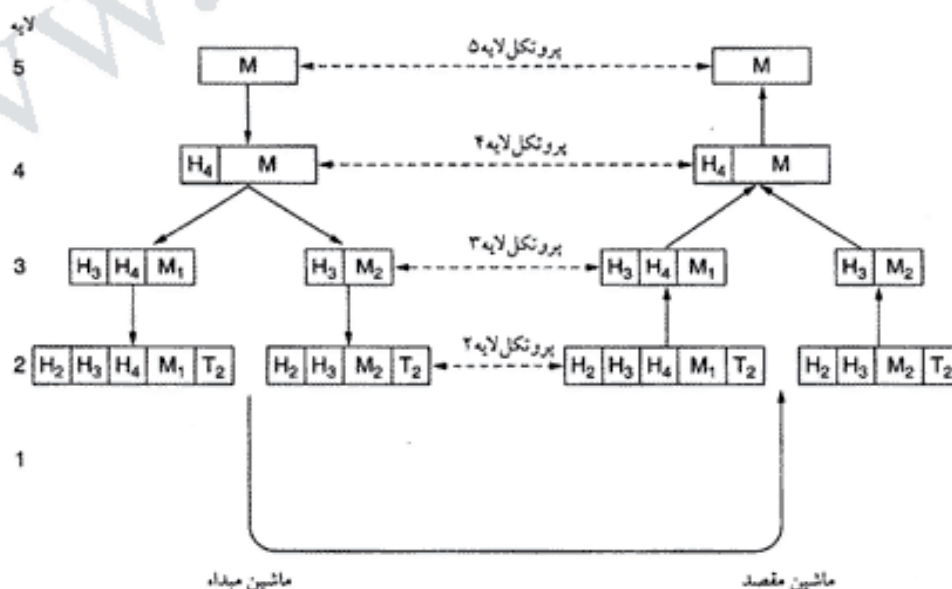
مترجم، سپس، این پیام را به منشی خود می دهد تا مثلاً از طریق فکس (پروتکل لایه 1) ارسال شود. وقتی این پیام به منشی سمت مقابل رسید، آنرا به مترجم (لایه 2) تحویل داده، و مترجم نیز پس از ترجمه به زبان فرانسه آنرا، از طریق واسط 2/3، به فیلسوف 2 می دهد. توجه کنید که تا زمانیکه واسطها تغییر نکرده باشند، پروتکلها کاملاً از یکدیگر مستقل هستند. برای مثال، مترجمها می توانند زبان توافقی خود را به فنلاندی تغییر دهند، مشروط باینکه واسط آنها با لایه های 1 و 3 هیچ تغییری نکند. یا اینکه منشیها می توانند از فکس به ایمیل یا تلفن سوئیچ کنند، بدون اینکه این کار هیچ تأثیری روی لایه های دیگر بگذارد. هر پروسس می تواند اطلاعات دلخواه خود را (که فقط بدرد پروسس همتای آن می خورد) به پیام اضافه کند. این اطلاعات به لایه های بالاتر منتقل نخواهند شد. حال یک مثال فنی تر را در نظر می گیریم: نحوه ارتباط در یک شبکه پنج لایه (شکل ۱-۱۵ را ببینید). برنامه ای

که در لایه 5 اجرا می شود، یک پیام (M) تولید کرده و برای ارسال به لایه 4 می دهد. لایه 4 یک سرآیند (header) به این پیام اضافه، و آنرا به لایه 3 تحویل می دهد. سرآیند حاوی اطلاعات کنترلی بین لایه های متناظر است؛ برای مثال، لایه 4 می تواند به هر پیام یک عدد ترتیبی نسبت بدهد، تا اگر لایه های پائینتر آنها را بدون نظم و ترتیب ارسال کردند، لایه متناظر در سمت مقابل بتواند آنها را به ترتیب صحیح بازبایی کند. در برخی از لایه ها، این سرآیندها حاوی اندازه بسته، زمان ارسال و اطلاعاتی از این قبیل است.

در بسیاری از شبکه اندازه پیام در لایه 4 هیچ محدودیتی ندارد، ولی (تقریباً همیشه) این محدودیت در لایه 3 وجود دارد. در نتیجه، لایه 3 باید پیام را به قطعات کوچکتر بشکند، و به هر قطعه یک سرآیند لایه 3 اضافه کند. در این مثال، پیام M به دو قطعه M_1 و M_2 شکسته شده است.

سپس، لایه 3 این بسته ها را به لایه بعدی (یعنی لایه 2) تحویل می دهد. لایه 2 علاوه بر اضافه کردن سرآیند خاص خود به ابتدای هر بسته، به انتهای آنها نیز یک دنباله (trailer) چسبانده، و آنها را به لایه بعدی (که لایه انتقال فیزیکی است) می دهد. در سمت مقابل، ماشینی که این بسته ها را دریافت می کند، آنها را لایه به لایه بالا می فرستد، و هر لایه (قبل از تحویل به لایه بالاتر) اطلاعات خاص خود را از بسته ها برمی دارد. بدین ترتیب، هیچ لایه ای سرآیندها و دنباله های لایه های زیرین خود را دریافت نخواهد کرد.

نکته مهمی که در شکل 1-15 باید بدان توجه کنید، رابطه ارتباطات مجازی و حقیقی بین ماشینها، و تفاوت پروتکل ها و واسطه ها است. برای مثال، پروتکل های همتا در لایه 4 فکر می کنند که ارتباط بین آنها یک ارتباط «افقی» است، که با استفاده از پروتکل لایه 4 برقرار می شود. هر یک از این پروتکل ها احتمالاً ابزارهایی بنام *SendToOtherSide* (برای ارسال پیام به سمت مقابل) یا *GetFromOtherSide* (برای دریافت پیام از سمت مقابل) نیز دارند، و حتی نمی دانند که پیامهای خود را از طریق واسطه 3/4 به لایه پائین تر می دهند، نه به طرف مقابل. تجریدی بودن رابطه پروتکل های همتا یک نکته کلیدی در طراحی شبکه است. با این تمهید، کار بسیار پیچیده و دشواری مانند طراحی کامل یک شبکه به کارهای کوچکتر و ساده تری (مانند طراحی لایه های جداگانه) شکسته می شود.



شکل 1-15. انتقال اطلاعات در یک شبکه پنج لایه.

با اینکه نام این بخش «نرم افزار شبکه» است، اما باید متوجه باشید که لایه‌های پائین‌تر معمولاً بصورت سخت‌افزاری پیاده‌سازی می‌شوند. با این حال، اینها نیز الگوریتمهای پیچیده نرم‌افزاری هستند، که فقط در سخت‌افزار حک و ثابت شده‌اند.

۲-۳-۱ ملاحظات در طراحی لایه‌ها

برخی از مفاهیم کلیدی طراحی شبکه‌های کامپیوتری در لایه‌های مختلفی حضور دارند، که در زیر برخی از مهمترین آنها را بطور مختصر بررسی خواهیم کرد.

هر لایه به مکانیزمی برای شناسایی فرستنده (sender) و گیرنده (receiver) نیاز دارد. از آنجائیکه یک شبکه معمولاً تعداد زیادی کامپیوتر دارد، و در هر کامپیوتر پروسس‌های متعددی در حال اجرا هستند، باید ابزاری وجود داشته باشد که هر پروسس بتواند پروسس‌های خود را دقیقاً شناسایی و مشخص کند. بعبارت دیگر، برای تعیین دقیق مقصد به یک نظام آدرس‌دهی (addressing) نیاز داریم.

نحوه انتقال داده‌ها نیز نیازمند قواعد و مقررات خاص خود است. در برخی از سیستمها داده‌ها فقط در یک جهت حرکت می‌کنند، اما در برخی دیگر در هر دو جهت. همچنین پروتکل باید تعیین کند که هر ارتباط فیزیکی معادل چند کانال منطقی است، و اولویت‌بندی آنها چگونه است. در بسیاری از شبکه‌ها هر ارتباط فیزیکی حداقل دو کانال منطقی را شامل می‌شود، یکی برای داده‌های معمولی و دیگری برای داده‌های اضطراری.

مسئله مهم دیگر کنترل خطا (error control) است، چون هیچ ارتباط فیزیکی صد در صد کامل و عاری از خطا نیست. گداهای بسیاری برای کشف و تصحیح خطا وجود دارد، ولی هر دو طرف باید بر سر یکی از آنها توافق کنند. همچنین گیرنده باید بتواند به طریقی به فرستنده اعلام کند که کدام پیامها را درست دریافت کرده و کدامها را غلط.

همانطور که قبلاً هم گفتیم، در بسیاری از موارد بسته‌های تشکیل دهنده یک پیام بصورت منظم ارسال نمی‌شوند؛ در این حالت، پروتکل باید طوری طراحی شده باشد که گیرنده بتواند قطعات پیام را دوباره بنحو صحیح بهم بچسباند. یک راه حل شماره‌گذاری قطعات (بسته‌های) پیام است، ولی باز این سوال باقیست که گیرنده باید چگونه با این بسته‌ها عمل کند.

یکی از مسائل که در تمام لایه‌ها وجود دارد اینست که، فرستنده با چه سرعتی باید اطلاعات را ارسال کند تا گیرنده‌های گنبد در گرداب داده‌ها غرق نشوند. برای این مشکل نیز راه حل‌های مختلفی وجود دارد که بعداً مفصلاً درباره آنها توضیح خواهیم داد. برخی از این راه حل‌ها شامل نوعی فیدبک (بازخور) از گیرنده به فرستنده است، که در هر لحظه وضعیت گیرنده را (بطور مستقیم یا غیرمستقیم) به فرستنده اعلام می‌کند. در برخی دیگر، دو طرف از قبل بر سر نرخ انتقال اطلاعات توافق می‌کنند. این مبحث کنترل جریان (flow control) نام دارد.

مسئله دیگری که در لایه‌های متعدد باید حل شود اینست که، اغلب پروسسها قادر نیستند پیامهایی با هر طول دلخواه دریافت کنند. این وضعیت باعث شده تا مکانیزمهایی برای شکستن پیامها به قطعات کوچک، ارسال، و سپس مونتاژ آنها در مقصد ابداع شود. مشکل دیگری که از اینجا ناشی می‌شود آنست که برخی از پروسسها اصرار دارند پیامها را آنقدر ریز کنند، بگونه‌ایکه کارایی کل سیستم را مختل می‌کند. در اینجا راه حل چسباندن چند قطعه به یکدیگر و شکستن دوباره آنها در مقصد است.

در اغلب موارد ایجاد کانالهای ارتباطی جداگانه برای هر زوج پروسس کاری پرهزینه (و گاهی غیرممکن) است. در این موارد، لایه‌های پائین‌تر برای برقراری ارتباط بین چند پروسس مستقل، از یک کانال استفاده می‌کنند. این عمل (که مالتی‌پلکس - multiplexing - و دمالتی‌پلکس - demultiplexing - نام دارد) معمولاً بصورت شفاف انجام می‌شود، که در اینصورت می‌توان آنرا در هر لایه‌ای پیاده‌سازی کرد. برای مثال، وقتی تعداد خطوط

مخابراتی موجود محدود است، در لایه فیزیکی از تکنیکهای مالتی پلکس استفاده می شود. وقتی بین مبدأ و مقصد مسیرهای مختلفی وجود دارد، یکی از آنها باید انتخاب شود. گاهی تصمیم گیری در این مورد باید در چند لایه انجام شود. مثلاً، برای ارسال اطلاعات از لندن به رُم اولاً (در لایه های بالاتر) باید تصمیم بگیریم که از مسیر فرانسه استفاده کنیم یا از مسیر آلمان؛ سپس در لایه های پایینتر انتخاب مسیرهای خلوت تر از میان مسیرهای موجود پیش می آید. به این مبحث مسیریابی (routing) گفته می شود.

۳-۳-۱ سرویسهای اتصال-گرا و غیرمتصل

هر لایه می تواند دو نوع سرویس در اختیار لایه بالاتر از خود بگذارد: سرویس اتصال-گرا (connection-oriented) و سرویس غیرمتصل (connectionless). در این قسمت این سرویس ها و تفاوت های آنها را بررسی خواهیم کرد.

سرویس اتصال-گرا بر اساس مدل سیستمهای تلفن کار می کند. وقتی می خواهید با یک نفر تماس بگیرید، گوشی تلفن را برداشته، شماره می گیرید، صحبت می کنید، و بعد گوشی را می گذارید. در یک سرویس اتصال-گرا هم ابتدا اتصال برقرار شده، و بعد از تبادل اطلاعات مورد نظر، اتصال قطع می شود. مهمترین نکته در مورد سرویسهای اتصال-گرا اینست که آنها مانند یک لوله عمل می کنند: فرستنده از یک طرف داده ها (بیت ها) را به داخل لوله می فرستند، و گیرنده در طرف دیگر آنها را می گیرد. در اغلب موارد داده ها بهمان ترتیبی که فرستاده شده اند، دریافت می شوند.

در برخی موارد بعد از برقراری اتصال، فرستنده، گیرنده و زیر شبکه ابتدا یک سری مذاکرات اولیه (negotiation) انجام می دهند تا بر سر مواردی از قبیل حداکثر اندازه پیامها، کیفیت سرویس مورد نظر، و مانند آن توافق کنند. معمولاً، یک طرف پیشنهادی می دهد و طرفهای دیگر آنرا قبول یا رد کرده، و با بکلی پیشنهاد جدیدی ارائه می کنند.

از سوی دیگر، سرویس غیرمتصل بر اساس مدل پُست بنا شده است. هر پیام (نامه) دارای آدرس مشخصی است، و مسیری که برای رسیدن به مقصد طی می کند، کاملاً مستقل از پیامهای دیگر است. معمولاً وقتی دو نامه به یک مقصد می فرستید، اولین نامه زودتر از دومی به آنجا می رسد؛ ولی گاهی پیش می آید که اولی با تأخیر و بعد از دومی به مقصد برسد.

برای هر سرویس می توان یک کیفیت سرویس (quality of service) در نظر گرفت. برخی از سرویسها مطمئن و قابل اعتماد هستند، بگونه ای که هیچ داده ای در حین انتقال از بین نمی رود. یک سرویس قابل اعتماد معمولاً بگونه ای طراحی می شود که گیرنده دریافت صحیح داده ها را به فرستنده اعلام کند. این تصدیق دریافت (acknowledgement) باعث تحمیل یک بار اضافی و تأخیر در انتقال پیامها می شود، که اغلب ارزش آنرا دارد، ولی گاهی به زحمتش نمی ارزد.

انتقال فایل (file transfer) از جمله مواردیست که به یک سرویس مطمئن اتصال-گرا نیاز دارد؛ صاحب فایل معمولاً میل دارد تمام بیت های فایلش (با همان نظم و ترتیب) به مقصد برسد. کمتر کسی را پیدا می کنید که راضی شود یک فایل فروقاتی دریافت کند، حتی اگر اینکار به معنای سرعت بیشتر باشد.

سرویس اتصال-گرای قابل اعتماد بر دو گونه مختلف است: توالی پیام (message sequence) و جریان بایت (byte stream). در سرویس توالی پیام حد و مرز پیامها همیشه حفظ می شود؛ وقتی دو پیام ۱۰۲۴ بایتی می فرستید، طرف مقابل همیشه دو پیام ۱۰۲۴ بایتی دریافت خواهد کرد، نه یک پیام ۲۰۴۸ بایتی، یا چهار پیام ۵۱۲ بایتی. اما در سرویس دوم، چیزی بنام حد و مرز پیام وجود ندارد، و فقط جریانی از بایتها دیده خواهد شد. در این حالت وقتی ۲۰۴۸ بایت به مقصد می رسد، به هیچ طریقی نمی توان گفت که آیا این یک پیام ۲۰۴۸ بایتی بوده، یا دو

پیام ۱۰۲۴ بایتی، و یا ۲۰۴۸ پیام ۱ بایتی. برای مثال، اگر بخواهید صفحات یک کتاب را به یک دستگاه حرفه‌چینی الکترونیکی بفرستید، شاید برایتان مهم باشد که حد و مرز هر صفحه مشخص باشد. از طرف دیگر، وقتی از راه دور به یک کامپیوتر متصل می‌شوید، جریان بایت‌ها از کامپیوتر مبدأ به مقصد تمام آن چیز است که نیاز دارید، و حد و مرز پیامها هیچ اهمیتی ندارد.

اما همانطور که گفتیم، در برخی از موارد تأخیری که در نتیجه تصدیق دریافت (acknowledgement) پدید می‌آید، غیر قابل قبول است؛ مکالمه دیجیتال یکی از این موارد است. اغلب کاربران ترجیح می‌دهند گاهی صدای طرف مقابل را با نویز بشنوند، تا اینکه (در نتیجه مکانیزم تصدیق دریافت) مکالمه با تأخیر و وقفه انجام شود. و یا در کنفرانسهای ویدئویی کمی برفک و نویز قابل تحمل‌تر است، تا اینکه تصاویر با پرشهای اعصاب خردکن دریافت شود.

از طرف دیگر، همه کاربردها به ارتباط متصل نیاز ندارند؛ پست الکترونیک (ایمیل) یکی از مواردیست که نیازی به سرویس اتصال-گرا ندارد، و بویژه در مواقعی که هزینه عملی تعیین‌کننده است، سرویس قابل اعتماد نیز چندان الزامی نیست. در این موارد فقط کفایت اولویت تحویل پیام بالا باشد. سرویس غیر متصل غیر قابل اعتماد (سرویس‌ی که به تصدیق دریافت از طرف مقابل متکی نیست) اغلب بعلت شباهتی که با سیستم تلگراف دارد، سرویس دیتاگرام (datagram service) نامیده می‌شود.

اما گاهی با اینکه نیازی به برقراری یک اتصال نیست (مثلاً برای ارسال یک پیام کوتاه)، ولی قابل اعتماد بودن ارتباط اهمیت دارد. در این قبیل موارد می‌توان از سرویس دیتاگرام همراه با تصدیق دریافت (acknowledged datagram service) استفاده کرد. این مانند پست کردن یک نامه سفارشی است، که فرستنده می‌خواهد از رسیدن نامه بدست گیرنده مطمئن شود. با دریافت این تأییدیه، فرستنده مطمئن می‌شود که نامه گم نشده و بدست طرف مقابل رسیده است.

سرویس دیگری نیز وجود دارد، که سرویس درخواست-پاسخ (request-reply service) نام دارد. در این سرویس، فرستنده دیتاگرامی که حاوی یک درخواست است ارسال می‌کند، و پاسخ آنرا می‌گیرد. برای مثال، وقتی پیامی به کتابخانه محلی می‌فرستید و سؤال می‌کنید که «زبان سواحیلی در کدام کشور تکلم می‌شود»، از این سرویس استفاده می‌کنید. از سرویس درخواست-پاسخ معمولاً در سیستمهای مشتری-سرویس دهنده استفاده می‌شود؛ مشتری درخواست خود را به سرویس دهنده فرستاده، و پاسخ آنرا دریافت می‌کند. در شکل ۱-۱۶ خلاصه‌ای از سرویسهای توضیح داده شده در این قسمت را مشاهده می‌کنید.

	سرویس	مثال
اتصال گرا	استریم پیام قابل اعتماد	چند صفحه متوالی
	استریم بایت قابل اعتماد	ورود آزاره دور
	اتصال غیر قابل اعتماد	صدای دیجیتالی
غیر متصل	دیتا گرام غیر قابل اعتماد	
	دیتا گرام تصدیق شده	ایمیل ثبت شده
	درخواست - پاسخ	جستجوی پایگاه داده

شکل ۱-۱۶. شش نوع سرویس مختلف.

شاید در نگاه اول تعجب کنید که اصولاً چرا باید از یک سرویس غیرقابل اعتماد استفاده کنیم، و اصلاً چه کسی چنین چیزی را لازم دارد؟ اول از همه اینکه، امکان دارد در مواردی سرویس قابل اعتماد اساساً در دسترس نباشد. برای مثال، اینترنت یک ارتباط قابل اعتماد نیست، و گاهی ممکنست داده‌ها در حین انتقال صدمه ببینند؛ این بر عهده پروتکل‌های لایه‌های بالاتر است که این مشکل را حل کنند. دوم اینکه، تأخیر ذاتی سرویس‌های مبتنی بر تصدیق دریافت در مواردی (مانند برنامه‌های چندرسانه‌ای) پذیرفتنی نیست. به این دلایل، وجود سرویس‌های قابل اعتماد و غیرقابل اعتماد هر دو لازم است.

۳-۱ عملکردهای پایه سرویس

هر سرویس با یک سری عملکردهای پایه (primitives) که در اختیار کاربر خود می‌گذارد، شناخته می‌شود. این عملکردهای پایه یا خود کاری را انجام می‌دهند، و یا انجام کاری را در طرف مقابل گزارش می‌کنند. اگر پشته پروتکل (protocol stack) جزئی از سیستم عامل باشد (که اغلب نیز چنین است)، عملکردهای پایه نیز معمولاً جزء فراخوانی‌های سیستم (system call) هستند. این فراخوانی‌ها باعث فعال شدن کُدی در هسته سیستم عامل شده، و ارسال بسته‌های پیام انجام می‌شود.

عملکردهای پایه هر سرویس به خصلت آن سرویس بستگی دارد. برای مثال، عملکردهای پایه یک سرویس اتصال-گرافتفاوت از سرویس‌های غیرمتصل است. در شکل ۱-۱۷ حداقل عملکردهای پایه لازم برای پیاده‌سازی یک سرویس اتصال-گرای جریان بایت قابل اعتماد را در یک محیط مشتری-سرویس دهنده ملاحظه می‌کنید. طرز استفاده از این عملکردهای پایه مانند زیر است. ابتدا، کامپیوتر سرویس دهنده LISTEN را اجرا می‌کند تا نشان دهد که آماده پذیرش ارتباطات ورودی است. عملکرد LISTEN معمولاً بصورت یک فراخوانی مسدودشونده (blocking) پیاده‌سازی می‌شود؛ بدین معنا که بعد از اجرای این عملکرد، پروسس سرویس دهنده تا زمان دریافت درخواست اتصال مسدود می‌شود.

سپس، مشتری عملکرد CONNECT را اجرا می‌کند تا به سرویس دهنده متصل شود. فراخوانی CONNECT معمولاً باید مشخص کند که مقصد اتصال کجاست، بهمن دلیل ممکنست پارامتری داشته باشد که آدرس سرویس دهنده را بدست می‌دهد. با این عمل، سیستم عامل مشتری پیامی را به همتای خود می‌فرستد، و درخواست اتصال می‌کند - این مرحله با شماره (1) در شکل ۱-۱۸ نشان داده شده است. پس از آن، پروسس مشتری تا زمان دریافت پاسخ به حالت تعلیق (suspend) درمی‌آید. وقتی این بسته به سرویس دهنده رسید، توسط سیستم عامل پردازش می‌شود، و وقتی می‌بیند یک درخواست اتصال است، بدنبال یک پروسس شنونده (listener) می‌گردد. اگر چنین پروسسی را پیدا کند، آنرا از حالت انسداد خارج کرده، و یک پیام تصدیق دریافت (acknowledgement) به مشتری پس می‌فرستد - مرحله (2). دریافت این پیام توسط مشتری باعث می‌شود تا پروسس از حالت تعلیق در آید. در این لحظه پروسس‌های سرویس دهنده و مشتری هر دو در حال اجرا هستند، و

عملکرد پایه	مفهوم
LISTEN	انتظار برای دریافت اتصال
CONNECT	برقراری ارتباط با همتای منتظر
RECEIVE	انتظار برای دریافت اتصال
SEND	ارسال پیام به همتا
DISCONNECT	پایان اتصال

شکل ۱-۱۷. پنج عملکرد پایه لازم برای پیاده‌سازی یک سرویس اتصال-گرای ساده.



شکل ۱-۱۸. تبادل بسته‌ها در یک شبکه اتصال-گرای مشتری-سرویس دهنده.

اتصال برقرار شده است. توجه به این نکته ضروریست که پیام تصدیق دریافت (2) توسط کُد پروتکل (که در سطح هسته - kernel level - اجرا می‌شود) ایجاد می‌شود، نه کُد عملکرد پایه (که یک پروسس سطح کاربر - user level - است). اگر هنگام دریافت درخواست اتصال توسط سرویس دهنده، هیچ پروسس شنونده‌ای وجود نداشته باشد، نتیجه نامشخص است. در برخی از سیستمها، این درخواست برای مدتی در صف (queue) می‌ماند، به امید اینکه شاید یک پروسس LISTEN اجرا شود.

این فرآیند بسیار شبیه تماس تلفنی مشتری با مدیر قسمت پشتیبانی مشتریان در یک شرکت است. نشستن مدیر قسمت پشتیبانی در کنار تلفن بنوعی اعلام آمادگی برای دریافت تقاضاها است (همان پروسس LISTEN). سپس یکی از مشتریان زنگ می‌زند (پروسس CONNECT)؛ و به محض اینکه مدیر پشتیبانی گوشی تلفن را برداشت، ارتباط برقرار می‌شود.

قدم بعدی را باید سرویس دهنده بردارد: اجرای عملکرد RECEIVE برای دریافت اولین درخواست. معمولاً سرویس دهنده این کار را بلافاصله بعد از برطرف شدن انسداد (و حتی قبل از اینکه پاسخ تصدیق دریافت آن به مشتری برسد) انجام می‌دهد. فراخوانی RECEIVE نیز جزء پروسسهای مسدودشونده است. وقتی مشتری اعلام آمادگی سرویس دهنده را دریافت کرد، درخواست خود را در قالب یک عملکرد SEND به آن می‌فرستد - مرحله (3).

رسیدن بسته SEND به سرویس دهنده آنرا از حالت انسداد خارج کرده، و باعث می‌شود تا بتواند به درخواست مشتری رسیدگی کند. پس از آماده شدن پاسخ، سرویس دهنده آنرا با اجرای عملکرد SEND به مشتری پس می‌فرستد - مرحله (4). رسیدن این بسته به مشتری باعث می‌شود تا از حالت انسداد خارج شده، و محتوای پاسخ را بررسی کند. اگر مشتری درخواست‌های بیشتری داشته باشد، آنها را در همین مرحله انجام می‌دهد؛ در غیر اینصورت، با اجرای عملکرد DISCONNECT اتصال را قطع می‌کند. فراخوانی DISCONNECT نیز معمولاً یک فراخوانی مسدودکننده است، و باعث تعلیق پروسس مشتری و اعلام پایان ارتباط به سرویس دهنده می‌شود - مرحله (5). پروسس سرویس دهنده با دریافت پیام DISCONNECT از مشتری، عملکرد DISCONNECT را در سمت خود اجرا کرده، و (پس از اعلام به مشتری) ارتباط را قطع می‌کند. وقتی بسته سرویس دهنده به مشتری رسید، مشتری نیز اتصال را رها می‌کند - مرحله (6). این بود ماجرای ساده یک ارتباط اتصال-گرا!

البته، زندگی همیشه این قدر شیرین نیست، و خیلی چیزها می‌تواند آنرا تلخ کند. برای مثال، فرض کنید عملکرد CONNECT قبل از LISTEN اجرا شود، یکی از بسته‌ها وسط راه گم شود، و اتفاقاتی از این قبیل. بعداً درباره این مسائل بیشتر صحبت خواهیم کرد، اما فعلاً به همین ارتباط ساده شکل ۱-۱۸ دقت کنید.

با توجه به این نکته که برای کامل شدن سیکل این پروتکل به شش بسته نیاز داریم، شاید بپرسید که چرا بجای آن از یک ارتباط غیرمتصل (که فقط به دو بسته - یک درخواست و یک پاسخ - نیاز دارد) استفاده نکنیم. جواب اینست

که در یک دنیای بی عیب و نقص می توان چنین کرد؛ اما آیا اوضاع همیشه بر وفق مراد ماست؟ اگر فایل ها خیلی بزرگ باشند، خط ارتباطی پُر از خطا باشد، بسته ها مدام گم شوند، چکار باید کرد؟ اگر پای صدها و هزارها بسته در میان باشد، و فقط چند بسته ناقابل گم شود، مشتری چگونه باید این موضوع را متوجه شود؟ مشتری از کجا بفهمد آخرین بسته ای که گرفته، واقعاً آخرین بسته بوده، یا خیلی ساده ارتباط قطع شده است؟ فرض کنید، مشتری فایل دومی را هم درخواست می کند، و بعد یکبار بسته ای می رسد که شماره ۱ دارد. آیا این اولین بسته فایل دوم است، یا اولین بسته فایل اول که تا حالا سرگردان بوده، و همین الان راه خود را پیدا کرده؟ خلاصه، در دنیای واقعی یک شبکه غیر قابل اعتماد مبتنی بر پروتکل درخواست-پاسخ نمی تواند کافی باشد. در فصل ۳ پروتکل هایی را که برای رفع این مشکلات ابداع شده اند، به تفصیل بررسی خواهیم کرد. فعلاً همین قدر کفایت بدانید که، گاهی پروتکل هایی که به روش جریان بایت با هم ارتباط برقرار می کنند، بهترین گزینه اند.

۵-۳-۱ رابطه سرویس و پروتکل

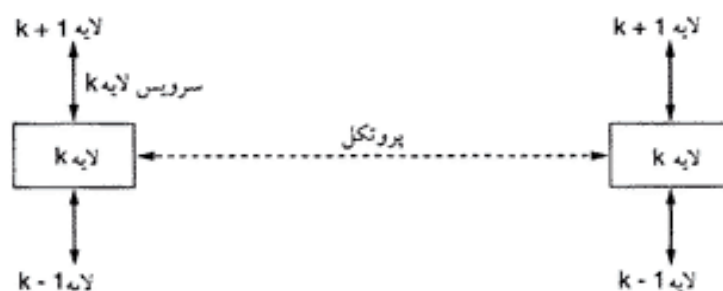
سرویس و پروتکل دو چیز متفاوتند، که اغلب افراد آنها را با هم اشتباه می گیرند. اما این تفاوت بقدری مهم است، که جا دارد باز هم بر آن تأکید کنیم. سرویس (service) عبارتست از مجموعه ای از عملکردهای پایه که یک لایه در اختیار لایه بالاتر از خود قرار می دهد. سرویس فقط می گوید که یک لایه چه کارهایی می تواند برای کاربر خود انجام دهد، ولی هیچ چیز درباره چگونگی آن نمی گوید. سرویس در واقع به واسطه دو لایه مربوط می شود، که در آن لایه پائینتر ارائه دهنده سرویس و لایه بالاتر مصرف کننده سرویس است.

اما، پروتکل (protocol) عبارتست از مجموعه قواعد حاکم بر فرمت، مفهوم و نحوه تبادل بسته ها و پیامها بین دو لایه همتا. در واقع این پروتکل است که سرویسهای تعریف شده در هر لایه را پیاده سازی می کند. همتایان هر لایه می توانند پروتکل ارتباطی خود را عوض کنند، مشروط بر اینکه سرویسهایی که به کاربران خود می دهند، تغییری نکنند. با این تعریف، سرویس و پروتکل کاملاً از هم مستقل هستند.

همانطور که در شکل ۱-۱۹ می بینید، سرویس به واسطه دو لایه مربوط می شود، در حالیکه پروتکل پیامهای مبادله شده بین دو لایه همتا (در دو کامپیوتر مختلف) را کنترل می کند. بسیار اهمیت دارد که این دو مفهوم را با هم مخلوط نکنید.

یک مثال مشابه از زبانهای برنامه نویسی می تواند در درک تمایز سرویس و پروتکل کمک کند. سرویس در واقع شبیه نوع داده (data type) یا شیء (object) در زبانهای شیء گرا است. با اینکه می دانیم یک شیء چه خواصی دارد، اما نمی دانیم آنها را چگونه پیاده سازی کرده است. این پیاده سازی همان پروتکل است که از چشم کاربر مخفی می ماند.

البته در بسیاری از پروتکل های قدیمی این تمایز بروشنی وجود ندارد. در این پروتکلها هر تغییری در پروتکل بلافاصله به چشم کاربر خواهد آمد. اما امروزه طراحان شبکه سعی می کنند این تمایز را رعایت کنند.



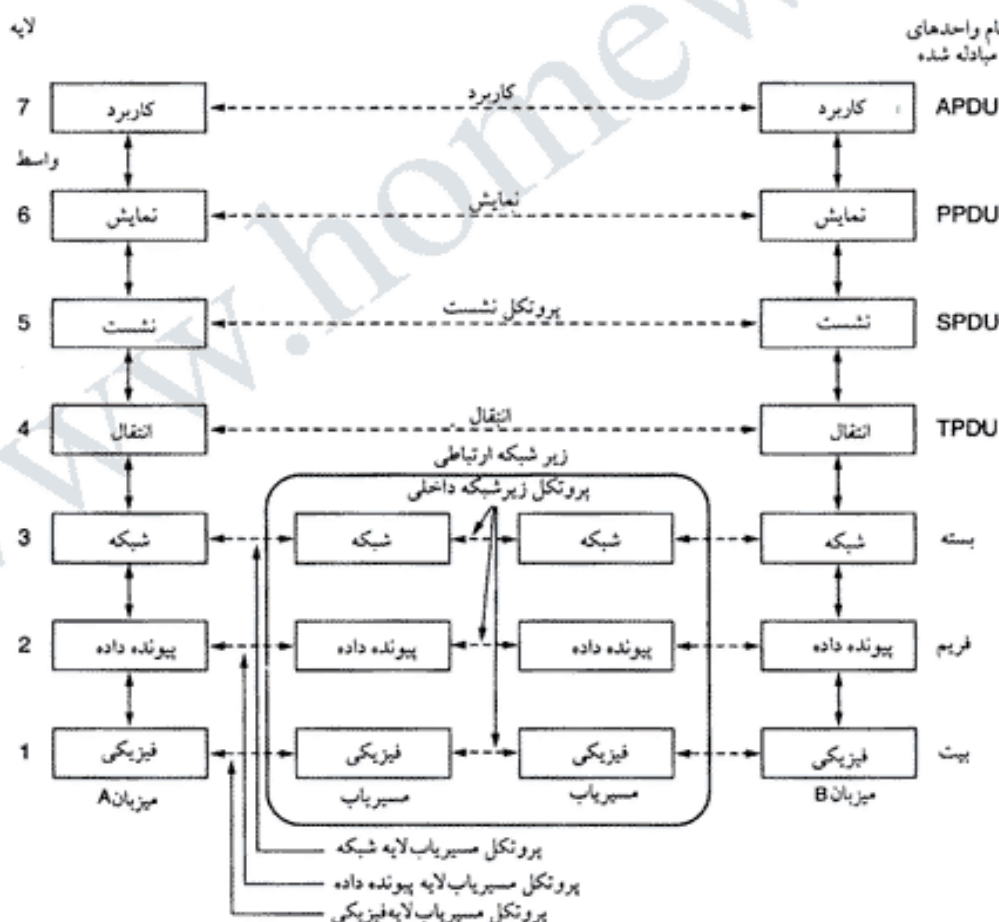
شکل ۱-۱۹. رابطه بین سرویس و پروتکل.

۴-۱ مدل های مرجع

حال که با شبکه های چند لایه بصورت تئوری آشنا شدید، وقت آنست که نگاهی به چند نمونه از این نوع شبکه ها بیندازیم. در دو قسمت آینده دو تا از مهمترین معماری های شبکه، مدل مرجع OSI و مدل مرجع TCP/IP، را بررسی خواهیم کرد. با اینکه پروتکل های مدل OSI امروزه بندرت مورد استفاده عملی دارند، اما این مدل همچنان معتبر، و مشخصات لایه های آن از اهمیت زیادی برخوردار است. وضعیت مدل TCP/IP بر عکس است: با اینکه خود این مدل کمتر مورد استفاده قرار می گیرد، ولی پروتکل های آن کاربرد وسیعی دارند. به همین دلیل، مدل های فوق را مفصلاً مورد بررسی قرار داده ایم.

۱-۴-۱ مدل مرجع OSI

در شکل ۱-۲۰-۱ مدل OSI را (منهای لایه رسانه فیزیکی) ملاحظه می کنید. این مدل بر اساس نظرات پیشنهادی سازمان بین المللی استاندارد ها (International Standards Organization - ISO) - بعنوان اولین استاندارد بین المللی شبکه های چند لایه - توسعه داده شد (Day and Zimmermann, 1983). این مدل در سال ۱۹۹۵ مورد تجدیدنظر قرار گرفت (Day, 1995). این مدل که نام کامل آن مدل مرجع ارتباطات سیستم های باز (ISO OSI - Open System Interconnection) است، با ارتباطات سیستم های باز - سیستم هایی که قادر به ارتباط با سیستم های دیگر هستند - سر و کار دارد؛ ولی ما آنرا بسادگی مدل OSI خواهیم نامید.



شکل ۱-۲۰. مدل مرجع OSI.

مدل OSI هفت لایه دارد. اما، چرا هفت لایه؟ برخی از مهمترین دلایل انتخاب هفت لایه عبارتند از:

۱. هر کجا به تجرید خاصی نیاز باشد، باید یک لایه ایجاد کرد.
۲. هر لایه باید وظیفه کاملاً مشخصی را انجام دهد.
۳. وظیفه هر لایه باید با در نظر گرفتن تعریف پروتکل‌های استاندارد بین‌المللی انتخاب شود.
۴. مرزهای هر لایه باید طوری انتخاب شود، که کمترین انتقال اطلاعات از آنها لازم باشد.
۵. تعداد لایه‌ها باید آنقدر زیاد باشد، که نیازی به تعریف توابع مشابه در یک لایه نباشد؛ و باید آنقدر کم باشد که معماری شبکه بیش از حد بزرگ و پیچیده نشود.

در زیر لایه‌های مدل OSI را (از پائین به بالا) مورد بحث قرار داده‌ایم. توجه داشته باشید که مدل OSI خود یک معماری شبکه نیست، چون هیچ سرویس یا پروتکلی در آن تعریف نمی‌شود. این مدل فقط می‌گوید که هر لایه چه کاری باید انجام دهد. با اینکه سازمان استانداردهای بین‌المللی (ISO) پروتکل‌های هر لایه را نیز تعریف کرده است، ولی آنها جزء مدل OSI نیستند، و جداگانه بصورت استانداردهای بین‌المللی منتشر می‌شوند.

لایه فیزیکی

لایه فیزیکی (physical layer) وظیفه انتقال بیت‌های خام را از طریق کانال مخابراتی بر عهده دارد. مهمترین نکته در طراحی این لایه اینست که وقتی یک طرف یک بیت 1 می‌فرستد، طرف مقابل یک بیت 1 دریافت کند، نه یک بیت 0. سؤالات اساسی در این لایه عبارتند از اینکه، برای 1 و 0 از چه ولتاژهایی استفاده کنیم، هر بیت باید چند نانوثانیه (یک میلیاردیم ثانیه) روی خط دوام بیاورد، آیا انتقال همزمان در هر دو جهت امکانپذیر باشد یا خیر، اتصال اولیه چگونه شروع شود و چگونه پایان یابد، رابط شبکه (network connector) چند پایه باید داشته باشد و وظیفه هر پایه چیست. مسائل طراحی در این لایه عمدتاً از نوع مکانیکی، الکتریکی، تایمینگ (همزمانی)، و رسانه فیزیکی انتقال (که زیر لایه فیزیکی قرار دارد) هستند.

لایه پیوند داده

مهمترین وظیفه لایه پیوند داده (data link layer) عبارتست از تبدیل خط فیزیکی پُر از خطا به یک خط ارتباطی عاری از خطا برای لایه بالاتر، یعنی لایه شبکه. لایه پیوند داده این کار را با شکستن داده‌های ورودی به بسته‌های کوچک چند صد یا هزار بیتی (که فریم داده - data frame - نامیده می‌شوند)، و ارسال آنها انجام می‌دهد. وقتی گیرنده هر بسته را دریافت می‌کند، یک فریم تصدیق دریافت (acknowledgement frame) به فرستنده باز پس می‌فرستد، تا آنرا از دریافت صحیح بسته مطلع کند.

مسئله دیگری که در لایه پیوند داده (و حتی لایه‌های بالاتر) باید حل شود اینست که، چگونه یک گیرنده کند را با یک فرستنده سریع هماهنگ کند. برای این منظور باید مکانیزمی تعبیه شود تا فرستنده در هر لحظه از مقدار پافر (buffer - حافظه موقتی) گیرنده مطلع باشد. در اغلب موارد این دو ویژگی - کنترل جریان اطلاعات و مقابله با خطا - در هم ادغام می‌شوند.

لایه پیوند داده در شبکه‌های پخش (broadcast) باید با مسئله دیگری نیز دست و پنجه نرم کند: کنترل دسترسی به یک کانال مشترک. برای این منظور از زیرلایه‌ای بنام کنترل دسترسی رسانه (medium access control) در لایه پیوند داده استفاده می‌شود.

لایه شبکه

لایه شبکه (network layer) عملکرد زیر شبکه را کنترل می‌کند. یکی از مسائلی که باید در این لایه حل شود، نحوه مسیریابی بسته‌ها از مبدأ به مقصد است. این مسیرها می‌توانند مسیرهای استاتیک باشند (مسیرهایی که بطور

ثابت و پندرت متغیر در شبکه تعبیه شده‌اند)، یا مسیرهای نیمه‌استاتیک (مسیرهایی که در ابتدای هر نشست تعیین و مشخص می‌شوند)، و یا مسیرهای دینامیک (مسیرهایی که در هر لحظه و برای هر بسته از نو - و با توجه به بار شبکه - جستجو و مشخص می‌شوند).

اگر تعداد بسته‌های در حال حرکت در یک زیرشبکه بیش از حد باشد، آنها راه یکدیگر را بند آورده و وضعیتی را بوجود می‌آورند که به آن گلوگاه (bottleneck) یا ازدحام (congestion) گفته می‌شود. کنترل این وضعیت نیز بر عهده لایه شبکه است. به بیان کلی، کیفیت سرویس (تأخیر انتشار بسته‌ها، زمان انتقال آنها، و حالت‌های گذرا در شبکه) همگی جزء مسئولیت‌های لایه شبکه است.

اگر یک بسته برای رسیدن به مقصد خود باید از یک شبکه خارج و وارد شبکه دیگری شود، مسائل جدیدی بروز خواهد کرد. اول اینکه، امکان دارد آدرس دهی در این شبکه‌ها متفاوت باشد. دوم اینکه، ممکن است شبکه دوم این بسته را بکلی نپذیرد، چون مثلاً اندازه آن بیش از حد بزرگ است؛ و یا اینکه پروتکلها با هم فرق داشته باشند، و مسائلی از این قبیل. حل همه این مسائل (و بهم پیوستن شبکه‌های ناهمگن) نیز از وظایف لایه شبکه است. در شبکه‌های پخش مسیریابی بسته‌ها ساده است، بهمین دلیل لایه شبکه با بسیار کوچک است یا اصلاً نیست.

لایه انتقال

اصلی‌ترین وظیفه لایه انتقال (transport layer) گرفتن داده‌ها از لایه بالاتر، تقسیم آن به قطعات کوچکتر (در صورت نیاز)، ارسال آن به لایه شبکه، و حصول اطمینان از دریافت صحیح آنها در طرف مقابل است. علاوه بر آن، همه این کارها باید بگونه‌ای مؤثر و طوری انجام شود که لایه‌های بالاتر را از تغییرات اجتناب‌ناپذیر در سخت‌افزار ایزوله کند.

این لایه همچنین تعیین می‌کند که چه سرویسهایی باید در اختیار لایه نشست (و از آنجا، در اختیار کاربران شبکه) قرار گیرد. متداولترین نوع انتقال کانالهای نقطه-به-نقطه عاری از خطاست، که در آن بایتها بهمان ترتیبی که فرستاده شده‌اند، در طرف مقابل دریافت می‌شوند. با این حال انواع دیگری از سرویسهای انتقال وجود دارد، که از میان آنها می‌توان به انتقال پیام بدون اطمینان از دریافت منظم آنها، و ارسال همزمان پیامهای پخشی به چندین نقطه اشاره کرد. نوع این سرویسه در لحظه برقراری اتصال مشخص می‌شود. (البته باید تصریح کنیم که کانال عاری از خطا در دنیای واقعی وجود خارجی ندارد، و آنچه از این اصطلاح برداشت می‌شود پائین بودن نرخ خطاست، بگونه‌ای که بتوان در عمل آنرا نادیده گرفت).

لایه انتقال یک لایه نقطه-به-نقطه واقعی است، که در آن کامپیوتر فرستنده (مبدأ) مستقیماً با کامپیوتر گیرنده (مقصد) ارتباط دارد. در لایه‌های پائینتر ارتباط ماشین مبدأ معمولاً با ماشینهای همسایه (و نه ماشین مقصد) است. تفاوت لایه‌های ۱ تا ۳ (که بصورت زنجیره‌ای هستند) با لایه‌های ۴ تا ۷ (که نقطه-به-نقطه هستند) در شکل ۱-۲۰ نشان داده شده است.

لایه نشست

لایه نشست (session layer) اجازه می‌دهد تا بین کاربران در ماشینهای مختلف نشست برقرار شود. نشست سرویسهای مختلفی ارائه می‌کند، از جمله: کنترل دیالوگ (dialog control - کنترل اینکه نوبت چه کسی است)، مدیریت نشانه (token management - جلوگیری از تداخل اعمال مهم)، و سنکرون کردن (synchronization - کنترل عملیات انتقال طولی‌المدت، و از سرگیری آن از نقطه قطع شده در صورت بروز اختلال).

لایه نمایش

بر خلاف لایه‌های پائینتر، که عمدتاً با بیت‌ها سروکار دارند، لایه نمایش (presentation layer) توجه خود را

روی ساختار پیامها و مفهوم آنها متمرکز می‌کند. برای اینکه کامپیوترهایی با ساختارهای متفاوت داده بتوانند با هم ارتباط برقرار کنند، ساختار پیامهای مبادله شده بایستی کاملاً مشخص و استاندارد باشد. وظیفه لایه نمایش مدیریت این ساختارها در سطح بالاست.

لایه کاربرد

بسیاری از پروتکل‌های مورد نیاز کاربران در لایه کاربرد (application layer) قرار دارد، که از معروفترین آنها می‌توان به پروتکل HTTP (HyperText Transfer Protocol) - پروتکل اصلی وب - اشاره کرد. وقتی مرورگر وب می‌خواهد صفحه‌ای را بار کند، نام آن صفحه را با استفاده از پروتکل HTTP به سرور می‌دهد و وب می‌فرستد؛ سرور دهنده وب نیز با همین پروتکل صفحه را به مرورگر برمی‌گرداند. پروتکل انتقال فایل (FTP)، پروتکل انتقال خبر (NNTP)، و پروتکل‌های پست الکترونیک (SMTP و POP) نیز جزء پروتکل‌های کاربردی هستند.

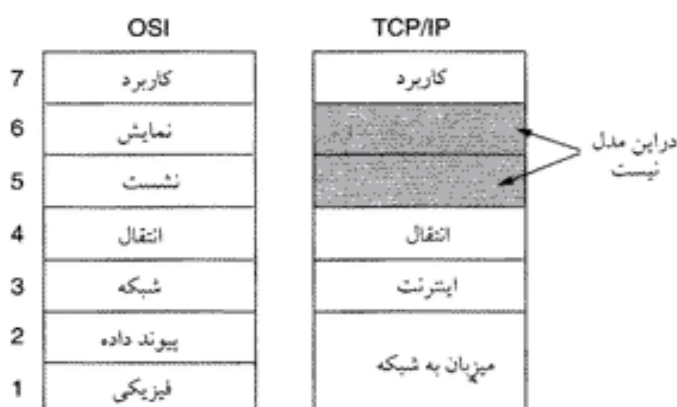
۲-۴-۱ مدل مرجع TCP/IP

اکنون اجازه دهید به مدل مرجع بکار رفته در پذیرش شبکه‌های کامپیوتری، آرپانت (ARPANET)، و خلف آن اینترنت، بپردازیم. آرپانت یک شبکه تحقیقاتی بود که توسط DoD (وزارت دفاع ایالت متحده آمریکا) پایه‌ریزی شد. بتدریج صدها دانشگاه و مرکز دولتی بوسیله خطوط اجاره‌ای تلفن (leased line) به این شبکه ملحق شدند. با پیشرفت مخابرات رادیویی و ماهواره، مشکلاتی در پروتکل‌های ارتباطی آرپانت بوجود آمد، که انتخاب یک معماری مرجع جدید را الزامی می‌کرد. یکی از اولین هدفهای آرپانت ارتباط یکپارچه شبکه‌های مختلف بود، که بالاخره (بعد از بررسی چندین پروتکل) توسط مدل مرجع TCP/IP محقق شد. این مدل برای اولین بار توسط (Cerf and Kahn, 1974) تعریف شد، که در سال ۱۹۸۵ مورد تجدیدنظر قرار گرفت (Leiner et al., 1985). فلسفه طراحی مدل مرجع TCP/IP در (Clark, 1988) مورد بحث قرار گرفته است.

دغدغه همیشگی وزارت دفاع این بوده که بخشی از شبکه و وسایل با ارزش آن در یک لحظه (احتمالاً در یک حمله اتمی) نیست و نابود شود، و به همین دلیل همواره تأکید داشته که این شبکه باید بگونه‌ای طراحی شود که حتی در صورت از بین رفتن بخشی از زیر شبکه‌های آن، بتواند بدون وقفه به کار خود ادامه دهد. بعبارت دیگر، هدف این است که دو کامپیوتر مادامیکه که کار می‌کنند، باید بتوانند با هم ارتباط داشته باشند (حتی اگر تعدادی از ماشینهای واسط بین آنها از مدار خارج شوند). علاوه بر آن، این مدل باید بتواند از عهده طیف وسیعی از کاربردهای متنوع (از انتقال فایل گرفته، تا مکالمه زمان واقعی) برآید.

لایه اینترنت

تمام این الزامات باعث شد تا در نهایت یک شبکه سوئیچینگ بسته (packet-switching) مبتنی بر یک لایه ارتباطات غیر متصل (connectionless) انتخاب شود. این لایه، که لایه اینترنت (internet layer) نام دارد، سنگ بنای معماری TCP/IP است. وظیفه این لایه اینست که به ماشینها اجازه دهد بسته‌های خود را روی شبکه و به سمت مقصد بفرستند. این لایه رسیدن پیامها را با همان ترتیبی که فرستاده شده‌اند، تضمین نمی‌کند؛ وظیفه مرتب کردن پیامها (در صورت نیاز) بر عهده لایه‌های بالاتر است. (دقت کنید که در اینجا «اینترنت» یک کلمه عام است.) لایه اینترنت تا حد زیادی شبیه سیستم پست است. اگر چند نامه را به مقصد کشوری دیگر در صندوق پست بیندازید، با کمی شانس همه آنها به دست گیرنده خواهند رسید. البته احتمال دارد که هر یک از این نامه از مسیر متفاوتی به مقصد رسیده باشد، که این موضوع از دید کاربران پنهان است (و علاقه‌ای هم به دانستن آن ندارند). فرمت بسته‌های پیام و پروتکل آنها در لایه اینترنت تعریف می‌شود، که IP (Internet Protocol) نام دارد.



شکل ۱-۲۱. مدل مرجع TCP/IP.

وظیفه لایه اینترنت اینست که بسته‌های IP را به مقصد برساند. مسیریابی بسته‌ها (و جلوگیری از ازدحام در مسیره‌های شلوغ) برعهده این لایه است، که از این نظر می‌توان آنرا معادل لایه شبکه در مدل OSI دانست (شکل ۱-۲۱ را ببینید).

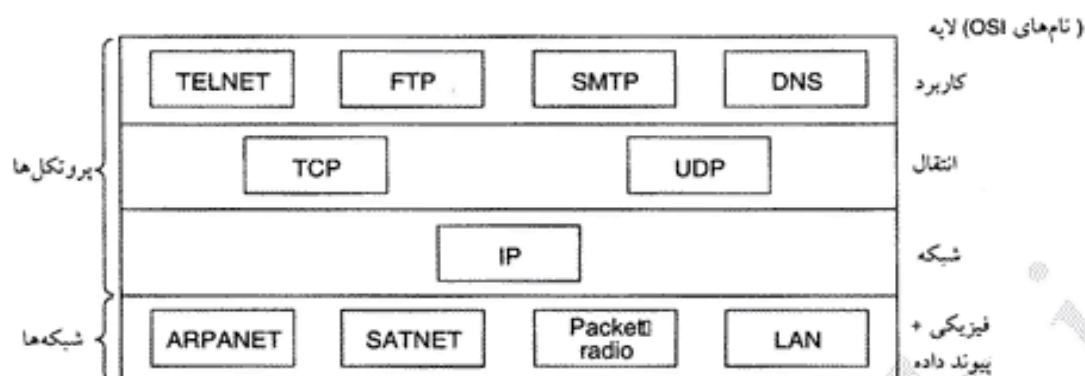
لایه انتقال

لایه بالای لایه اینترنت در مدل TCP/IP، لایه انتقال (transport layer) نام دارد. این لایه شبیه لایه انتقال در مدل OSI است، و اجازه می‌دهد تا عناصر هم‌تا در کامپیوترهای مبدأ و مقصد با هم مکالمه انجام دهند. در این لایه دو پروتکل انتقال نقطه-به-نقطه تعریف شده است. پروتکل اول، که TCP (Transmission Control Protocol) نام دارد، یک پروتکل اتصال-گرای قابل اعتماد است که اجازه می‌دهد تا جریانی از بایتهای بدون خطا از یک کامپیوتر در اینترنت به کامپیوتر دیگر فرستاده شود. این پروتکل جریان بایتهای را بصورت بسته بسته در آورده، و به لایه اینترنت تحویل می‌دهد. در ماشین مقصد عکس این عمل انجام می‌شود: بسته‌ها به هم چسبانده شده، و بصورت جریانی از بایتهای به لایه بالاتر فرستاده می‌شود. در پروتکل TCP کنترل جریان داده‌ها (flow control) نیز وجود دارد، بدین معنا که فرستنده داده‌ها را سریعتر از آنچه گیرنده توان دریافت آنرا دارد، ارسال نخواهد کرد.

پروتکل دوم این لایه، که UDP (User Datagram Protocol) نام دارد، یک پروتکل غیرمتصل غیر قابل اعتماد است، که در مواردی که نیازی به سخت‌گیریهای TCP نیست از آن استفاده می‌شود. این پروتکل بیشتر در مواردی که سرعت اهمیت بیشتری دارد تا دقت (مانند انتقال صوت و تصویر)، یا در جاهاییکه فرآیند درخواست-پاسخ فقط یک بار انجام می‌شود، بکار می‌رود. در شکل ۱-۲۲ رابطه پروتکل‌های IP، TCP و UDP را مشاهده می‌کنید. پروتکل IP اکنون در شبکه‌های بسیاری پیاده‌سازی شده است.

لایه کاربرد

مدل TCP/IP لایه‌های نشست یا نمایش ندارد، یعنی در واقع معتقد است که نیازی به آنها نیست. تجربه مدل OSI نیز نشان می‌دهد که این نظر درست است، و این دو لایه بندرت کاربردی پیدا کرده‌اند. در بالای لایه انتقال لایه کاربرد (application layer) قرار می‌گیرد، که تمام پروتکل‌های سطح بالا در آن قرار دارند. پروتکل‌های ترمینال مجازی (TELNET)، انتقال فایل (FTP) و پست الکترونیک (SMTP) از پروتکل‌هایی هستند که از سالها قبل در این لایه پیاده‌سازی شده‌اند (شکل ۱-۲۲). پروتکل ترمینال مجازی اجازه



شکل ۱-۲۲. پروتکل‌ها و شبکه‌ها در مدل TCP/IP.

می‌دهد تا کاربر وارد کامپیوترهای راه دور شده، و با آنها مانند یک کامپیوتر محلی کار کند. پروتکل انتقال فایل نیز ابزار است مؤثر برای انتقال اطلاعات از یک ماشین به ماشین دیگر. پست الکترونیک در ابتدا چیزی بیش از یک انتقال فایل ساده نبود، ولی بعدها یک پروتکل خاص بنام SMTP برای آن توسعه داده شد. اکنون پروتکل‌های معروف دیگری نیز در این لایه وجود دارند، که برخی از آنها عبارتند از: پروتکل نام ناحیه (DNS) برای ترجمه نام کامپیوترها به آدرس شبکه، پروتکل انتقال خبر (NNTP) برای خواندن مقالات پوزنت (USENET)، پروتکل انتقال صفحات ابرمتن (HTTP) برای خواندن صفحات وب، و دهها پروتکل دیگر.

لایه میزبان - به - شبکه

در زیر لایه اینترنت یک شکاف بزرگ دیده می‌شود؛ در واقع مدل TCP/IP درباره این قسمت تا حد زیادی سکوت کرده است، و فقط انتظار دارد که میزبان بنحوی به شبکه وصل شده، و بتواند بسته‌های IP را ارسال کند. پروتکل انجام این کار در مدل TCP/IP تعریف نمی‌شود، و در موارد مختلف متفاوت است (حتی کتابها و مقالاتی که درباره TCP/IP نوشته شده‌اند، بندرت در این باره صحبت می‌کنند).

۳-۴-۱ مقایسه مدل‌های OSI و TCP/IP

مدل‌های TCP/IP و OSI نقاط مشترک زیادی دارند. هر دوی آنها مبتنی بر مجموعه‌ای از پروتکل‌های مستقل هستند، و عملکرد لایه‌ها نیز تا حد زیادی شبیه یکدیگر است. برای مثال، در هر دو مدل لایه‌های بالای لایه انتقال (و از جمله خود آن) بصورت نقطه-به-نقطه عمل می‌کنند، مستقل از شبکه هستند، و سرویس‌های خود را (به شکلی کاربرد-گرا) در اختیار لایه‌های بالاتر می‌گذارند.

علیرغم شباهت‌های اساسی، این دو مدل تفاوت‌های بسیاری نیز با هم دارند، که در این قسمت به آنها خواهیم پرداخت. شایان ذکر است که ما در اینجا مدل‌ها را با هم مقایسه می‌کنیم، نه مجموعه پروتکل‌های آنها را (در این باره نیز بعداً صحبت خواهیم کرد). برای یک مقایسه کامل و جامع بین TCP/IP و OSI به کتاب (Piscitello and Chapin, 1993) مراجعه کنید.

در مدل OSI سه مفهوم محوری وجود دارد:

۱. سرویس (service)
۲. واسط (interface)
۳. پروتکل (protocol)

شاید بزرگترین دستاورد مدل OSI روشن ساختن مفاهیم فوق (و تفکیک آنها) باشد. هر لایه سرویس‌هایی در اختیار لایه‌های بالاتر از خود قرار می‌دهد. تعریف این سرویسها فقط می‌گوید که یک لایه چه کاری انجام می‌دهد، و هیچ حرفی درباره نحوه انجام آنها و چگونگی استفاده از سرویسها نمی‌زند.

تعریف چگونگی دسترسی به سرویسهای یک لایه بر عهده واسطه است. واسطه پارامترهای ورودی لازم، و نتیجه‌ای را که باید منتظر آن باشید، تعریف می‌کند. حتی واسطه هم نمی‌گوید که یک لایه چگونه کار خود را انجام می‌دهد.

و بالاخره، کاری که لایه انجام می‌دهد را پروتکل‌های آن لایه تعریف می‌کنند. یک لایه مادامیکه کارش را بدرستی انجام دهد، می‌تواند از هر پروتکلی استفاده کند. تغییر پروتکل‌های یک لایه هیچ تأثیری روی ارتباط آن با لایه‌های بالاتر نخواهد گذاشت.

ایده‌های فوق بسیار شبیه مفاهیم مدرن برنامه‌نویسی شیء‌گرا هستند. هر شیء، مانند یک لایه، متدها (عملکردها)ی دارد که اشیاء دیگر از آنها استفاده می‌کنند. نحوه استفاده از این متدها در واقع همان سرویسها نیست که این شیء در اختیار دیگران می‌گذارد. ورودیها و خروجیهای شیء واسطه آن با دنیای خارج هستند. کد اجرایی شیء نیز شبیه همان پروتکل است، که نحوه عملکرد آن از دید دیگران مخفی است.

در مدل اولیه TCP/IP تمایز بین سرویسها، واسطه‌ها و پروتکلها واضح و مشخص نبود، اگرچه افرادی (با توجه به تجربه موفق OSI) سعی کرده بودند آنرا هر چه بیشتر شبیه OSI کنند. برای مثال، لایه اینترنت فقط دو سرویس واقعی بنامهای SEND IP PACKET و RECEIVE IP PACKET داشت. با توجه به این وضع، پروتکل‌های OSI بهتر از TCP/IP مخفی شده‌اند، و امکان تغییر آنها براحتی وجود دارد، چیزی که هدف غایی طراحی لایه‌ای محسوب می‌شود.

مدل OSI قبل از اختراع پروتکل‌های آن طراحی و ابداع شد. این بدان معناست که مدل OSI وابستگی و تمایل خاصی به هیچ مجموعه پروتکلی ندارد، چیزی که در سایر مدلها بسیار دیده می‌شود. البته این وضعیت یک نقطه ضعف نیز دارد، و آن اینست که طراحان تجربه‌چندانی در زمینه موضوع کار ندارند، و واقعاً نمی‌دانند کدام عملکرد را باید در کدام لایه قرار دهند. برای مثال، لایه پیوند داده در ابتدا فقط برای شبکه‌های نقطه-به-نقطه طراحی شده بود، و وقتی شبکه‌های پهنی وارد بازار شد، مجبور شدند یک زیرلایه به آن اضافه کنند.

وقتی افراد شروع به طراحی شبکه با استفاده از مدل OSI و پروتکل‌های موجود کردند، بزودی دریافتند که این شبکه‌ها با سرویسهای موردنیاز انطباق ندارند!، بنابراین مجبور شدند زیرلایه‌های زیادی به آن وصله‌پینه کنند. بالاخره، کمیته استاندارد مقرر کرد که هر کشور برای خود یک شبکه منطبق با مدل OSI (تحت نظارت دولت) داشته باشد - شبکه‌ای که به هیچ عنوان آینده (اینترنت) در آن دیده نشده بود. خلاصه، کارها آنطوری که انتظار داشتند از آب در نیامد.

در مورد TCP/IP وضع بر عکس بود: اول پروتکلها اختراع و توسعه داده شدند، و سپس مدلی برای توصیف آنها ساخته شد. هیچ مشکلی در زمینه انطباق پروتکلها با مدل وجود نداشت؛ همه چیز جفت و جور بود. تنها مشکل این بود که این مدل با هیچ مجموعه پروتکل دیگری جور در نمی‌آمد. این بدان معنا بود که مدل TCP/IP بدرد توصیف شبکه‌های غیر TCP/IP نمی‌خورد.

جدای از مسائل فلسفی قضیه، تفاوت دیگر در تعداد لایه‌های این دو مدل است: مدل OSI هفت لایه دارد و مدل TCP/IP چهار لایه. لایه‌های شبکه، انتقال و کاربرد در هر دو مشترکند، ولی لایه‌های دیگر فرق دارند.

تفاوت دیگر در زمینه ارتباطات اتصال-گرا و غیرمتصل است. مدل OSI از هر دو نوع ارتباط اتصال-گرا و غیرمتصل در لایه شبکه پشتیبانی می‌کند، ولی در لایه انتقال فقط سرویس اتصال-گرا دارد (چون این سرویس در

معرض دید کاربران است). مدل TCP/IP در لایه شبکه فقط سرویس غیر متصل دارد، ولی در لایه انتقال از هر دو نوع ارتباط پشتیبانی می‌کند، و دست کاربر را برای انتخاب باز می‌گذارد (که بویژه برای پروتکل‌های ساده درخواست-پاسخ بسیار مهم است).

۱.۴.۱ نگاهی انتقادی به مدل OSI و پروتکل‌های آن

مدل‌های OSI و TCP/IP (و پروتکل‌هایشان) هیچکدام کامل نیستند، و جا دارد برخی از نقاط ضعف آنها را برشماریم. در این قسمت و قسمت آینده، برخی از نقاط ضعف مهم مدل‌های OSI و TCP/IP را بررسی خواهیم کرد. با مدل OSI شروع می‌کنیم.

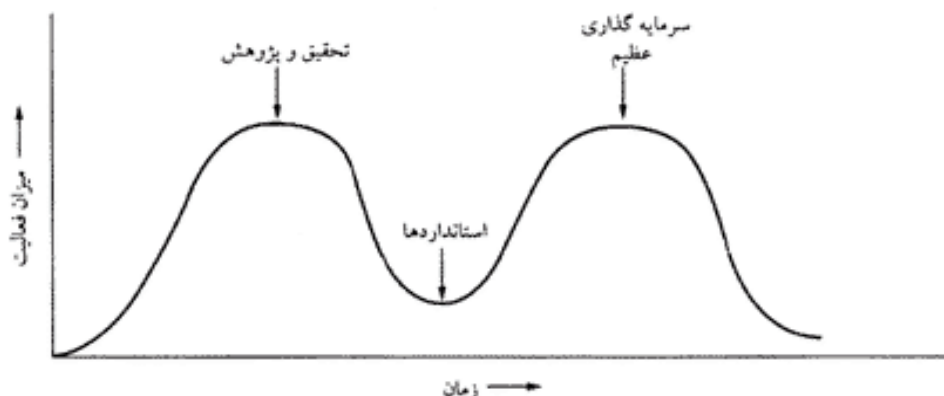
در زمان چاپ ویرایش دوم این کتاب (سال ۱۹۸۹)، بسیاری از متخصصان برجسته شبکه بر این باور بودند که آینده در بست متعلق به مدل OSI و پروتکل‌های آن است، و هیچ چیز نمی‌تواند در مقابل پیشرفت آن مقاومت کند. اما این اتفاق نیفتاد. چرا؟ نگاهی به گذشته درسهای بسیاری برای چشمان عبرت‌بین دارد، که می‌توان آنها را چنین خلاصه کرد:

۱. زمان نامناسب
۲. تکنولوژی نامناسب
۳. پیاده‌سازی نامناسب
۴. سیاست‌های نامناسب

زمان نامناسب

اولین عامل شکست مدل OSI زمان نامناسب بود. زمانی که یک استاندارد وضع می‌شود، اهمیت حیاتی در موفقیت یا عدم موفقیت آن دارد. دیوید کلارک از دانشگاه M.I.T فرضیه‌ای در زمینه استانداردها دارد که به ملاقات فیل‌ها معروف است، و در شکل ۱-۲۳ آنرا مشاهده می‌کنید.

این شکل میزان فعالیتهای حول یک موضوع جدید را نشان می‌دهد. وقتی موضوعی برای اولین بار کشف می‌شود، گرداگرد آنرا سبیلی از فعالیتهای تحقیقی (به شکل بحث، مقاله و سخنرانی) فرا می‌گیرد. بعد از مدتی این موج فروکش می‌کند، و بعد از اینکه صنعت به آن موضوع علاقمند شد، موج سرمایه‌گذاری‌ها از پی می‌آید. بسیار مهم است که در نقطه تلاقی این دو فیل (موج تحقیق و موج سرمایه‌گذاری) استانداردها بطور کامل وضع شوند. اگر استاندارد زودتر از موعد (قبل از پایان تحقیقات) نوشته شود، خطر آن هست که موضوع بدرستی



شکل ۱-۲۳. فرضیه ملاقات فیل‌ها

درک نشده باشد، و استاندارد ضعیف از آب در آید. اگر استاندارد دیرتر از موعد (بعد از شروع موج سرمایه‌گذاری) نوشته شود، شرکت‌های بسیاری قبلاً - در مسیرهای مختلف - در آن سرمایه‌گذاری کرده‌اند، و این خطر هست که استاندارد آنها را نادیده بگیرد. اگر فاصله این دو فیل خیلی کم باشد (همه عجله داشته باشند که زودتر کار را شروع کنند)، خطر آن هست که استاندارد نویسان بین آنها له شوند.

اکنون معلوم شده است که پروتکل‌های استاندارد OSI بین فیله‌ها شدند. وقتی پروتکل‌های OSI پایه عرصه وجود گذاشتند، پروتکل‌های رقیب (TCP/IP) مدت‌ها بود که در دانشگاه‌ها و مراکز تحقیقاتی پذیرفته شده بودند. با اینکه هنوز موج سرمایه‌گذاری صنعتی در TCP/IP شروع نشده بود، اما بازار آکادمیک آنقدر بزرگ بود که شرکت‌های بسیاری را تشویق به تولید محصولات TCP/IP کند. و وقتی OSI بالاخره از راه رسید، کسی نبود که داوطلبانه از آن پشتیبانی کند. همه منتظر بودند دیگری قدم اول را بردارد؛ قدمی که هرگز برداشته نشد، و OSI در نطفه خفه شد.

تکنولوژی نامناسب

دلیل دیگری که OSI هرگز با نگرش آن بود که، این مدل و پروتکل‌های آن هر دو ناقص و معیوب بودند. انتخاب هفت لایه برای این مدل بیشتر یک انتخاب سیاسی بود تا فنی، و در حالیکه دو لایه آن (نشست و نمایش) تقریباً خالی بودند، در لایه‌های دیگر (لینک داده و شبکه) جای نفس کشیدن نبود.

مدل OSI (و سرویس‌ها و پروتکل‌های آن) بطرزی باورنکردنی پیچیده است. اگر کاغذهای چاپی این استاندارد را روی هم بچینید، ارتفاع آن از نیم متر هم بیشتر خواهد شد! پیاده‌سازی پروتکل‌های OSI بسیار دشوار، و عملکرد آنها ناقص است. در این رابطه، نقل جمله جالبی از پاول موکاپتریس (Rose, 1993) خالی از لطف نیست:

سؤال: از ترکیب یک گانگستر با یک استاندارد بین‌المللی چه چیزی بدست می‌آید؟

جواب: کسی پیشنهادی به شما می‌کند که از آن سر در نمی‌آورید.

مشکل دیگر مدل OSI، علاوه بر غیر قابل فهم بودن آن، اینست که برخی از عملکردهای آن (مانند آدرس دهی، کنترل جریان داده‌ها، و کنترل خطا) در تمام لایه‌ها تکرار می‌شود. برای مثال، سالترز و همکارانش (1984) نشان دادند که کنترل خطا باید در بالاترین لایه انجام شود تا بیشترین تأثیر را داشته باشد، بنابراین تکرار آن در لایه‌های پائینتر نه تنها غیر ضروری است، بلکه باعث افت کارایی هم خواهد شد.

پیاده‌سازی نامناسب

با توجه به پیچیدگی بیش از حد مدل OSI و پروتکل‌های آن، جای تعجب نبود که اولین پیاده‌سازیهای آن حجیم، سنگین و کند باشد. آنهایی که با این مدل کار کرده بودند، بزودی پشیمان شدند، و طولی نکشید که کلمه OSI مترادف شد با «کیفیت بد». بعدها محصولات بهتری به بازار آمد، اما آوازه منفی OSI فراموش نشد.

از طرف دیگر، اولین پیاده‌سازی TCP/IP (که بخشی از سیستم عامل یونیکس برکلی بود) بسیار خوب از کار درآمد (و لازم به گفتن نیست که مجانی هم بود). افراد بسیاری سرعت شروع به استفاده از آن کردند، هواخواه آن شدند، آنرا توسعه دادند، و این باعث شد که باز هم به خیل طرفداران آن اضافه شود. در اینجا، برخلاف OSI، ماریج رو به بالا می‌رفت، نه پائین.

سیاست‌های نامناسب

بدلیل اولین پیاده‌سازی TCP/IP، بسیاری از افراد (بویژه در محیط‌های دانشگاهی) تصور می‌کردند که TCP/IP جزئی از یونیکس است، و یونیکس هم در آن دوران محبوبیتی فوق‌العاده داشت.

از سوی دیگر، این عقیده رواج داشت که OSI یک مخلوق دولتی (مخصوصاً دولتهای اروپایی و آمریکایی) است البته این عقیده فقط تا حدی درست بود، اما همین تصور هم که عده‌ای دیوانسالار دولتی بخواهد یک

استاندارد فنی را بزور جا بیندازند، باعث شد تا برنامه‌نویسان و طراحان شبکه تمایلی به همکاری از خود نشان ندهند. زبانهای برنامه‌نویسی PL/1 (که در دهه ۱۹۶۰ از سوی IBM بعنوان زبان آینده توسعه داده شد) و Ada (که وزارت دفاع آمریکا حامی آن بود) بهمین دلیل دچار سرنوشتی مشابه شدند.

۵-۴-۱ نگاهی انتقادی به مدل TCP/IP

مدل TCP/IP و پروتکل‌های آن نیز مشکلات خاص خود را دارند. اول اینکه، در این مدل مفاهیم سرویس، واسط و پروتکل بروشنی از یکدیگر تفکیک نشده‌اند؛ کاری که در مدل OSI بخوبی انجام شده است. به همین دلیل نمی‌توان از TCP/IP بعنوان ابزاری برای طراحی و توسعه شبکه‌های جدید استفاده کرد.

دوم اینکه، مدل TCP/IP به هیچ عنوان یک مدل کلی نیست، و نمی‌توان از آن برای توصیف شبکه‌های غیر TCP/IP استفاده کرد. برای مثال، توصیف بلوتوث با مدل TCP/IP بکلی غیرممکن است.

سوم اینکه، با در نظر گرفتن مفاهیم شبکه‌های چند لایه، لایه میزبان-به-شبکه اساساً یک لایه واقعی نیست، بلکه فقط یک واسط (بین لایه‌های شبکه و لینک داده) است. در واقع، این یکی از مهمترین جاهانیست که مدل TCP/IP مفاهیم واسط و لایه را با هم قاطی کرده است.

چهارم اینکه، در مدل TCP/IP هیچ تمایزی بین لایه‌های فیزیکی و لینک داده نیست (و حتی حرفی از آنها بمیان نیامده است). اینها دو لایه کاملاً متفاوت هستند - لایه فیزیکی با مشخصات کابل و فیبر نوری و کانالهای مخابراتی سروکار دارد، در حالیکه وظیفه لایه پیوند داده شکستن داده‌ها به قطعات کوچکتر و اطمینان از تحویل صحیح آنها به مقصد است. در یک مدل کامل این دو لایه باید از هم جدا باشند؛ کاری که در مدل TCP/IP انجام نشده است.

و بالاخره اینکه، اگر چه پروتکل‌های TCP و IP بسیار خوب طراحی و پیاده‌سازی شده‌اند، بسیاری از دیگر پروتکل‌های این مدل چنین نیستند، و اغلب توسط دانشجویان کنجکاو (و در ساعات بیکاری) نوشته شده‌اند. این پروتکلها بعلت انتشار سریع (که اغلب دلیلی جز مجانی بودن ندارد) بسرعت جامی افتند، و بهمین دلیل جایگزین کردن آنها بسیار دشوار می‌شود. برخی از این پروتکلها امروز چیزی جز شرمساری نیستند. مثلاً، پروتکل TELNET اساساً برای ترمینالهای کُند و سنگین تله‌تایپ نوشته شد، و هیچ نشانی از گرافیک و ماوس در آن نیست. اما، بعد از ۲۵ سال خیلی‌ها همچنان از آن استفاده می‌کنند.

بطور خلاصه، مدل OSI (علیرغم برخی از مشکلات آن، و منهای لایه‌های نشست و نمایش) ثابت کرده که بهترین ابزار برای توصیف شبکه‌های کامپیوتریست - اما متأسفانه همچنان روی کاغذ باقی مانده است. از طرف دیگر، با اینکه چیزی بنام مدل TCP/IP در واقع وجود خارجی ندارد، اما پروتکل‌های آن در مقیاس وسیعی مورد استفاده قرار می‌گیرند. از آنجائیکه در دنیای کامپیوتر هر کسی حرف خود را می‌زند، ما هم در این کتاب از یک مدل اصلاح شده OSI استفاده خواهیم کرد، ولی توجه خود را بیشتر روی TCP/IP و پروتکل‌های وابسته به آن معطوف می‌کنیم - البته درباره پروتکل‌های جدید مانند SONET، 802 و بلوتوث هم صحبت خواهیم کرد. در حقیقت، مدلی که ما در این کتاب از آن استفاده خواهیم کرد، یک مدل ترکیبی (شکل ۱-۲۴) است.

5	لایه کاربرد
4	لایه انتقال
3	لایه شبکه
2	لایه پیوند داده
1	لایه فیزیکی

شکل ۱-۲۴. یک مدل مرجع ترکیبی، که در این کتاب از آن استفاده خواهد شد.

۵-۱ شبکه های نمونه

بحث شبکه های کامپیوتری انواع مختلفی از شبکه ها (کوچک و بزرگ، شناخته شده و مهجور) را در بر می گیرد. این شبکه ها در اندازه، کاربرد و تکنولوژیهای بکار رفته با هم متفاوتند. برای آن که تصویری از تنوع موجود در شبکه های کامپیوتری داشته باشید، در این قسمت نگاهی به چند نمونه از آنها خواهیم انداخت.

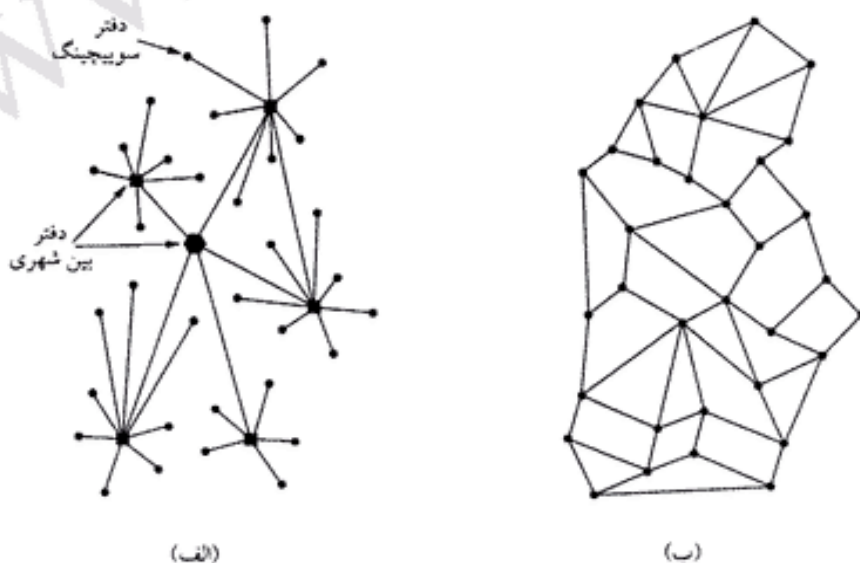
این قسمت را با معرفی اینترنت (که احتمالاً شناخته شده ترین شبکه هاست) شروع می کنیم، و نگاهی به تاریخچه، سیر تکامل و تکنولوژی آن خواهیم داشت. پس از آن به سراغ ATM (که تفاوت های چشمگیری با اینترنت دارد، و حتی می توان گفت ضد آن است) می رویم. و بالاخره، نگاهی به IEEE 802.11 (استاندارد شبکه های محلی بیسیم) می اندازیم.

۱-۵-۱ اینترنت

اینترنت (Internet) در واقع اصلاً یک شبکه نیست، بلکه مجموعه ایست از شبکه های مختلف که از پروتکل های خاصی استفاده کرده، و سرویس های مشخصی را ارائه می کند. ویژگی غیر عادی اینترنت اینست که توسط فرد خاصی طراحی نشده، و هیچکس هم آنرا کنترل نمی کند. برای درک بهتر این مطلب، اجازه دهید بینیم اینترنت از کجا شروع شد، و علت آن چه بود. یکی از جالبترین تاریخچه های اینترنت را می توانید در کتاب جان نافتون - 2000 ببینید. این کتاب نه تنها برای افراد عادی، بلکه برای مورخان نیز جالب است؛ برخی از مطالب ذیل از این کتاب اقتباس شده است. البته کتاب های فنی بیشتری نیز درباره اینترنت و پروتکل های آن نوشته شده، که از آن میان می توان به (Maufer, 1999) اشاره کرد.

آرپانت (ARPANET)

داستان ما از اواخر دهه ۱۹۵۰ شروع می شود. در اوج جنگ سرد، وزارت دفاع ایالات متحده آمریکا به فکر ایجاد یک شبکه فرماندهی و کنترل افتاد که بتواند حتی در مقابل حملات هسته ای دوام بیاورد. در آن زمان تمامی مخابرات نظامی به شبکه تلفن عمومی متکی بود، که مستعد آسیب تشخیص داده شده بود. با یک نگاه به شکل ۱-۲۵ (الف) می توانید مبنای این استدلال را دریابید. در این شکل نقاط سیاه نماینده مراکز سونیچینگ شهری



شکل ۱-۲۵. (الف) ساختار شبکه تلفن. (ب) طرح بارن برای یک سیستم سونیچینگ توزیع شده.

هستند که هزاران خط تلفن از آنها منشعب می‌شود. این مراکز نیز بنوبه خود به مراکز بین شهری بزرگتر متصل هستند، که در مجموع شبکه تلفن کشوری را می‌سازند. آسیب‌پذیری این سیستم از آنجا ناشی می‌شد که تخریب چند مرکز بین شهری کلیدی می‌توانست تماس تلفنی را در کل کشور مختل کند.

در سال ۱۹۶۰ وزارت دفاع قراردادی را با شرکت راند (RAND Corporation) امضا کرد، که در آن وظیفه یافتن یک راه‌حل به آن محول شده بود. یکی از متخصصان این شرکت، بنام پل بارن (Paul Baran)، طرح یک شبکه توزیع شده (distributed) و تحمل‌پذیر خطا (fault-tolerant) را پیشنهاد کرد، که آنرا در شکل ۱-۲۵ (ب) می‌بینید. از آنجائیکه در این شبکه طول مسیر بین مراکز سونیچینگ طولانیتر از آن بود که بتوان از سیگنالهای آنالوگ استفاده کرد، بارن پیشنهاد کرد در این سیستم از تکنولوژی سونیچینگ بسته دیجیتال (digital packet-switching) استفاده شود. بارن گزارشات متعددی برای وزارت دفاع نوشت، و جزئیات سیستم پیشنهادی خود را تشریح کرد. مقامات رسمی پتاگون به ایده نهفته در این سیستم علاقمند شدند، و از AT&T (که در آن زمان انحصار شبکه تلفن کشوری را در دست داشت) خواستند که یک نمونه اولیه از آن بسازد. AT&T طرح بارن را رد کرد؛ بزرگترین و ثروتمندترین شرکت دنیا تحمل نمی‌کرد که یک جوان تازه از راه رسیده به آنها بگوید چگونه شبکه تلفن بسازند! آنها ادعا کردند که طرح بارن قابل اجرا نیست، و بدین ترتیب ایده آن را در نطفه خفه کردند.

سالها گذشت، و وزارت دفاع همچنان بدنبال سیستم فرماندهی و کنترل ایده‌آل خود بود. برای درک بهتر اتفاقات بعدی، باید کمی به عقب برگردیم: به اکتبر ۱۹۵۷، زمانی که اتحاد جماهیر شوروی (سابق) با پرتاب اولین قمر مصنوعی بنام اسپوتنیک در مسابقه فضایی از ایالات متحده پیشی گرفت. آیزنهاور، رئیس جمهور وقت ایالات متحده، در جستجو برای یافتن علت عقب‌افتادگی کشورش، با وحشت دریافت که نیروهای زمینی، دریایی و هوایی آمریکا مشغول دعا بر سر تقسیم بودجه تحقیقاتی پتاگون هستند. وی بلافاصله تصمیم گرفت که یک مرکز واحد برای تحقیقات نظامی بوجود آورد؛ مرکزی که آرپا (آژانس پروژه‌های تحقیقاتی پیشرفته Advanced Research Projects Agency - ARPA) نام گرفت. آرپا هیچ دانشمند یا آزمایشگاهی نداشت؛ در واقع، آرپا چیزی نبود جز یک دفتر هماهنگی کوچک با بودجه‌ای ناچیز (البته با معیارهای پتاگون). آرپا کارش را با عقد قرارداد با واگذاری امتیاز به شرکتها یا دانشگاههایی که ایده‌های جالبی داشتند، انجام می‌داد.

در سالهای اول، آرپا بیشتر سعی داشت خطوط کلی مأموریت خود را روشن و ترسیم کند، ولی در سال ۱۹۶۷ توجه مدیرعامل آن، لاری رابرتس، به موضوع شبکه جلب شد. او با متخصصان بسیاری مشورت کرد؛ و یکی از همین متخصصان، بنام وسلی کلارک، بود که پیشنهاد ایجاد یک زیرشبکه سونیچینگ بسته را مطرح کرد (شکل ۱-۱۰).

بعد از مقداری بحثهای اولیه، رابرتس ایده را پسندید و آنرا طی یک مقاله نسبتاً مبهم به گردهمایی اصول سیستم عامل (که در اواخر ۱۹۶۷ در گاتلین‌بورگ، تنسی برگزار شده بود) ارائه کرد (Roberts, 1967). در میان ناباوری رابرتس، مقاله دیگری نیز به این کنفرانس ارائه شده بود که نه تنها سیستم مشابهی را توصیف می‌کرد، بلکه حتی صحبت از پیاده‌سازی آن تحت مدیریت فردی بنام دونالد دیویس در آزمایشگاه ملی فیزیک (NPL) در انگلستان بمیان آمده بود. سیستم NPL در واقع سیستمی در سطح ملی نبود، بلکه فقط چند کامپیوتر را در محوطه NPL به هم متصل می‌کرد، اما نکته مهم این بود که نشان می‌داد سونیچینگ بسته در عمل کار می‌کند. از همه جالبتر اینکه، سیستم NPL بر اساس کارهای بارن پایه‌گذاری شده بود. وقتی رابرتس از گاتلین‌بورگ برگشت، دیگر مصمم بود چیزی را بسازد که بعدها به آرپانت (ARPANET) معروف شد.

این زیرشبکه تعدادی مینی‌کامپیوتر بنام (Interface Message Processor) IMP را با خطوط انتقال

56-kbps به هم متصل می کرد. برای رسیدن به قابلیت اعتماد بالا، هر IMP به حداقل دو IMP دیگر متصل می شد. این زیرشبکه در واقع یک زیرشبکه دیتاگرام (datagram subnet) بود، بنابراین اگر تعدادی از خطوط یا IMP ها از بین می رفتند، پیامها می توانستند از طریق مسیرهای جایگزین به مقصد برسند.

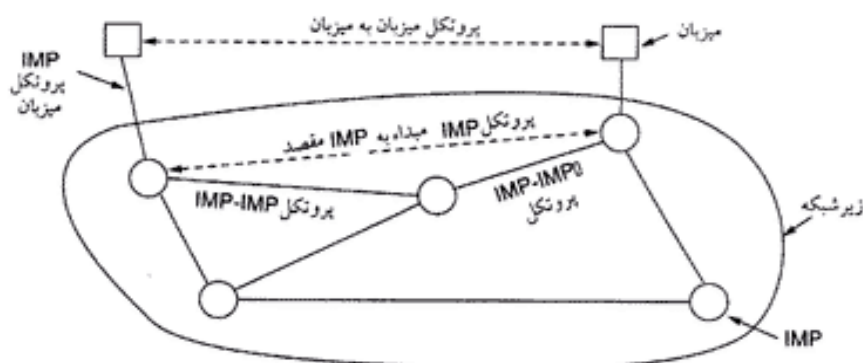
هر گره (node) این شبکه عبارت بود از یک کامپیوتر میزبان و یک IMP، که با سیمی کوتاه به هم وصل می شدند. هر میزبان می توانست پیامهایی تا سقف ۸۰۶۳ بیت به IMP خود بفرستد، و این IMP سپس پیام را به بسته های ۱۰۰۸ بیتی شکسته و آنها را بصورت مستقل به سمت مقصد می فرستاد. هر بسته قبل از اینکه به گره بعدی هدایت شود، بایستی بطور کامل دریافت می شد؛ بدین ترتیب، آرپانت اولین زیرشبکه سوئیچینگ بسته بود که بصورت ذخیره-هدایت (store-and-forward) کار می کرد.

پس از آن آرپا مناقصه ای برای ساخت این زیرشبکه اعلام کرد، که دوازده شرکت اسناد آنرا خریدند. بعد از بررسی پیشنهادات رسیده، در دسامبر ۱۹۶۸ آرپا شرکت BBN را (که یک شرکت مشاوره در کمبریج، ماساچوست بود) برای ساخت این زیرشبکه و نوشتن نرم افزارهای آن برگزید. شرکت BBN مینی کامپیوترهای اصلاح شده هانی ول DDP-316 را (که ۱۲ کیلوبایت حافظه ۱۶ بیتی داشت) بعنوان IMP انتخاب کرد. از آنجائیکه قطعات مکانیکی ذاتاً غیر قابل اعتماد فرض می شدند، این IMP ها اصلاً دیسک نداشتند، و با خطوط اجاره ای 56-kbps به هم متصل می شدند. با اینکه امروزه حتی بچه ها هم دیگر خطوط 56-kbps را قبول ندارند (و به کمتر از ADSL راضی نمی شوند)، آنروزها خطوط 56-kbps بالاترین چیزی بود که می شد آرزو کرد.

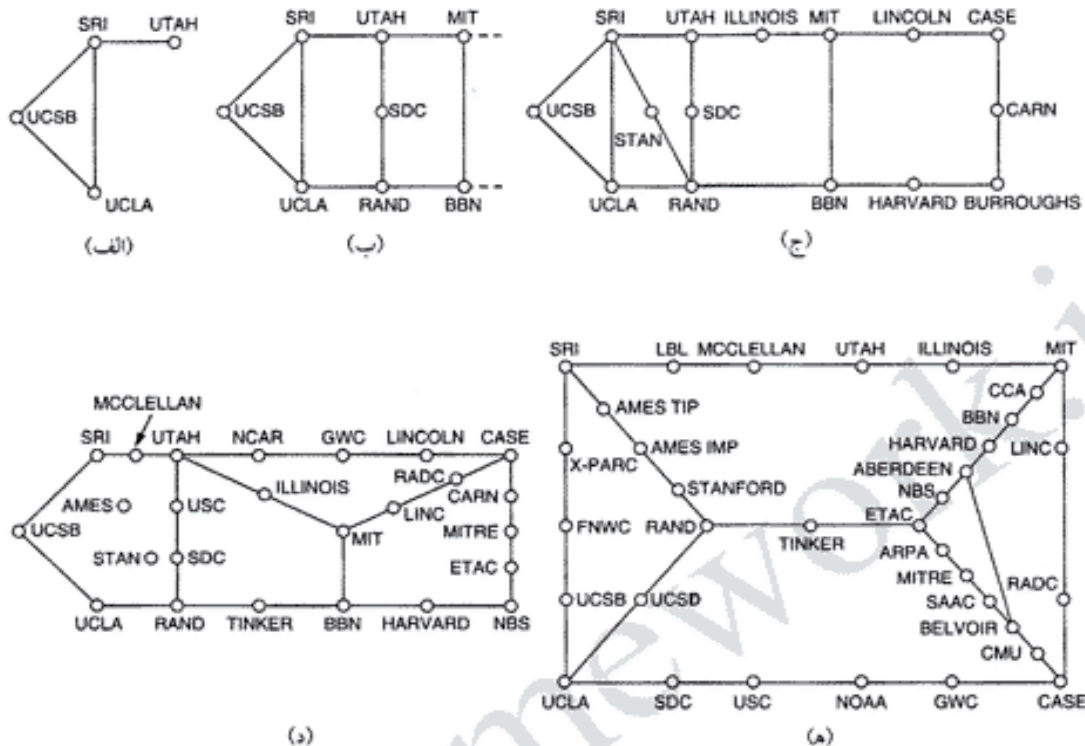
نرم افزار نیز در دو بخش مجزا طراحی شد: زیرشبکه، میزبان. نرم افزار زیرشبکه عبارت بود از پروتکل ارتباط IMP با میزبان، پروتکل IMP-IMP، و نرم افزاری برای بهبود ارتباط IMP مبدأ با IMP مقصد. در شکل ۱-۲۶ طراحی اولیه آرپانت را ملاحظه می کنید.

در خارج از زیرشبکه، میزبانها نیز به نرم افزار نیاز داشتند: پروتکل ارتباط میزبان با IMP، پروتکل میزبان-میزبان، و نرم افزارهای کاربردی. بزودی معلوم شد که BBN احساس می کند با گرفتن پیام در نقطه واسط میزبان-IMP، و تحویل آن در نقطه IMP-میزبان سمت مقابل کارش پایان یافته است.

اما رابرتس مشکل دیگری داشت: کامپیوترهای میزبان هم نیازمند نرم افزار بودند. برای حل این مشکل، در تابستان ۱۹۶۹ رابرتس همایشی از متخصصان شبکه (که عمدتاً دانشجویان تازه فارغ التحصیل بودند) در اسنوپرد، یوتا تشکیل داد. این دانشجویان فکر می کردند کسی وجود دارد که طرح کلی شبکه را برای آنها توضیح دهد، و بعد از آن می توانستند نوشتن نرم افزار را شروع کنند. آنها بسیار شگفت زده شدند وقتی فهمیدند که نه متخصصی برای توضیح طرح شبکه وجود دارد، و نه اساساً چیزی بنام طرح شبکه! آنها دریافتند که باید کار را از صفر شروع کنند.



شکل ۱-۲۶. طراحی اولیه آرپانت.



شکل ۱-۲۷. مراحل رشد آرپانت. (الف) دسامبر ۱۹۶۹. (ب) ژوئیه ۱۹۷۰. (ج) مارس ۱۹۷۱. (د) آوریل ۱۹۷۲. (ه) سپتامبر ۱۹۷۲.

با وجود همه این مشکلات، بالاخره آرپا موفق شد در دسامبر ۱۹۶۹ یک شبکه آزمایشی متشکل از چهار گره (دانشگاههای UCSB، UCLA، SRI و یوتا) راه اندازی کند. علت انتخاب این چهار دانشگاه آن بود که همگی آنها قراردادهای متعددی با آرپا داشتند، و از طرف دیگر (صرفاً برای زورآزمایی فنی) کامپیوترهای آنها بکلی با هم ناسازگار بود. با نصب IMP های جدید این شبکه گسترش یافت، و بزودی سراسر ایالات متحده را تحت پوشش گرفت. در شکل ۱-۲۷ رشد آرپانت را در طی سه سال پس از تولد آن ملاحظه می کنید.

آرپا برای کمک به رشد این نوزاد تازه متولد شده (آرپانت)، در زمینه شبکه های ماهواره ای و مخابرات رادیویی نیز سرمایه گذاریهایی انجام داد. در یک آزمایش معروف، با استفاده از یک شبکه رادیویی پیامهایی از یک کامیون در حال حرکت در جاده های کالیفرنیا به دانشگاه SRI، و از آنجا از طریق آرپانت به ساحل شرقی ایالات متحده فرستاده شد، که سپس از آنجا از طریق شبکه ماهواره ای به دانشگاه کالج در لندن هدایت شد. بدین ترتیب محققانی که در کامیونی در جاده های کالیفرنیا نشسته بودند، توانستند با کامپیوترهایی در لندن کار کنند.

این آزمایش همچنین نشان داد که پروتکل های موجود آرپانت برای کار روی شبکه های مختلف مناسب نیستند. این نتایج منجر به تحقیقات بیشتر روی پروتکلها شد، که با اختراع TCP/IP و پروتکل های آن به اوج رسید (Cerf and Kahn, 1974). مدل TCP/IP بویژه برای ارتباطات روی شبکه های مختلف و ناهمگن (که آرپانت روز بروز به سمت آن حرکت می کرد) طراحی شده بود.

بمنظور تشویق و ترغیب پذیرش این پروتکل های جدید، آرپا قراردادهایی با شرکت BBN و دانشگاه کالیفرنیا در برکلی (UCB) منعقد کرد، تا این پروتکلها را با یونیکس برکلی یکپارچه کنند. محققان برکلی هم کار خود را با

نوشتن برنامه‌های واسط شبکه (که به سوکت - socket - معروف شدند)، و برنامه‌های کاربردی و مدیریتی بنحوی احسن انجام دادند.

زمانه نیز با TCP/IP یار بود؛ بسیاری از دانشگاهها تازه کامپیوترهای جدید VAX را خریده، و آنها را در شبکه‌های LAN به هم متصل کرده بودند، اما هیچ نرم‌افزاری برای شبکه کردن آنها نداشتند. وقتی یونیکس 4.2BSD (با پروتکل‌های TCP/IP، سوکتها و نرم‌افزارهای کمکی خود) بعنوان یک بسته نرم‌افزاری کامل به بازار آمد، بلافاصله مورد قبول جامعه دانشگاهی قرار گرفت. از همه مهمتر اینکه، با TCP/IP می‌شد به آرپانت وصل شد، اتفاقی که بسیاری منتظر آن بودند.

در دهه ۱۹۸۰ شبکه‌های بسیاری (بویژه شبکه‌های محلی) به آرپانت ملحق شدند. با افزایش تعداد کامپیوترهای آرپانت مشکل جدیدی پدید آمد، و آن پیدا کردن یک کامپیوتر در میان خیل عظیم کامپیوترها بود. برای حل این مشکل سیستم نام ناحیه (Domain Name System - DNS) ابداع شد، که نام کامپیوترها را به آدرس IP آنها تبدیل می‌کرد. از آن به بعد، DNS تبدیل به یک پایگاه داده عمومی و توزیع شده شد، که علاوه بر آدرس IP کامپیوترها، اطلاعات دیگری را نیز در اختیار کاربران خود قرار می‌داد. در فصل ۷ درباره DNS مفصلاً صحبت خواهیم کرد.

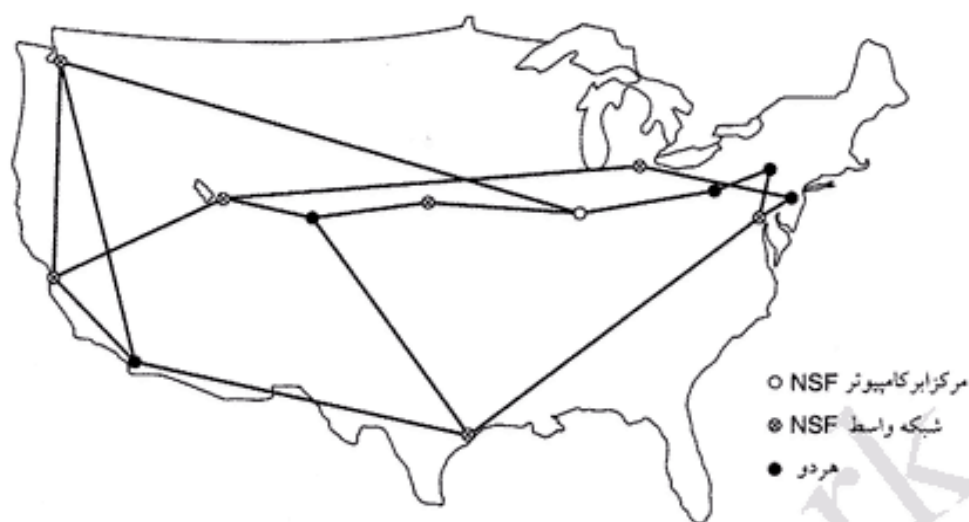
NSFNET

در اواخر دهه ۱۹۷۰ بنیاد ملی علوم ایالات متحده (U.S. National Science Foundation - NSF) شاهد تأثیر روزافزون آرپانت بر تحقیقات دانشگاهی بود. اما هر دانشگاهی که می‌خواست به آرپانت دسترسی داشته باشد، بایستی قراردادی با وزارت دفاع داشته باشد (که بسیاری از آنها نداشتند). پاسخ NSF به این وضعیت، راه‌اندازی شبکه‌ای مشابه آرپانت بود که تمام دانشگاهها به آن دسترسی داشته باشند. بمنظور ایجاد زیربنایی محکم برای این شبکه، NSF با متصل کردن شش ابرکامپیوتر خود در دانشگاههای سان‌دیوگو، بولدر، کامپاین، پیتسبورگ، ایتاکا و پرینستون، یک ستون فقرات (backbone) بوجود آورد. هر یک از این ابرکامپیوترها یک برادر کوچکتر (یک مینی‌کامپیوتر LSI 11، معروف به فازبال) داشت. این فازبال‌ها به خطوط اجاره‌ای 56-kbps متصل بودند، و زیرشبکه را می‌ساختند - یعنی، شبکه NSF از نظر سخت‌افزاری شبیه آرپانت بود. اما، تکنولوژی نرم‌افزاری آن با آرپانت متفاوت بود: فازبال‌ها از همان ابتدا به TCP/IP صحبت می‌کردند، که آنرا تبدیل به این اولین شبکه گسترده TCP/IP می‌کرد.

بعدها NSF تعداد زیادی شبکه منطقه‌ای تأسیس کرد، که به هزاران دانشگاه، آزمایشگاه تحقیقاتی، کتابخانه، و موزه اجازه می‌داد تا به هر یک از ابرکامپیوترهای آن دسترسی داشته باشند، یا اینکه مستقیماً با یکدیگر تماس برقرار کنند. این شبکه (شامل ستون فقرات و شبکه‌های محلی) NSFNET نامیده شد. از طریق لینکی بین یک IMP و یک فازبال در دانشگاه کارنگی-ملون، NSFNET به آرپانت نیز متصل شده بود. اولین ستون فقرات NSFNET را در شکل ۱-۲۸ مشاهده می‌کنید.

NSFNET یک موفقیت آنی بود، و از همان ابتدا با تراکم کاری روبرو شد. NSF بلافاصله به فکر گسترش NSFNET افتاد، و به همین منظور قراردادی با کنسرسیوم MERIT بست. برای ایجاد دومین ستون فقرات. کانالهای فیبرنوری با ظرفیت 448-kbps از MCI (که اکنون در WorldCom ادغام شده است) اجاره شد. برای مسیریاب‌های شبکه نیز از IBM PC-RT استفاده شد. این شبکه نیز بسیار زود با تراکم کاری روبرو شد، و در سال ۱۹۹۰ ظرفیت ستون فقرات آن به 1.5-Mbps ارتقاء داده شد.

با ادامه رشد NSFNET، بزودی NSF متوجه شد که دولت نمی‌تواند برای همیشه به سرمایه‌گذاری در شبکه ادامه دهد. از طرف دیگر، شرکتهای تجاری نیز مایل بودند به شبکه NSFNET ملحق شوند، ولی مقررات NSF



شکل ۱-۲۸. ستون فقرات NSFNET در سال ۱۹۸۸.

کاربردهای انتفاعی شبکه را ممنوع کرده بود. متعاقب آن، NSF بعنوان اولین قدم بسوی تجاری کردن شبکه، شرکت‌های IBM، MERIT و MCI را به ایجاد یک مؤسسه غیرانتفاعی (بنام Advanced Networks and Services - ANS) ترغیب کرد. در سال ۱۹۹۰، ANS کنترل NSFNET را بدست گرفت، و با ارتقاء لینکهای 1.5-Mbps به 45-Mbps شبکه ANSNET را بوجود آورد. این شرکت بعد از ۵ سال کار به AOL (America Online) فروخته شد. اما در آن زمان دیگر شرکت‌های بسیاری سرویسهای تجاری IP ارائه می‌کردند، و روشن شده بود که دولت باید پای خود را از تجارت شبکه بیرون بکشد.

برای تسهیل امور (و اطمینان از اینکه تمام شبکه‌های منطقه‌ای می‌توانند با هم تماس بگیرند)، NSF چهار قرارداد با شرکت‌های بزرگ برای ایجاد نقطه دسترسی شبکه (Network Access Point - NAP) امضا کرد. این چهار شرکت عبارت بودند از: PacBell (در منطقه سانفرانسیسکو)، Ameritech (در منطقه شیکاگو)، MFS (در منطقه واشینگتن دی.سی.)، Sprint (در منطقه نیویورک). هر اپراتور شبکه که بخواهد سرویسهای ستون فقرات به شبکه‌های منطقه‌ای NSF بدهد، بایستی به تمام NAP ها متصل باشد.

بدین ترتیب، هر بسته که بخواهد از یک منطقه به منطقه دیگر برود، می‌تواند از هر یک از این ستونهای فقرات استفاده کند، که نتیجه آن ایجاد رقابت برای سرویس بهتر و قیمت کمتر است. با این تمهید، ستون فقرات منحصر بفرد دولتی جای خود را به یک زیرساخت متنوع و رقابتی داد. بسیاری از افراد دولت فدرال را به گناه عدم خلاقیت سرزنش می‌کنند، ولی در واقع این بنیاد ملی علوم و وزارت دفاع بودند که زیرساخت‌های اینترنت را شکل داده و سپس اداره آنرا به بخش خصوصی سپردند.

در دهه ۱۹۹۰ مناطق و کشورهای بسیاری، با تأثیرپذیری از الگوی آرپانت و NSFNET، شبکه‌های ملی تحقیقاتی خود را بوجود آوردند. در اروپا، این شبکه‌ها (که EBONE و EuropaNET نام داشتند) از لینکهای 2-Mbps شروع کردند، و به 34-Mbps ارتقاء یافتند. در آنجا نیز زیرساخت‌های شبکه بتدریج به بخش خصوصی محول شد.

کاربردهای اینترنت

بعد از آنکه در اول ژانویه ۱۹۸۳ TCP/IP بعنوان تنها پروتکل رسمی آرپانت معرفی شد، تعداد شبکه‌ها، کامپیوترها و کاربران متصل به آن بسرعت افزایش یافت؛ و وقتی آرپانت و NSFNET به هم متصل شدند، رشد آن

حالت نمایی بخود گرفت. بسیاری از مناطق و کشورها (از جمله کانادا، اروپا و اقیانوسیه) به شبکه ملحق شدند. در اواسط دهه ۱۹۸۰ دیگر افراد به این مجموعه به عنوان شبکه‌ای از شبکه‌ها (که بعدها به اینترنت معروف شد) نگاه می‌کردند، بدون آنکه هیچگونه بخشنامه رسمی در کار باشد، یا حتی مراسم افتتاحیه‌ای (با قیچی و نوارهای رنگی، و تشویق و هورا) برگزار شده باشد.

چسبی که اینترنت را به هم متصل نگه می‌دارد، مدل TCP/IP و مجموعه پروتکل‌های آن است. پذیرش TCP/IP باعث شد تا سرویس‌های جهانی بتوانند جنبه عملی بخود بگیرند.

اما واقعاً «روی اینترنت بودن» چه معنایی دارد؟ طبق تعریف ما، ماشینی روی اینترنت است که مجموعه پروتکل‌های TCP/IP را اجرا کند، یک آدرس IP داشته باشد، و بتواند بسته‌های IP را به تمام ماشینهای دیگری که روی اینترنت هستند، بفرستد. صرف توانایی ارسال و دریافت ایمیل به معنای بودن روی اینترنت نیست، چون سرویسهای ایمیل می‌تواند به شبکه‌های خارج از اینترنت هدایت شود. با این حال، اوضاع با وضعیتی که در حال حاضر وجود دارد (میلیونها کامپیوتر شخصی می‌توانند با مودم به یک ISP وصل شده، یک آدرس IP موقتی بگیرند، و بسته‌های IP رد و بدل کنند)، کمی مغشوش و مبهم است. اما، مادامیکه این کامپیوترها به مسیریاب ISP متصل هستند، پُر بیراه نیست که آنها را روی اینترنت بدانیم.

اینترنت سستی (از ۱۹۷۰ تا اوایل دهه ۱۹۹۰) چهار کاربرد عمده داشت:

۱. ایمیل (e-mail) - نوشتن، ارسال و دریافت نامه‌های پُست الکترونیک از همان روزهای اول آرپانت جزء سرویسهای آن بود، و همچنان یکی از محبوبترین‌هاست. امروزه بسیاری از افراد روزانه دهها و صدها ایمیل دریافت می‌کنند، و به آن بعنوان درجه‌ای برای ارتباط با دنیای خارج نگاه می‌کنند - بسیار بیشتر از تلفن یا پُست معمولی.
۲. اخبار (news) - گروه خبری (newsgroup) یک محفل اختصاص یافته برای تبادل پیام در یک زمینه خاص است. امروزه هزاران گروه خبری در زمینه‌های فنی و غیرفنی (از جمله کامپیوتر، علوم، هنر و سیاست) وجود دارند. هر گروه خبری برای خود قواعد و مقرراتی دارد، که سرپیچی از آنها را برنمی‌تابد.
۳. ورود از راه دور (remote login) - هر روز هزاران نفر در سراسر دنیا برای ورود به کامپیوترهای دیگر از طریق اینترنت (البته آنهایی که حق ورود به آنها را داشته باشند)، از برنامه‌هایی مانند rlogin، telnet یا ssh استفاده می‌کنند.
۴. انتقال فایل (file transfer) - با استفاده از برنامه‌های FTP، کاربران اینترنت می‌توانند فایل‌های خود را از یک ماشین به ماشین دیگر کپی کنند. جریان انتقال دانش از این طریق بسیار گسترده و متنوع است.

تا اوایل دهه ۱۹۹۰ اینترنت جولانگاه دانشگاهیان، کارمندان دولت و محققان صنعتی بود؛ اما یک برنامه کاربردی جدید بنام WWW (World Wide Web) این وضعیت را بکلی تغییر داد، و میلیونها نفر افراد عادی نیز توانستند به کاربران حرفه‌ای اینترنت ملحق شوند. این برنامه، که توسط تام برنرز-لی (Tom Berners-Lee) از فیزیکدانان مرکز تحقیقات هسته‌ای اروپا (CERN) ابداع شد، هیچ یک از سرویسهای اینترنت را عوض نکرد، ولی کاربرد آنها را ساده‌تر کرد. به کمک این تکنولوژی جدید، و برنامه مرورگر موزایک (Mosaic browser)، که توسط مارک آندرسن در مرکز ملی کاربردهای ابرکامپیوتر (NCSA) نوشته شد، WWW ایجاد سایتیهای مشکل از صفحات مختلف (و با اطلاعاتی در قالبهای متن، تصویر، صدا، و حتی ویدئو)، با لینکهایی به صفحات دیگر را امکانپذیر کرد. با کلیک کردن روی یک لینک (link)، کاربر مستقیماً به صفحه‌ای که مشخص شده، می‌پرد. حتماً سایتیهای بسیاری متعلق به شرکتیهای بزرگ را دیده‌اید، که می‌توانید با کلیک کردن هر یک از لینکهای آن، به صفحه مربوطه (مثلاً، صفحه مربوط به محصولات شرکت، لیست قیمت‌های آن، پشتیبانی فنی، فروش و غیره) وارد شوید. در زمانی کوتاه صفحات جدید و متنوعی به WWW اضافه شد، صفحاتی مانند نقشه شهرها و کشورها،

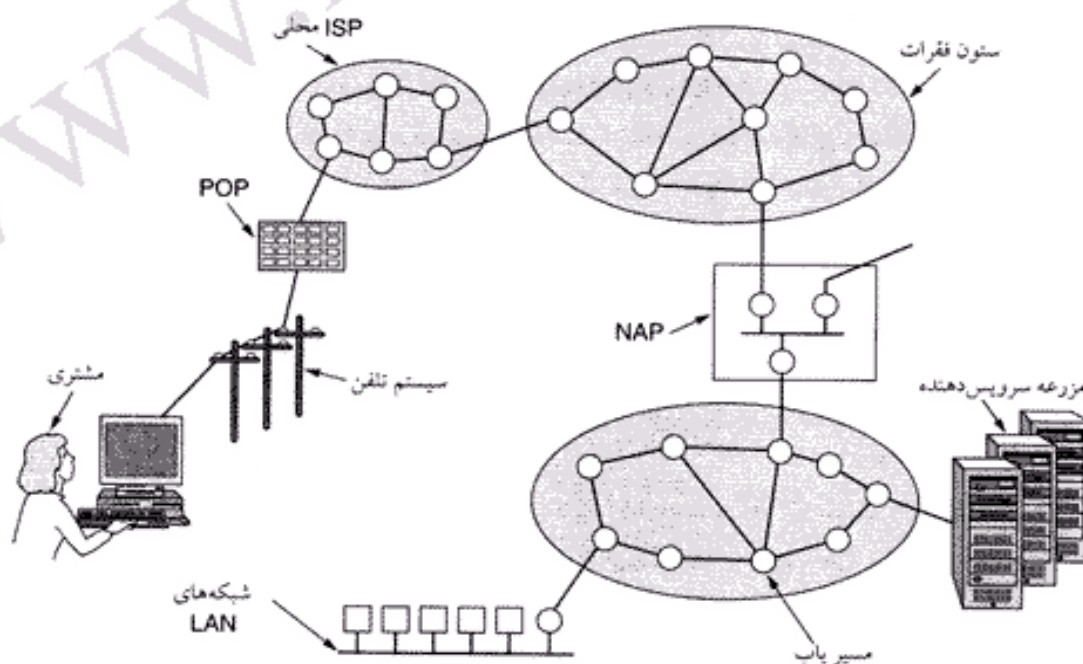
جدول قیمت سهام، کاتالوگ کارتهای کتابخانه‌ها، برنامه‌های رادیویی، و حتی متن کامل کتابهایی که از شمول قانون حق‌التألیف خارج شده‌اند (مانند کتابهای مارک تواین، چارلز دیکنز، و امثالهم). حتی بسیاری از افراد عادی نیز برای خود سایت (صفحات خانگی) ساخته‌اند.

موتور محرکه این رشد، شرکتهای ارائه‌دهنده سرویس اینترنت (ISP - Internet Service Provider) بودند. این شرکتها به افراد اجازه می‌دادند تا از خانه و با کامپیوترهای شخصی خود به اینترنت متصل شده، و از سرویسهای آن استفاده کنند. در دهه ۱۹۹۰، این شرکتها هر ساله برای دهها میلیون نفر امکان دسترسی اینترنت فراهم کردند، و چهره آنرا از محیطی دانشگاهی و نظامی به یک شبکه عمومی تغییر دادند. تعداد دقیق کاربران اینترنت در حال حاضر معلوم نیست، ولی محققاً سر به صدها میلیون نفر می‌زند، و خیلی زود از مرز یک میلیارد خواهد گذشت.

معماری اینترنت

در این قسمت سعی می‌کنیم تصویری کلی از اینترنت بدست دهیم (شکل ۱-۲۹ را ببینید). به دلیل شباهتها و تداخل وظایف زیادی که بین شرکتهای مخابرات و ISP ها وجود دارد، امروزه اوضاع بسیار در هم و مغشوش است، و بسختی می‌توان گفت کی چکاره است - به همین دلیل توضیحات ذیل ساده‌تر از آن چیز است که در واقعیت وجود دارد. اجازه دهید شکل ۱-۲۹ را جزء به جزء بررسی کنیم.

بهترین نقطه برای شروع، خانه مشتری (client) است. در اینجا فرض را بر این گذاشته‌ایم که مشتری با استفاده از یک مودم و خط تلفن به ISP متصل می‌شود. مودم (modem) وسیله‌ایست که سیگنالهای دیجیتالی کامپیوتر را به سیگنالهای آنالوگ تبدیل می‌کند، تا این سیگنالها بتوانند بدون اعوجاج روی خطوط تلفن مستقل شوند. این سیگنالها در نقطه تماس ISP (که به POP - Point Of Presence - معروف است) مجدداً تبدیل به سیگنالهای دیجیتال شده، و وارد شبکه منطقه‌ای ISP می‌شود. از این نقطه به بعد سیستم کاملاً دیجیتال است، و بر مبنای سوئیچ بسته کار می‌کند. اگر این ISP همان شرکت مخابرات باشد، POP در مرکز سوئیچینگ تلفن واقع



شکل ۱-۲۹. یک تصویر کلی از اینترنت.

خواهد بود؛ اما اگر ISP و شرکت مخابرات یکی نباشند، POP یک مرکز سوییچینگ کوچک بین راهی خواهد بود، که از آنجا به شبکه تلفن وصل می‌شود.

شبکه منطقه‌ای هر ISP از چند مسیریاب، که به شهرهای مختلف تحت پوشش آن ISP سرویس می‌دهند، تشکیل می‌شود. اگر مقصد بسته ارسال شده از مشتری یکی از کامپیوترهای واقع در همان شبکه منطقه‌ای ISP باشد، بلافاصله به آن تحویل داده می‌شود. ولی اگر چنین نباشد، بسته به اپراتور ستون فقرات ISP داده خواهد شد. اپراتور ستون فقرات (backbone operator) بالاترین نقطه این زنجیره است (شرکتهای At&T و Sprint جزء اپراتورهای عمده هستند). هر اپراتور یک شبکه بزرگ از ستونهای فقرات بین‌المللی (متشکل از هزاران مسیریاب که با فیبرهای نوری پرسرعت به هم متصلند) را اداره می‌کند. شرکتهای بزرگ و آنهایی که سرویسهای میزبانی اینترنت ارائه می‌کنند، معمولاً مستقیماً به ستون فقرات متصل هستند. اپراتورهای ستون فقرات این نوع خدمات را تشویق می‌کنند، و برای آن تسهیلات ویژه‌ای بنام کاربر کرایه‌ای فراهم می‌آورند، که باعث ارتباط نزدیک با ستون فقرات از سرعتهای بسیار بالایی برخوردارند.

اگر مقصد بسته ارسال شده یکی از ISP های متصل به ستون فقرات باشد، به نزدیکترین مسیریاب فرستاده می‌شود و از آنجا بدست وی خواهد رسید. با این حال در دنیا ستونهای فقرات متعددی (با سرعتهای مختلف) وجود دارند، و احتمال دارد که این بسته وارد یکی از این شاهراههای رقیب شود. برای اینکه بسته‌ها بتوانند براحتی بین شاهراهها حرکت کنند، تمام آنها باید به یک NAP متصل باشند - NAP چیزی نیست بیش از اتاکی پُر از مسیریاب‌های متعدد (حداقل یکی به ازای هر شاهراه)، که در یک LAN ساده به هم متصل شده‌اند. شاهراههای بزرگ، غیر از اتصال از طریق NAP، معمولاً بصورت مستقیم نیز به شاهراههای دیگر راه دارند (که به آن ارتباط دوجانبه گفته می‌شود). یکی از تناقضهای بزرگ اینترنت آن است که شرکتهایی که در انتظار عموم با هم رقابت سخت دارند، در خفا با یکدیگر ارتباطات نزدیک و تنگاتنگ برقرار می‌کنند (Metz, 2001).

این هم از مروری اجمالی بر اینترنت. البته در فصلهای آینده تک تک این اجزاء، الگوریتمها، و پروتکلها را مفصلاً بررسی خواهیم کرد. نکته‌ای که جالبست بدانید اینست که امروزه بسیاری از شرکتهای ارتباطات داخلی شبکه خود را بر اساس مدل و تکنولوژیهای اینترنت بنا می‌کنند، چیزی که به اینترانت (intranet) معروف است.

۲-۵-۱ شبکه‌های اتصال-گرا: X.25، Frame Relay و ATM

از همان اولین روزهایی که شبکه پا به عرصه وجود گذاشت، جنگ بین طرفداران زیرشبکه‌های اتصال-گرا و شبکه‌های غیرمتصل (دیتاگرام) نیز شروع شد. مهمترین برگ برنده طرفداران زیرشبکه‌های غیرمتصل همان آرپانت/اینترنت است. بیاد دارید که قصد اولیه وزارت دفاع آمریکا از بنیانگذاری آرپانت، ایجاد شبکه‌ای بود که بتواند در مقابل ضربات هسته‌ای (و منهدم شدن بخش بزرگی از خطوط و تجهیزات انتقال) دوام بیاورد (در واقع، هدف اصلی این طرح بالا بردن ضریب تحمل خرابی شبکه بود). این رهیافت منجر به طراحی شبکه‌ای شد که در آن هر بسته راه خود را مستقل از بسته‌های دیگر طی می‌کند. بدین ترتیب، اگر تعدادی از مسیریاب‌های شبکه از مدار خارج شوند، مادامیکه شبکه بتواند مسیرهای جدید خود را از نو پیکربندی کند، در ارسال بسته‌ها از مبدأ به مقصد خللی پیش نخواهد آمد.

طرفداران زیرشبکه‌های اتصال-گرا معمولاً همان شرکتهای تلفن هستند. در این سیستم، آغازکننده ارتباط قبل از آنکه بتواند ارسال اطلاعات را شروع کند، بایستی منتظر برقراری ارتباط مستقیم با طرف مقابل بماند. این ارتباط فیزیکی در تمام طول تماس برقرار می‌ماند، و تمام بسته‌های اطلاعات از همین مسیر واحد عبور خواهند کرد. اگر هر یک از تجهیزات این مسیر به هر دلیلی از کار بیفتند، تماس قطع خواهد شد - چیزی که وزارت دفاع مسلماً نمی‌پسندد.

پس علت علاقه شرکت‌های تلفن به این سیستم چیست؟ دو دلیل اصلی این علاقه عبارتند از:

۱. کیفیت سرویس
۲. حساسی مصرف‌کنندگان

در شبکه‌های اتصال-گرا هر تماس مقداری از منابع زیرشبکه (از قبیل توان پردازشی مسیریاب‌ها) را بخود اختصاص می‌دهد، و در صورتیکه این منابع به حالت اشباع برسند، تماس جدید امکانپذیر نبوده، و کاربر با بوق اشغال روبرو خواهد شد. در این روش تماس‌ها (بدلیل اختصاص منابع کافی) از کیفیت بالایی برخوردار هستند. از طرف دیگر، اگر در شبکه‌های غیرمتصل تعداد زیادی بسته به یکباره وارد یک مسیریاب شوند، ممکنست برخی از آنها (بدلیل کمبود امکانات پردازشی) در داخل روتر از بین بروند. البته فرستنده متوجه این نقص خواهد شد، و بسته‌های گمشده را از نو ارسال خواهد کرد، ولی همین موضوع باعث افت کیفیت شبکه (بویژه در مورد صدا و تصویر) می‌شود. لازم به گفتن نیست که شرکت‌های تلفن بیش از هر چیز نگران کیفیت صدا هستند، و به همین دلیل همچنان زیرشبکه‌های اتصال-گرا را ترجیح می‌دهند.

دلیل دومی که شرکت‌های تلفن سرویس‌های اتصال-گرا را بیشتر می‌پسندند، امکان صدور صورتحساب برای مشترکان است (کاری که مدت‌هاست به آن عادت کرده‌اند). هزینه تماس‌های بین شهری و خارج از کشور معمولاً بر حسب مدت مکالمه محاسبه می‌شود (علت اتخاذ این روش هم بیشتر سادگی آن بوده است). اگر تماس مستقیمی بین دو طرف مکالمه برقرار نباشد، طبیعتاً این شرکت‌ها نمی‌توانند برای مشترکان خود صورتحساب صادر کنند.

سیستم‌های محاسبه صورتحساب هزینه بسیار سنگینی به شرکت‌های تلفن تحمیل می‌کند. اگر یک شرکت تلفن بر اساس یک شارژ ثابت ماهیانه از مشترکان خود پول دریافت کند، علیرغم بالا رفتن مصرف مشترکان، می‌تواند پول زیادی صرفه‌جویی کند. اما عوامل زیادی (که بیشتر آنها هم سیاسی هستند) با اتخاذ این روش مخالفت می‌کنند. جالبست بدانید که در اغلب سیستم‌های مشابه (مانند تلویزیون‌های کابلی و برخی پارک‌های تفریحی) از روش محاسبه ثابت استفاده می‌شود. در این سیستم‌ها هم امکان پرداخت به ازای مصرف وجود دارد، ولی بدلیل هزینه بالای صدور صورتحساب معمولاً از آن اجتناب می‌شود.

به دلایل فوق، جای تعجب نیست که شرکت‌های تلفن طرفدار زیرشبکه‌های اتصال-گرا باشند؛ اما تعجب‌برانگیز این است که، اینترنت هم دارد به همین سمت پیش می‌رود - البته با این استدلال که این کار باعث بالا رفتن کیفیت سرویس‌های صدا و تصویر آن خواهد شد. اکنون اجازه دهید چند شبکه اتصال-گرا را بهتر بشناسیم.

Frame Relay و X.25

اولین شبکه اتصال-گرا که وارد سرویس عمومی شد، شبکه X.25 بود. این شبکه در اوایل دهه ۱۹۷۰، و در زمانی طراحی شد که شرکت‌های تلفن بصورت انحصاری عمل می‌کردند، و هر کشور شبکه ملی خاص خود را داشت. برای استفاده از X.25، ابتدا کامپیوتر مبدأ با ماشین مقصد تماس تلفنی برقرار می‌کرد. از آنجائیکه در آن واحد تماس‌های مختلفی می‌توانست وجود داشته باشد، به هر تماس تلفنی یک شماره داده می‌شد. بسته‌های داده بسیار ساده بودند: یک سرآیند ۳ بیتی و بدنه‌ای متشکل از ۱۲۸ بایت. سرآیند (header) تشکیل می‌شد از یک شماره تماس ۱۲ بیتی، یک شماره ترتیب بسته (packet sequence number)، یک عدد تصدیق دریافت (acknowledgement number)، و چند بیت متفرقه. شبکه‌های X.25 به مدت نزدیک به یک دهه با موفقیتی نسبی کار کردند.

در دهه ۱۹۸۰ شبکه‌های X.25 جای خود را به نوع جدیدی از شبکه‌های اتصال-گرا بنام frame relay (رله فریم) دادند. این شبکه جدید اساساً هیچ نوع کنترل خطا و کنترل جریانی نداشت، و بسته‌ها به همان نوبت

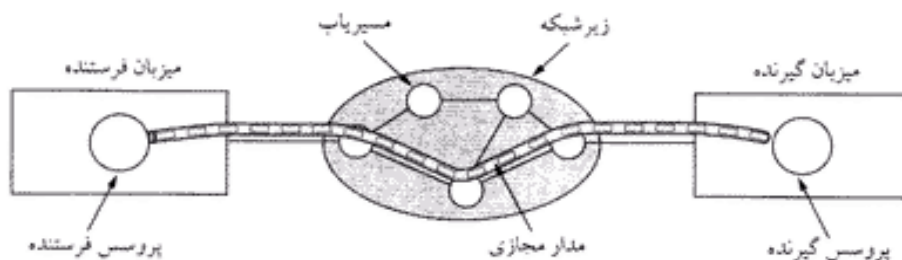
دریافت در مقصد تحویل می شدند (البته اگر به مقصد می رسیدند). این سه خصوصیت (فقدان کنترل خطا، فقدان کنترل جریان، و تحویل ترتیبی بسته ها) شبکه های frame relay را بسیار شبیه یک LAN بزرگ می کند، و در واقع بزرگترین کاربرد آن هم همین است: اتصال چند LAN دور از هم، و ایجاد یک LAN بزرگ. شبکه های frame relay هم نسبتاً موفق بودند، و هنوز در برخی جاها از آنها استفاده می شود.

حالت انتقال آسنکرون (ATM)

یکی دیگر از شبکه های اتصال-گرا (که اهمیت بسیار بیشتری نیز دارد) شبکه حالت انتقال آسنکرون (Asynchronous Transfer Mode - ATM) است. علت این نامگذاری عجیب آن است که در شبکه های تلفن اکثر تماسها بصورت سنکرون (synchronous - وابسته به پالس ساعت) هستند، در حالیکه ATM چنین نیست. شبکه ATM در اوایل دهه ۱۹۹۰ طراحی شد، و سروصدای زیادی نیز پدید آورد (Ginburg, 1996; Gorlaski, 1995; Ibe, 1997; Kime et al., 1994; Stallings, 2000). شبکه های ATM با ادعای ادغام تمام انواع شبکه و سیستمهای مخابراتی (صدا، داده، تلویزیون کابلی، تلکس، تلگراف، کیبورهای نامه بر، قوطیهای حلیمی سیمی، طیفهای افریقایی، علامتهای دودی سرخپوستان، و خلاصه هر چیزی که به نوعی اطلاعات منتقل می کند) به میدان آمدند - اتفاقی که هرگز نیفتاد. علت آن هم تا تقریباً شبیه همان بلایی بود که سر OSI آمد (زمان نامناسب، تکنولوژی بد، پیاده سازی نامناسب، و سیاستهای غلط). شرکتهای اینترنتی که منتظر وسیله ای بودند تا شرکتهای تلفن را در همان راند اول از پا در آورند، به ATM امید بستند. اما این امید دیری نپایید، و شرکتهای اینترنتی (حتی سرسخت ترین آنها) بزودی دریافتند که تارشدن به سرویسهای مطلوب راه درازی در پیش دارند. البته ATM از OSI بسیار موفقتر بود، و حتی امروز هم در شبکه های تلفن (و برای انتقال بسته های IP) مورد استفاده قرار می گیرد. از آنجائیکه این زیرشبکه فقط برای ارتباطات داخلی بکار می رود، اغلب کاربران معمولی از وجود آن اطلاعی ندارد، ولی ATM زنده و سر حال است.

مدار مجازی ATM

از آنجائیکه شبکه های ATM از نوع اتصال-گرا هستند، برای برقراری ارتباط اولیه ابتدا باید یک بسته خاص بفرستند. با عبور این بسته از زیرشبکه، تمام مسیرهای هابی که در مسیر آن قرار دارند، آنرا در جدولهای خود ثبت می کنند و منابع لازم را برای آن کنار می گذارند. به ارتباطی که بدین طریق برقرار می شود، مدار مجازی (virtual circuit) می گویند، چون بسیار شبیه مدارهای فیزیکی در شبکه های تلفن است (شکل ۱-۳۰ را ببینید). بسیاری از شبکه های ATM از مدارهای مجازی دائمی بین دو نقطه پشتیبانی می کنند (که بسیار شبیه خطوط اجاره ای در سیستم تلفن معمولی است). هر اتصال (موقت یا دائم) دارای یک شماره شناسایی است. بعد از برقراری ارتباط، دو طرف می توانند شروع به فرستادن داده کنند. ایده اصلی در ATM ارسال داده ها در بسته های کوچک و با اندازه ثابت، پنجم سلول (cell)، است. هر سلول ۵۳ بایت طول دارد، که ۵ بایت آن سرآیند، و



شکل ۱-۳۰. یک مدار مجازی.

بایت	5	48
	سرآیند	داده‌های کاربر

شکل ۱-۳۱. یک سلول ATM.

۴۸ بایت باقیمانده داده‌هاست (شکل ۱-۳۱). شماره شناسایی اتصال در سرآیند سلولها نوشته می‌شود، بطوریکه تمام مسیرهای مسیریابی می‌توانند تشخیص دهند که هر سلول متعلق به کدام اتصال است، و چگونه باید آنرا هدایت کنند. هدایت سلولها بصورت سخت‌افزاری (و با سرعت فوق‌العاده بالا) صورت می‌گیرد - در واقع، علت اصلی اندازه ثابت سلولها در شبکه‌های ATM اینست که ساخت مسیرهای سخت‌افزاری برای آن بسیار ساده است (هدایت بسته‌های IP با اندازه متغیر به مسیرهای نرم‌افزاری نیاز دارد، که بسیار کندتر هستند).

مزیت دیگر ATM توانایی آن در ارسال همزمان یک سلول به مسیرهای مختلف است - که این ویژگی در سیستمهای پخش تلویزیونی بسیار مفید است. از طرف دیگر، کوچک بودن سلولها باعث می‌شود تا هیچ خطی برای مدت طولانی اشغال نشود، و کیفیت سرویس افزایش یابد.

در ATM تمام سلولها از یک مسیر به مقصد هدایت می‌شوند. البته تضمینی برای رسیدن یک سلول به مقصد وجود ندارد، ولی ترتیب آنها حتماً رعایت می‌شود. اگر سلول ۲ بعد از سلول ۱ فرستاده شده باشد، به همان ترتیب به مقصد می‌رسند، و هرگز سلول ۲ پیش از سلول ۱ به مقصد نخواهد رسید. اگر هر یک از این سلولها (یا هر دوی آنها) در بین راه از بین بروند، این بر عهده پروتکل‌های لایه‌های بالاتر است که آنها را بازیابی کنند. از این نظر ATM حداقل یک پله بالاتر از اینترنت می‌ایستد (که نه ترتیب بسته‌ها ضمانت می‌شود، نه حتی رسیدن صحیح و سالم آنها).

سازماندهی شبکه‌های ATM شبیه WAN های قدیمی (متشکل از خطوط تلفنی و سونیچها) است. متداولترین سرعتها در شبکه‌های ATM عبارتند از 155-Mbps و 622-Mbps (البته ATM از سرعتهای بالاتر هم پشتیبانی می‌کند). علت انتخاب سرعت 155-Mbps آنست که تلویزیونهای با وضوح بالا (HDTV) به چنین سرعتی نیاز دارند - مقدار دقیق این سرعت 155.52-Mbps است، که دقیقاً معادل سرعت سیستم AT&T SONET می‌باشد. علت انتخاب سرعت 622-Mbps نیز اینست که از ترکیب چهار کانال 155.52-Mbps یک کانال 622-Mbps بوجود می‌آید.

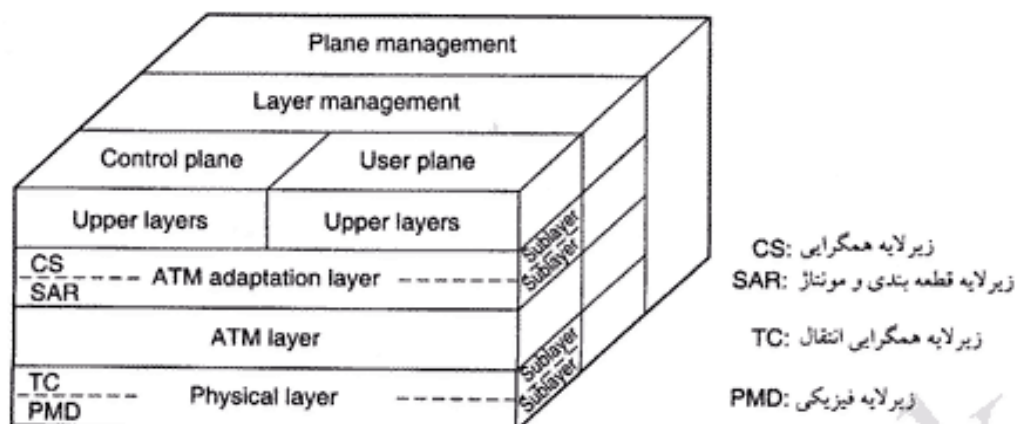
مدل مرجع ATM

شبکه ATM برای خود یک مدل مرجع مستقل دارد، که با مدل‌های OSI و TCP/IP فرق دارد - این مدل را در شکل ۱-۳۲ ملاحظه می‌کنید. این مدل سه لایه دارد: لایه فیزیکی، لایه ATM، لایه انطباق ATM (و هر چند لایه که کاربر مایل باشد روی این لایه‌ها سوار کند).

لایه فیزیکی با مشخصات فیزیکی سیستم (ولتاژها، زمانبندی بیت‌ها و غیره) سروکار دارد. مدل ATM هیچ پیشیازی در مورد این مشخصات ندارد، و می‌گوید که ارسال سلولها می‌تواند بصورت مستقیم یا از طریق سیستمهای انتقال دیگر انجام شود. بعبارت دیگر، ATM مستقل از سیستم انتقال است.

لایه ATM با خود سلولها و انتقال آنها سروکار دارد. ایجاد و رها کردن مدار مجازی، تعریف فیلدهای سرآیند سلول، و کنترل ازدحام (congestion control) از وظایف این لایه است.

از آنجائیکه اکثر برنامه‌های کاربردی تمایلی به کار کردن با بسته‌هایی به کوچکی سلولهای ATM ندارند، یک لایه دیگر بالای لایه ATM تعبیه شده تا این قبیل برنامه‌ها بتوانند بسته‌های بزرگتری به ATM بفرستند. این لایه



شکل ۱-۳۲. مدل مرجع ATM.

(در سمت فرستنده) بسته های داده را به سلولهای ۵۳ بایتی می شکند، و در طرف گیرنده آنها را دوباره سر هم می کند. نام این لایه، لایه انطباق ATM (ATM Adaptation Layer - AAL) است. بر خلاف مدل های قبلی که دو بُعدی بودند، مدل ATM یک مدل مرجع سه بُعدی است (شکل ۱-۳۲ را ببینید). صفحه کاربر (user plane) با انتقال داده، کنترل جریان، تصحیح خطا، و دیگر عملکردهای کاربر سروکار دارد. از طرف دیگر، صفحه کنترل (control plane) مدیریت اتصال را بر عهده دارد. وظیفه صفحه های مدیریت لایه (layer management) و مدیریت صفحه (plane management) مدیریت منابع سیستم و هماهنگ کردن لایه های بینایی است. لایه های فیزیکی و AAL هر یک به دو زیر لایه تقسیم شده اند، یکی در پائین برای انجام عملکردهای محوله، و دیگری در بالا برای ارتباط با لایه بالاتر (که زیر لایه همگرایی - convergence sublayer - خوانده می شود). وظیفه هر یک از این لایه ها و زیر لایه ها را در شکل ۱-۳۳ ملاحظه می کنید.

OSI لایه	ATM لایه	ATM زیر لایه	کارکرد
3/4	AAL	CS	واسط استاندارد
		SAR	قطعه بندی و مونتاژ
2/3	ATM		کنترل جریان تولید سرآیند سلول مدار مجازی - مدیریت مسیر مانتی پلکس/دو مانتی پلکس سلول.
2		TC	ایزوله کردن سرعت سلول تولید مجموع تطبیقی تولید سلول بسته بندی و باز کردن بسته ها تولید فریم
1	Physical		
		PMD	زمان بندی بیت دسترسی فیزیکی

شکل ۱-۳۳. وظایف لایه ها و زیر لایه های ATM.

زیرلایه PMD (Physical Medium Dependent) مستقیماً به کابل شبکه وصل می‌شود، و کار آن ارسال و دریافت بیت‌ها و ایجاد همزمانی بین آنهاست. هر نوع کابل و سیستم انتقال زیرلایه PMD خاص خود را دارد. زیرلایه دیگر لایه فیزیکی، TC (Transmission Convergence) نام دارد. وقتی یک سلول ارسال می‌شود، لایه TC آنرا بصورت جریانی از بیت‌ها به لایه PMD می‌فرستد - که این کاری ساده است. در طرف گیرنده، لایه TC باید جریان بیت‌هایی را که از لایه PMD دریافت می‌کند، دوباره بصورت سلول در آورد - بعبارت دیگر باید بتواند ابتدا و انتهای هر سلول را بدرستی تشخیص دهد. در مدل ATM این کار در لایه فیزیکی انجام می‌شود، وظیفه‌ی که در مدل OSI (و تقریباً تمام مدل‌های دیگر) بر عهده لایه پیوند داده است.

مانطور که قبلاً هم گفتیم، لایه ATM مدیریت ایجاد و انتقال سلولها را بر عهده دارد. مهمترین بخش از وظایف ATM نیز در همین لایه صورت می‌گیرد. این لایه تلفیقی است از لایه‌های لینک داده و شبکه در مدل OSI - که در ضمن هیچ زیرلایه‌ای هم ندارد.

لایه AAL به دو زیرلایه SAR (Segmentation And Reassembly) و CS (Convergence Sublayer) تقسیم شده است. لایه پائینی (SAR) در طرف فرستنده بسته‌های داده را به سلول می‌شکند، و در طرف گیرنده دوباره آنها را به هم می‌چسباند. لایه بالایی (CS) وظیفه ارائه سرویسهای مختلف به برنامه‌های کاربردی را بر عهده دارد.

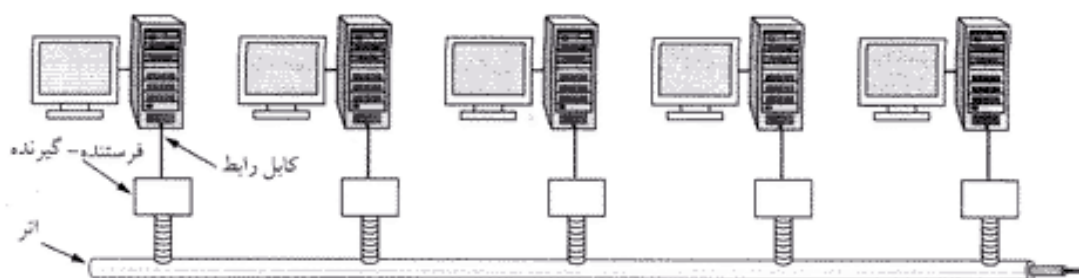
از آنجائیکه ATM در سراسری زوال قرار دارد، در این کتاب بیش از این درباره آن صحبت نخواهیم کرد. با این حال، بدلیل نصب در مقیاس وسیع، به احتمال زیاد تا چند سال دیگر نیز دوام خواهد آورد. برای کسب اطلاعات بیشتر درباره ATM به (Dobrowski and Grise, 2001; Gadecki and Heckart, 1997) مراجعه کنید.

۳-۵-۱ اینترنت

اینترنت و ATM هر دو شبکه‌های گسترده هستند، ولی در هر شرکت، سازمان و دانشگاه تعداد زیادی کامپیوتر وجود دارد که باید به هم متصل شوند. از همین جاست که نیاز به شبکه‌های محلی شکل می‌گیرد. در این قسمت می‌خواهیم کمی درباره متداولترین شبکه محلی، یعنی اینترنت (Ethernet)، صحبت کنیم.

داستان ما از ایالت بکر و دست نخورده هاوایی، در اوایل دهه ۱۹۷۰، شروع می‌شود - در اینجا منظور از بکر و دست نخورده فقدان شبکه تلفن است. با اینکه نبود شبکه تلفن برای کسانی که هاوایی را برای استراحت انتخاب می‌کردند، یک مزیت بود، ولی برای محقق بنام نورمن آبرامسون و همکارانش در دانشگاه هاوایی که می‌خواستند کاربران جزایر دورافتاده هاوایی را به کامپیوتر مرکزی در هونولولو (مرکز ایالت هاوایی) متصل کنند، چندان خوشایند نبود. کشیدن کابل از وسط اقیانوس آرام مسلماً نمی‌توانست راه حل مشکل آنها باشد، پس باید فکر دیگری می‌کردند.

یکی از راه‌حل‌هایی که آنها پیدا کردند، استفاده از امواج رادیویی با بُرد کوتاه بود. هر ترمینال به یک رادیوی کوچک وصل می‌شد، که دو فرکانس داشت: فرکانس ارسال (به کامپیوتر مرکزی)، فرکانس دریافت (از کامپیوتر مرکزی). وقتی کاربر می‌خواست به کامپیوتر مرکزی وصل شود، روی فرکانس ارسال یک بسته می‌فرستاد. اگر کس دیگری در همان لحظه در حال ارسال نبود، بسته مزبور به کامپیوتر مرکزی می‌رسید، و کاربر می‌توانست بسته تصدیق دریافت (acknowledgement) آنرا روی فرکانس دریافت بگیرد. اما اگر کانال اشغال بود، ترمینال بسته تصدیق دریافت را نمی‌گرفت، و متوجه می‌شد که باید دوباره سعی کند. از آنجائیکه روی کانال دریافت فقط یک کامپیوتر (کامپیوتر مرکزی) مجاز به ارسال اطلاعات بود، هرگز در آن انسداد پیش نمی‌آمد. این سیستم، که آلوهانت (ALOHA NET) نام گرفت، در شرایط ترافیک پائین خوب کار می‌کرد، ولی در ترافیک بالا بشدت ناکارآمد می‌شد.



شکل ۱-۳۴. معماری اترنت اولیه.

در همان زمان، دانشجویی بنام باب متکالف که تازه از M.I.T فارغ التحصیل شده بود، به قصد ادامه تحصیل در مقطع دکترا وارد دانشگاه هاروارد شد. باب در حین مطالعات خود با کارهای آبرامسون آشنا، و بشدت به آن علاقمند شد. این علاقه تا آن حد بود که باب تصمیم گرفت قبل از شروع به کار در مرکز تحقیقات زیراکس در پالو آلتو (Xerox PARC)، تابستان را در هاوایی بگذراند و با آبرامسون کار کند. وقتی باب متکالف به Xerox PARC برگشت، متوجه شد که محققان آنجا روی پروژه های کار می کنند که بعدها به کامپیوتر شخصی (Personal Computer - PC) معروف شد. ولی این یک ماشین ایزوله و جدا از همه جا بود. باب، با استفاده از تجارب آبرامسون، و به کمک یکی از همکارانش بنام دیوید باگزر، اولین شبکه محلی را طراحی و پیاده سازی کرد (Metcalfe and Boggs, 1976).

آنها این سیستم را (بیاد ماده ای خیالی بنام اتر - ether - که تا مدتها تصور می شد محیط انتشار امواج الکترومغناطیس است) اترنت نامیدند. (بعد از کشف معادلات انتشار امواج الکترومغناطیس توسط فیزیکدان بریتانیایی، جیمز کلارک ماکسول، دانشمندان فرض را بر این گذاشتند که این امواج در محیطی بنام اتر منتشر می شوند. فقط بعد از آزمایشات معروف مایکلسون-مورلی بود که فیزیکدانان دریافتند امواج الکترومغناطیس می توانند در خلاء منتشر شوند.)

رسانه انتشار در این سیستم خلاء نبود، بلکه از یک رشته کابل هم محور (coaxial) ضخیم بطول حداکثر ۲/۵ کیلومتر (با یک تکرارکننده در هر ۵۰۰ متر) استفاده می شد (این کابل در واقع همان اتر محسوب می شد) - شکل ۱-۳۴ را ببینید. حداکثر تا ۲۵۶ کامپیوتر را می شد به این کابل متصل کرد. به چنین کابلی، که چندین ماشین بصورت موازی به آن متصل شده اند، کابل چنداتصال (multidrop cable) گفته می شود. اترنت با سرعت 2.94 Mbps کار می کرد. اترنت یکی از اشکالات عمده آلو هانت را نیز برطرف کرده بود: هر کامپیوتر قبل از ارسال بسته خود ابتدا به کابل گوش می کرد، تا مطمئن شود کس دیگری در همان لحظه در حال استفاده از آن نیست. اگر چنین بود، کامپیوتر تا خالی شدن خط کار خود را عقب می انداخت. بدین ترتیب، اترنت توانست با اجتناب از تداخل های بی مورد به کارایی بالاتری دست یابد. البته آلو هانت امکان انجام چنین کاری را نداشت، چون فرستنده یک جزیره نمی توانست امواج فرستنده های جزایر دیگر را بشنود.

علیرغم اینکه کامپیوترهای اترنت تا وقتی خط خالی نباشد، اقدام به ارسال نمی کنند، اما مسئله دیگری ممکنست بروز کند: اگر چند کامپیوتر همزمان منتظر خالی شدن خط باشند، و به محض اینکه خط خالی شد، در یک لحظه شروع به ارسال اطلاعات خود کنند، چه خواهد شد؟ راه حل این مشکل آن است که هر کامپیوتر در تمام لحظات ارسال اطلاعات خود به خط گوش کند، و اگر متوجه تداخل امواج شد، ابتدا به دیگران اخطار می فرستد، و سپس برای مدتی کوتاه (که مقدار آنرا بطور تصادفی تعیین می کند) کنار می کشد، و بعد از این مدت دوباره سعی

می‌کند؛ اگر در مرتبه بعد باز هم تداخل پیش آمد، مدت انتظار را دو برابر می‌کند، تا بالاخره یکی از آنها فرصت ارسال امواج را پیدا کند.

ایترنت زیراکس چنان موفق بود که در سال ۱۹۷۸ شرکتهای DEC، ایتل و زیراکس استاندارد بنام DIX برای اینترنت 10-Mbps وضع کردند. استاندارد DIX در سال ۱۹۸۳ با دو تغییر جزئی به استاندارد IEEE 802.3 تبدیل شد.

متأسفانه زیراکس همواره یکی از شرکتهایی بوده است که اختراعات بسیار مهمی (مانند PC، اینترنت و ماوس) ز آنجا منشأ گرفته، ولی نتوانسته اقدامی برای تجاری کردن آنها بعمل آورد. وقتی زیراکس علاقه چندانی به اینترنت نشان نداد (و فقط در استاندارد کردن آن همکاری کرد)، باب متکالف شرکی بنام 3Com تأسیس کرد تا بتواند برای PC ها کارت شبکه اینترنت تولید کند.

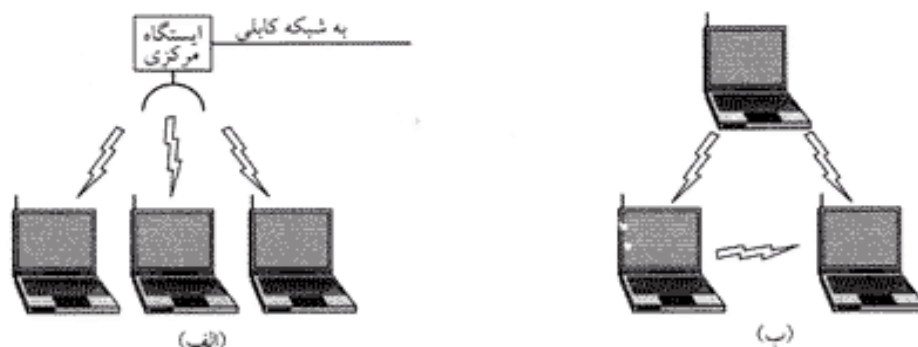
اینترنت به رشد و توسعه خود ادامه داد (رشدی که همچنان ادامه دارد)، و امروزه به سرعتی 100-Mbps و 1000-Mbps دست یافته است - سرعتی که انتظار می‌رود باز هم افزایش یابد. کابل کشی، سوئیچینگ و سایر جنبه‌های اینترنت نیز بهبود یافته، و ویژگیهای جدیدی به آن اضافه شده است. در فصل ۴ مفصلاً درباره اینترنت صحبت خواهیم کرد.

البته همین جا لازم به توضیح است که اینترنت (IEEE 802.3) تنها استاندارد LAN نیست. استانداردهای خط توکین (IEEE 802.4 - Token Bus) و حلقه توکین (IEEE 802.5 - Token Ring) نیز از جمله استانداردهای معروف LAN هستند. وجود سه استاندارد کمابیش ناسازگار برای شبکه‌های محلی بیشتر از اینکه مسئله‌ای فنی باشد، موضوعی سیاسی است. در همان زمان که اینترنت در حال استاندارد شدن بود، جنرال موتورز نیز در کار ایجاد شبکه‌ای بود که از همان توپولوژی اینترنت (یعنی کابل خطی) استفاده می‌کرد، ولی نوبت ارسال هر کامپیوتر با استفاده از بسته خاصی بنام توکین، که بین کامپیوترها دست به دست می‌گشت، تعیین می‌شد. هر کامپیوتر فقط زمانی می‌توانست اقدام به ارسال اطلاعات کند که توکین را در اختیار داشته باشد، و بدین ترتیب مشکل تداخل حل می‌شد. جنرال موتورز اعلام کرد که این تکنیک برای خط تولید کارخانجات اتومبیل سازی ضرورت مطلق دارد، و حاضر نبود حتی یک میلیمتر از موضع خود عقب‌نشینی کند. با وجود چنین ادعایی، 802.4 اکنون از صفحه روزگار محو شده است.

غول صنعت کامپیوتر، IBM هم سوگلی خود را داشت: حلقه توکین. تنها تفاوت این سیستم با شبکه جنرال موتورز آن بود که در اینجا کابل شبکه یک مسیر بسته (حلقه) را تشکیل می‌داد. بر خلاف 802.4، 802.5 هنوز در برخی از سایت‌های IBM مورد استفاده است (ولی خارج از IBM هیچکس از آن استفاده نمی‌کند). تحقیقاتی در زمینه نسل جدید و پرسرعت حلقه توکین با سرعت گیگابیت (802.5v) در جریان است، ولی بنظر نمی‌رسد بتواند با اینترنت رقابت کند. خلاصه اینکه، در جنگی که بین اینترنت، خط توکین و حلقه توکین در گرفت، اینترنت پیروز شد، چون اولین و در ضمن از رقبایش بهتر بود.

۵-۱۱ شبکه‌های محلی بیسیم: 802.11

تقریباً همزمان با به بازار آمدن کامپیوترهای کتابی، بسیاری افراد این رؤیا را در سر می‌پروراندند که بتوانند به محض ورود به جایی که دسترسی اینترنت وجود دارد، بلافاصله و بنحوی جادویی کامپیوترشان به اینترنت متصل شود - و همیشه وقتی رؤیایی وجود دارد، افرادی هم هستند که به فکر محقق کردن آن بیفتند. عملی‌ترین رهیافتی که برای به فعل در آوردن این ایده وجود داشت، مجهز کردن کامپیوترها به فرستنده-گیرنده‌های رادو، وی برد کوتاه بود - از همین جا بود که شرکتهای متعددی سرعت شبکه‌های محلی بیسیم را وارد بازار کردند.



شکل ۱-۳۵. شبکه بیسیم (الف) با ایستگاه مرکزی، و (ب) بدون ایستگاه مرکزی.

مشکل اصلی این بود که هیچکدام از این شبکه های بیسیم با هم سازگار نبودند، و کامپیوترهایی که به بیسیم های مختلف مجهز بودند، نمی توانستند با هم ارتباط برقرار کنند. بالاخره همه به این نتیجه رسیدند که استاندارد کردن شبکه های بیسیم می تواند ایده خوبی باشد، و بدنبال آن کمیته استاندارد IEEE مأمور تدوین این استاندارد شد - استاندارد 802.11 نام گرفت، و در میان عموم به WiFi معروف است. این یکی از استانداردهای مهم صنعت کامپیوتر است، و شایسته توجه کافی.

استاندارد پیشنهاد شده باید در دو حالت کار می کرد:

۱. در شرایط وجود یک ایستگاه مرکزی
۲. در شرایط فقدان ایستگاه مرکزی

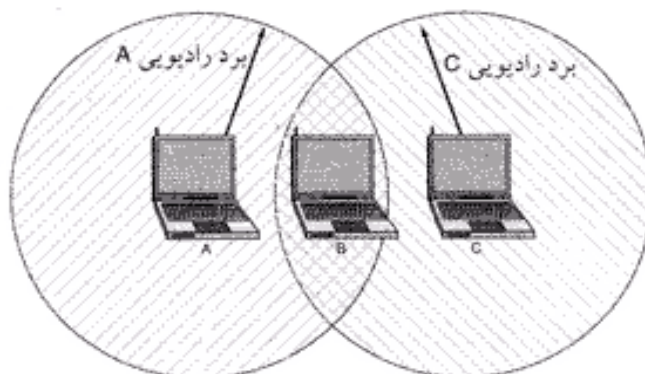
در حالت اول، تمام پیامها باید از طریق ایستگاه مرکزی، که در استاندارد 802.11 به آن نقطه دسترسی (access point) گفته می شود، مبادله شوند. اما در حالت دوم، کامپیوترها مستقیماً با یکدیگر ارتباط برقرار می کنند. این دو حالت را در شکل ۱-۳۵ مشاهده می کنید.

شروع می شود، و اعداد 1 تا 10 قبلاً استفاده شده بود، نام این استاندارد 802.11 شد. ولی بقیه کار به همین سادگی نبود.

برخی از مهمترین چالشهایی که کمیته تدوین استاندارد با آنها روبرو بود، عبارت بودند از: انتخاب یک باند فرکانسی مناسب (که ترجیحاً بین المللی نیز باشد)؛ محدود بودن بُرد سیگنالهای رادیویی؛ تأمین ایمنی مناسب؛ عمر محدود باتری در کامپیوترهای کتابی؛ مسائل بهداشتی (هنوز این مسئله بدرستی روشن نشده که آیا امواج رادیویی سرطانزا هستند یا خیر)؛ پیامدهای متحرک بودن کامپیوترها؛ و بالاخره، ایجاد سیستمی که از نظر پهنای باند ارزش اقتصادی داشته باشد.

با توجه به غالب بودن اینترنت در زمان تدوین این استاندارد، کمیته تصمیم گرفت که 802.11 باید در لایه های بالاتر از لایه پیوند داده با اینترنت سازگار باشد. بویژه، ارسال بسته های IP در شبکه های بیسیم بایستی دقیقاً به همان روش اینترنت باشد.

با این وجود، لایه های فیزیکی و لینک داده در این دو سیستم تفاوت های اساسی با هم دارند. اول اینکه، در اینترنت هر کامپیوتر قبل از شروع به ارسال اطلاعات به اتر (کابل شبکه) گوش می کند، و فقط در صورت خالی بودن آن شروع به ارسال می کند. در شبکه های بیسیم کار به همین سادگی نیست. برای درک علت آن، به شکل ۱-۳۶ نگاه کنید. فرض کنید کامپیوتر A در حال ارسال اطلاعات به کامپیوتر B است، ولی بُرد امواج آن به کامپیوتر C نمی رسد. اگر در این لحظه C بخواهد چیزی به B بفرستد، باید به اتر (در اینجا، فضا) گوش کند، ولی آیا عدم



شکل ۱-۳۶. گاهی بُرد امواج رادیویی برای پوشش دادن به تمام شبکه کافی نیست.

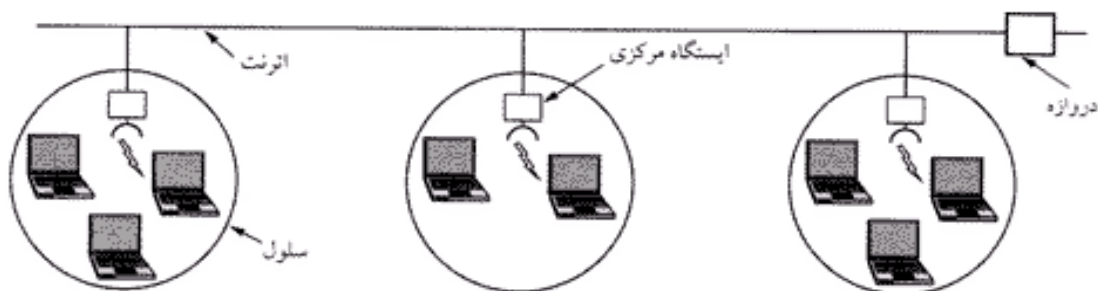
دریافت امواج به معنای آن است که می تواند با اطمینان شروع به ارسال کند؛ مسلماً خیر. استاندارد 802.11 باید این مسئله را حل می کرد.

مسئله دومی که باید حل می شد این بود که اجسام سخت امواج رادیویی را منعکس می کنند، و یک موج می تواند چندین بار (و از مسیرهای مختلف) به گیرنده برسد. این تداخل امواج باعث بروز حالتی می شود که به آن محوشدگی چندمسیره (multipath fading) می گویند.

مسئله سوم این است که بسیاری از نرم افزارهای موجود اساساً با چنین وضعیتی (متحرک بودن کامپیوتر) آشنا نیستند. برای مثال، بسیاری از برنامه ها دارای لیستی از چاپگرها هستند که می توانند روی آنها چاپ کنند. وقتی یکی از این برنامه ها (که روی یک کامپیوتر کتابی نصب شده) وارد محیط جدیدی می شود، نمی تواند تشخیص دهد که فهرست قبلی چاپگرها دیگر در این وضعیت اعتبار ندارد.

مسئله چهارم این است که وقتی یک کامپیوتر از بُرد یک ایستگاه مرکزی خارج و وارد محدوده ایستگاه دیگری می شود، باید مکانیزمی برای این جابجایی وجود داشته باشد. با اینکه تلفنهای همراه دارای چنین مکانیزمی هستند، اما این اتفاق در اینترنت نمی افتد، و باید بطریقی حل شود. یکی از راه حل های این مسئله، متصل کردن ایستگاههای مرکزی به یکدیگر از طریق کابل است (شکل ۱-۳۷ را ببینید). از دید دنیای خارج، این شبکه کاملاً شبیه یک شبکه اینترنت واحد است. نقطه اتصال سیستم 802.11 با دنیای خارج را درگاه (portal) می گویند.

بعد از مدتی کار سخت، کمیته موفق به تدوین استانداردی شد که این مسائل (و بسیاری مسائل دیگر) در آن حل شده بود. این شبکه بیسیم با سرعتهای 1-Mbps و 2-Mbps کار می کرد. تقریباً بلافاصله، همه لب به شکایت گشودند که این سرعتها بسیار کم است، و کار بر روی استانداردهای سریعتر آغاز شد. در همین زمان کمیته استاندارد به دو قسمت تقسیم شد، که در سال ۱۹۹۹ هر کدام استاندارد جداگانه ای وضع کردند. استاندارد



شکل ۱-۳۷. یک شبکه 802.11 چندسلولی.

802.11a از باند فرکانسی وسیعتری نسبت به 802.11 استفاده می کند، و سرعت آن به 54-Mbps می رسد. استاندارد 802.11b در همان باند فرکانسی 802.11 کار می کند، ولی با استفاده از مدولاسیون متفاوت به سرعت 11-Mbps دست می یابد. همانطور که می بینید، هر دوی این استانداردها از 802.11 (و حتی از اینترنت اولیه) سریعترند، ولی هنوز بدرستی معلوم نیست که کدامیک از آنها برنده نهایی این مسابقه اند. برای اینکه اوضاع از این هم پیچیده تر شود، کمیته 802.11 استاندارد جدیدی بنام 802.11g تدوین کرده، که در باند فرکانسی 802.11b، ولی با مدولاسیون 802.11a، کار می کند. در فصل ۴ مفصلاً درباره 802.11 صحبت خواهیم کرد.

شکی نیست که 802.11 انقلاب جدیدی در دنیای کامپیوتر و اینترنت بها کرده است. فرودگاهها، ایستگاههای قطار، بنادر، هتلها، فروشگاهها، دانشگاهها، مراکز آموزشی (و حتی کافه های کوچک) بسرعت در حال نصب شبکه های 802.11 هستند. در واقع، 802.11 همان چیزی را به اینترنت داده است، که کامپیوترهای کتابی به دنیای کامپیوتر دادند: تحرک پذیری.

۶-۱ استانداردهای شبکه

تعداد زیادی سازنده و تأمین کننده قطعات و تجهیزات شبکه وجود دارد، که فکر می کنند می دانند چگونه باید کار خود را انجام دهند. اما بدون یک عامل هماهنگ کننده، این وضعیت می تواند به یک آشوب واقعی بینجامد، و در نهایت هیچ کاری هم انجام نشود. تنها راه برای خلاصی از چنین وضعیتی، توافق بر سر استانداردهای شبکه است. استانداردها نه تنها اجازه می دهد تا تجهیزات مختلف بتوانند با هم کار کنند، بلکه فروش محصولات منطبق با استاندارد را نیز افزایش می دهند - و فروش بیشتر یعنی تولید انبوه، کاهش هزینه ها، طراحی بهتر، کاهش قیمت ها، و افزایش مجدد درخواست (و مگر اقتصاد چیزی غیر از این است). در این قسمت نگاهی به استانداردهای بین المللی (که از اهمیت زیادی برخوردارند، ولی کمتر شناخته شده اند) خواهیم داشت.

استانداردها بر دو نوعند: استانداردهای بالفعل (de facto)، و استانداردهای قانونی (de jure). استانداردهای بالفعل آنهایی هستند که بدون هیچ طرح رسمی بوجود آمده و پذیرفته شده اند. کامپیوترهای سازگار با IBM PC (که به اختصار PC خوانده می شوند) از جمله استانداردهای بالفعل هستند، چون شرکتهای بسیاری تصمیم گرفتند تا کپی های دقیق و کاملی از این نوع کامپیوتر بسازند. یکی دیگر از استانداردهای بالفعل، سیستم عامل یونیکس (UNIX) است، که در دانشکده های کامپیوتر بعنوان سیستم عامل استاندارد پذیرفته شده است. از طرف دیگر، استانداردهای قانونی آنهایی هستند که توسط مراجع مسئول بین المللی پذیرفته شده اند. مراجع بین المللی استاندارد به دو دسته تقسیم می شوند: آنهایی که طبق معاهدات بین المللی تأسیس شده اند، و آنهایی که بصورت داوطلبانه شکل گرفته اند. در زمینه استانداردهای شبکه های کامپیوتری، سازمانهایی از هر دو دسته وجود دارند، که در زیر آنها را معرفی خواهیم کرد.

۶-۱-۱ مراجع مسئول استانداردهای مخابرات

وضعیت قانونی شرکتهای تلفن از کشوری به کشور دیگر بطرز چشمگیری متفاوت است. در یک سو ایالات متحده آمریکا قرار دارد، که در آن متجاوز از ۱۵۰۰ شرکت خصوصی در این زمینه مشغول به کار هستند. شرکت AT&T، قبل از آنکه در سال ۱۹۸۴ طبق قانون ضد تراست به دو بخش تقسیم شود، بزرگترین شرکت خصوصی دنیا بود، و مخابرات این کشور را تحت سلطه کامل خود داشت. این شرکت سرویس تلفن ۸۰ درصد مناطق آمریکا را تأمین می کرد، در حالیکه بقیه شرکتهای تنها ۲۰ درصد بازار (آن هم اغلب در مناطق روستایی) را در اختیار داشتند. بعد از تقسیم AT&T، این شرکت اجازه یافت تا فقط در مخابرات راه دور فعالیت کند (آن هم در رقابت با

سایر شرکتها). هفت شرکت منطقه‌ای بل (Bell)، که از AT&T منشعب و با شرکتهای دیگر متحد شده بودند، نیز خدمات تلفن شهری و تلفن همراه را برعهده گرفتند. بدلیل ادغامها و انشعابات متعدد، صنعت مخابرات در ایالات متحده اکنون از وضعیت ثابتی برخوردار است.

در ایالات متحده، به شرکتهایی که سرویسهای مخابراتی در اختیار عموم قرار می‌دهند، کاربرد عمومی گفته می‌شود. نوع سرویسها و تعرفه خدمات این شرکتها توسط کمیسیون فدرال مخابرات (برای تماسهای بین‌ایالتی و بین‌المللی) و کمیسیونهای ایالتی (برای تماسهای داخل ایالتی) تعیین می‌شود.

در سوی دیگر، کشورهایی قرار دارند که در آنها تمامی سرویسهای مخابراتی (از جمله، پست، تلگراف، تلفن، و حتی رادیو و تلویزیون) تحت انحصار مطلق دولت قرار دارد. اکثر کشورهای دنیا هم از این دسته‌اند. در برخی از این کشورها، مخابرات را یک شرکت ملی اداره می‌کند، و در برخی دیگر دولت مستقیماً (از طریق وزارتخانه‌ای بنام پ.ت.ت: پست-تلگراف-تلفن) کارها را در دست دارد. اما، در کل دنیا حرکت یکپارچه‌ای به سمت آزادسازی، رقابت و حذف انحصار دولت آغاز شده است. امروزه اکثر کشورهای اروپایی مخابرات خود را به بخش خصوصی سپرده‌اند، و در دیگر کشورها این فرآیند (هر چند کند و بطئی) ادامه دارد.

با این همه شرکت و سرویسهای مخابراتی، وجود نوعی استاندارد بین‌المللی برای تضمین ارتباط بین افراد (و کامپیوترها) از کشوری به کشور دیگر ضروری بنظر می‌رسد. البته این نیاز از مدتها قبل آشکار شده بود: در سال ۱۸۶۵، نمایندگانی از کشورهای مختلف اروپایی آنچه را که امروز ITU (اتحادیه بین‌المللی مخابرات - International Telecommunication Union) نامیده می‌شود، بنا نهادند. وظیفه این اتحادیه استاندارد کردن مخابرات بین‌المللی (که در آن زمان فقط شامل تلگراف می‌شد) بود. حتی در آن زمان نیز مشخص بود که اگر کلیه کشورها در مورد استفاده از یک کُد یکسان (که همان کُد مورس بود) به توافق نرسند، مشکلات عدیده‌ای بروز خواهد کرد. با ورود تلفن به صحنه مخابرات بین‌المللی، ITU وظیفه استاندارد کردن آنرا نیز بر عهده گرفت. در سال ۱۹۴۷، ITU بصورت یکی از سازمانهای تابعه سازمان ملل متحد درآمد.

اتحادیه بین‌المللی مخابرات (ITU) سه بخش عمده دارد:

۱. بخش مخابرات رادیویی (ITU-R)
۲. بخش تدوین استانداردهای مخابراتی (ITU-T)
۳. بخش توسعه (ITU-D)

وظیفه ITU-R تخصیص فرکانسهای رادیویی به متقاضیان در سراسر دنیاست. بخش ITU-T (که بیشتر مد نظر ماست) وظیفه تدوین استاندارد برای سیستمهای مخابراتی (تلفن و داده) را بر عهده دارد. از سال ۱۹۵۳ تا ۱۹۹۳، ITU-T با نام CCITT (که از نام فرانسوی آن - Comité Consultatif International Télégraphique et Téléphonique - گرفته شده بود) شناخته می‌شد. در اول مارس ۱۹۹۳، CCITT برای کاهش بوروکراسی ساختار اداری‌اش را تغییر داد، و نام خود را نیز ITU-T گذاشت. ITU-T و CCITT هر دو توصیه‌هایی را در زمینه مخابرات تلفنی و داده منتشر می‌کردند. امروزه نیز افراد بسیاری به توصیه‌های CCITT مراجعه می‌کنند، اگر چه از سال ۱۹۹۳ این توصیه‌ها نام ITU-T را بر خود دارند.

اعضای ITU-T به چهار دسته تقسیم می‌شوند:

۱. دولتها
۲. اعضای بخش
۳. اعضای وابسته
۴. نمایندگیها

تقریباً تمام کشورهای عضو سازمان ملل متحد (یعنی حدود ۲۰۰ کشور) در ITU-T عضویت دارند. از آنجائیکه ایالات متحده آمریکا وزارتخانه‌ای بنام «پست و تلگراف و تلفن» ندارد، فرد دیگری باید نمایندگی آنرا در ITU-T بر عهده بگیرد، که این وظیفه به وزارت امور خارجه محول شده است (شاید به این دلیل که با کشورهای خارجی سروکار دارد). تعداد اعضای بخش در ITU-T به حدود ۵۰۰ می‌رسد، که شرکتهای تلفن (مانند AT&T، ودفون، ورلدکام)، تولیدکنندگان تجهیزات مخابراتی (مانند سیسکو، نوکیا، نورتل)، تولیدکنندگان کامپیوتر (مانند کامپک، سان، توشیبا)، سازندگان چیپ‌های میکروالکترونیک (مانند اینتل، موتورولا، IT)، شرکتهای رسانه‌ای (مانند AOL Time Warner، CBS، سونی)، و سایر شرکتهای علاقمند (مانند بوئینگ، سامسونگ، زیراکس) از آن جمله‌اند. تعدادی از سازمانهای علمی غیرانتفاعی و کنسرسیوم‌های صنعتی (مانند IFIP و IATA) نیز در ITU-T عضو بخش هستند. اعضای وابسته شرکتهای کوچکتري هستند، که به یکی از مباحث خاص ITU-T علاقمند هستند. نمایندگی‌ها نیز آنهایی هستند که بر امور مخابراتی نظارت دارند، مانند کمیسیون مخابرات فدرال ایالات متحده (FCC).

وظیفه ITU-T ارائه توصیه‌های فنی در زمینه تلفن، تلگراف و مخابرات داده است. این توصیه‌ها اغلب بصورت استانداردهای جهانی پذیرفته می‌شوند، مانند V.24 (در ایالات متحده به نام EIA RS-232 نیز شناخته می‌شود) که نقش و وظیفه پایه‌های رابط‌های مودمها و ترینالهای آسترون را مشخص می‌کند.

این نکته را باید تذکر داد که ITU-T فقط توصیه‌های فنی ارائه می‌کند، و دولت‌ها می‌توانند آنها را بپذیرند یا نپذیرند (همانطور که می‌دانید دولت‌ها مثل بچه‌های ده-یازده ساله هستند، یعنی دوست ندارند کسی به آنها دستور بدهد). البته کشوری که استاندارد غیر از استاندارد سایر کشورها را قبول کند، در عمل خود را از سایر کشورهای دنیا ایزوله کرده است - کاری که شاید فقط کره شمالی مایل به انجام آن باشد. احتمالاً نامگذاری استانداردهای ITU-T بنام «توصیه» فقط کلکی برای آرام کردن ملی‌گرایان کشورهای مختلف است.

کار واقعی ITU-T توسط ۱۴ گروه مطالعاتی (که اغلب آنها نزدیک به ۴۰۰ نفر عضو دارند) انجام می‌شود. سرفصل‌هایی که این گروه‌های مطالعاتی بررسی می‌کنند، از استانداردهای تهیه صورتحساب تلفن گرفته تا سرویسهای چندرسانه‌ای متغیر است. برای اینکه هر گروه بتواند کار خود را بهتر انجام دهد، به چند دسته کاری تقسیم می‌شود، که هر یک از این دسته‌ها بنوبه خود به چند تیم کارشناسی، و هر تیم تخصصی نیز به گروه‌های تخصصی تقسیم می‌شوند. (مثل اینکه هیچ وقت نمی‌توان از بوروکراسی خلاص شد!)

علیرغم این بوروکراسی، ITU-T کار خود را بخوبی انجام می‌دهد، و از اول تأسیس آن تاکنون نزدیک به ۳۰۰۰ توصیه (که بالغ بر ۶۰،۰۰۰ صفحه می‌شود) تدوین کرده است، که بسیاری از آنها بطور گسترده‌ای مورد استفاده قرار گرفته‌اند. برای مثال، استاندارد V.90 (که درباره مودمهای 56-kbps است) از توصیه‌های ITU-T می‌باشد.

با رشد و توسعه روزافزون مخابرات راه دور و جهانی شدن آن، استانداردها روز به روز اهمیت بیشتری می‌یابند، و سازمانهای بیشتری مایلند در تدوین آنها شرکت داشته باشند. برای کسب اطلاعات بیشتر درباره ITU به (Irmer, 1994) مراجعه کنید.

۲-۶-۱ مراجع مسئول استانداردهای بین‌المللی

استانداردهای بین‌المللی توسط سازمان بین‌المللی استاندارد (ISO) - که البته نام واقعی آن سازمان بین‌المللی برای تدوین استاندارد، IOS است) تهیه و منتشر می‌شوند. ISO یک سازمان غیربیمانی داوطلبانه است، در سال ۱۹۴۶ تأسیس شده، و دارای ۸۹ عضو می‌باشد. سازمانهای استاندارد ایالات متحده آمریکا (ANSI)، انگلستان (BSI)، فرانسه (AFNOR)، آلمان (DIN) و هشتاد و پنج کشور دیگر از اعضای ISO هستند.

تنوع استانداردهایی که ISO منتشر می‌کند، واقعاً جالب و حیرت‌آور است - استانداردهایی وجود دارد که کسی حتی فکر استاندارد بودن آنها را نمی‌کند، مانند استاندارد دانه‌های کاکائو (ISO 2451)، تورهای ماهیگیری (ISO 1530)، لباسهای زیر زنانه (ISO 4416) و غیره. تعداد استانداردهای ISO بالغ بر ۱۳,۰۰۰ است (که استانداردهای OSI نیز از آن جمله‌اند). ISO نزدیک به ۲۰۰ کمیته فنی (TC) دارد، که بترتیب زمان ایجاد شماره‌گذاری شده‌اند و هر کدام در موضوعی خاص تخصص دارند. برای مثال، TC1 با استانداردهای پیچ و مهره سروکار دارد، و TC97 با استانداردهای صنعت کامپیوتر و پردازش اطلاعات. هر TC به چند زیرکمیته (SC)، و هر زیرکمیته به چند گروه کاری (WG) تقسیم می‌شود.

کار اصلی ISO در WG ها (که متجاوز از ۱۰۰,۰۰۰ عضو داوطلب در سراسر دنیا دارند) انجام می‌شود. بسیاری از این «داوطلبان» در استخدام شرکتهایی هستند که محصول آنها قرار است استاندارد شود؛ برخی دیگر نیز مقامات رسمی کشورهایی هستند که مایلند روش ساخت محصولات کشورشان بصورت استانداردهای جهانی در آید. در برخی از WG ها مقامات متخصصان دانشگاهی نیز حضور دارند.

در زمینه استانداردهای صنعت مخابرات، ISO و ITU-T اغلب با یکدیگر تشریک مساعی دارند (در واقع، یکی از اعضای ITU-T است)، تا از توسعه استانداردهای ناسازگار اجتناب شود.

نماینده ایالات متحده آمریکا در ISO مؤسسه ملی استانداردهای آمریکا (ANSI) است، که برخلاف نامش یک مؤسسه غیردولتی و غیرانتفاعی است، و اعضای آن عبارتند از تولیدکنندگان، کاربرهای عمومی و شرکتهای ذینفع. ISO اغلباً استانداردهایی که توسط ANSI وضع می‌شود، را بعنوان استانداردهای بین‌المللی می‌پذیرد. رویه‌ای که در ISO برای پذیرش استانداردها مورد استفاده قرار می‌گیرد، بگونه‌ایست که بیشترین توافق اعضا را بدنبال داشته باشد. این رویه با اعلام نیاز یکی از سازمانهای عضو به یک استاندارد جدید آغاز می‌شود. بدنبال آن یک گروه کاری تشکیل می‌شود، تا برای استاندارد جدید یک پیش‌نویس کمیته (CD) تهیه کند. این CD بین سازمانهای عضو توزیع می‌شود، و آنها شش ماه فرصت دارند تا انتقادات خود را مطرح کنند. اگر اکثریت اعضا با این CD موافق باشند، یک سند اصلاح شده بنام پیش‌نویس استاندارد جهانی (DIS) برای اظهارنظر و رأی‌گیری منتشر می‌شود. بر اساس نتایج بدست آمده از این دور، متن نهایی بعنوان استاندارد جهانی (IS) تنظیم، اصلاح و منتشر می‌شود. در مورد استانداردهایی که اختلاف نظر جدی بین اعضا وجود داشته باشد، امکان دارد CD یا DIS چندین بار اصلاح و به رأی گذاشته شود، تا بتواند رأی کافی بدست آورد، و این فرآیندی است که ممکنست سالها طول بکشد.

در ایالات متحده آمریکا، مؤسسه ملی استانداردها و تکنولوژی (National Institute of Standards and Technology - NIST) که از توابع وزارت بازرگانی است، سازمان ملی استاندارد محسوب می‌شود، و تمام خریدهای دولتی (به استثنای خریدهای نظامی، که استانداردهای خاص خود را دارند) بایستی مطابق استانداردهای آن باشد.

یکی دیگر از بازیگران بزرگ در صحنه استانداردهای جهانی، مؤسسه مهندسان برق و الکترونیک (IEEE) است، که بزرگترین سازمان حرفه‌ای دنیا محسوب می‌شود. علاوه بر انتشار کتب و مجلات متعدد و برگزاری صدها کنفرانس علمی در سال، IEEE یک گروه تدوین استاندارد نیز دارد که در زمینه مهندسی برق و کامپیوتر فعالیت می‌کند. برای مثال، کمیته 802 وظیفه تدوین استاندارد شبکه‌های LAN را در IEEE بر عهده دارد، که گروههای کاری آنرا در شکل ۱-۳۸ ملاحظه می‌کنید. میزان موفقیت گروههای کاری 802 چندان بالا نیست، و داشتن عدد 802.x تضمینی برای موفقیت یک استاندارد نیست. البته استثناهایی هم در این زمینه وجود دارد، که استانداردهای 802.3 و 802.11 از آن جمله‌اند.

شماره	سرفصل
802.1	معماری LAN
802.2 ↓	کنترل لینک منطقی
802.3 *	اترنت
802.4 ↓	پاس توکن
802.5	حلقه توکن
802.6 ↓	دوباس-دو صف
802.7 ↓	گروه مشورتی تکنولوژی های پخشی
802.8 †	گروه مشورتی تکنولوژی های فیبرنوری
802.9 ↓	LAN ایزو سنکرون (برای کاربردهای زمان واقعی)
802.10 ↓	شبکه مجازی و امنیتی
802.11 *	LAN بی سیم
802.12 ↓	تقدم تقاضا (خاص)
802.13	عدد نحس! (کسی آنرا نمی خواهد)
802.14 ↓	مودم کابلی
802.15 *	شبکه های شخصی (بلووث)
802.16 *	بی سیم پاند وسیع
802.17	حلقه بسته برگشتی

شکل ۱-۳۸. گروههای کاری 802. گروههای مهم با * مشخص شده اند. آنهایی که با ↓ مشخص شده اند، به خواب زمستانی رفته اند، و گروههایی که با † مشخص شده اند، مدتهاست از صفحه روزگار محو شده اند.

۱-۳-۶ مراجع مسئول استانداردهای اینترنت

اینترنت نیز مکانیزمهای استانداردسازی خاص خود را دارد، که با ISO و ITU-T بسیار متفاوت است. مهمترین تفاوت آنها اینست که افرادی که در گردهمایی های ISO و ITU-T شرکت می کنند، کت و شلوار رسمی می پوشند، ولی شرکت کنندگان در نشست های استانداردسازی اینترنت لباس جین به تن می کنند (البته اگر تابستان نباشد، و محل اجلاس هم کنار دریا نباشد).

در گردهمایی های ISO و ITU-T اغلب نمایندگان شرکها و دولتها (که استانداردسازی شغل آنهاست) شرکت می کنند. آنها به استاندارد بعنوان «یک چیز آسمانی» نگاه می کنند، و حاضرند از جانشان برای آن مایه بگذارند. از طرف دیگر، اینترنتی ها بی نظمی را به عنوان یک اصل پذیرفته اند. اما از آنجائیکه بدون حداقلی از اشتراکات اصولاً امکان برقراری هیچ نوع تماسی وجود ندارد، استاندارد را (در کمال تأسف) چیزی لازم می دانند. وقتی آرپانت بوجود آمد، وزارت دفاع ایالات متحده آمریکا کمیته ای رسمی را مأمور نظارت بر آن کرد. در سال ۱۹۸۳، این کمیته به هیئت نظارت بر فعالیتهای اینترنتی (Internet Activity Board - IAB) تغییر نام داد، و وظیفه همسو نگه داشتن آرپانت و اینترنت بر عهده آن گذاشته شد (کاری که می توان آنرا به چوپانی یک گله گربه تشبیه کرد!). بعدها نام این کمیته به هیئت مدیره معماری اینترنت (Internet Architecture Board) تغییر کرد، بگونه ای که حروف اختصاری آن همچنان IAB باقی ماند.

به هر یک از نزدیک به ده عضو IAB موضوع مهمی برای دنبال کردن محول شد. این اعضا سالی چند بار با هم ملاقات و تبادل نظر می کردند، و گزارش کار خود را به NFS و وزارت دفاع (که پشتیبانی مالی آنرا بر عهده داشتند) ارائه می کردند. وقتی نیاز به تدوین استاندارد جدیدی حس می شد، اعضای IAB آنرا با جار و جنجال زیاد اعلام

می کردند، تا دانشجویانی که موتور نرم افزاری اینترنت محسوب می شدند، بتوانند آنرا پیاده سازی کنند. تبادل اطلاعات فنی توسط مقالات و گزارشهایی بنام نظرخواهی (Request For Comment - RFC) انجام می شد. این RFC ها (که به ترتیب انتشار شماره گذاری می شوند) در سایت www.ietf.org/rfc نگهداری می شوند، تا همه کسانی که مایلند بتوانند به آنها دسترسی داشته باشند. هم اکنون متجاوز از RFC ۳۰۰۰ در این سایت موجود است، که در این کتاب به بسیاری از آنها اشاره خواهیم کرد.

تا سال ۱۹۸۹ اینترنت چنان گسترش یافته بود، که این روش نسبتاً رسمی دیگر نمی توانست بکار آید. در آن زمان شرکتهای بسیاری محصولات TCP/IP خود را وارد بازار کرده بودند، و نمی خواستند آنها را بصرف اینکه چند پژوهشگر به خیال خود ایده های بهتری دارند، عوض کنند. در تابستان ۱۹۸۹، ساختار سازمانی IAB تغییر کرد، و به دو بخش نیروی پژوهشی اینترنت (Internet Research Task Force - IRTF) و نیروی مهندسی اینترنت (Internet Engineering Task Force - IETF) تقسیم شد، و نیروهای جدیدی (غیر از مجامع تحقیقاتی) وارد آن شدند. در سالهای اول این سازمانها دارای گردش بسته بودند: هر عضو برای دو سال انتخاب می شد، و اعضای جدید فقط توسط اعضای قدیمی منسوب می شدند. بعدها، انجمن اینترنت (Internet Society) بوجود آمد، که تمامی علاقمندان اینترنت می توانستند در آن عضویت یابند. انجمن اینترنت از جهاتی شبیه IEEE یا ACM است، و توسط یک هیئت امنای انتخابی، که اعضای IAB را منسوب می کند، اداره می شود. هدف از تقسیم IAB آن بود که IRTF بتواند به تحقیقات بلند مدت مشغول شود، در حالیکه IETF به کارهای مهندسی کوتاه مدت اشتغال دارد. IETF خود به چند گروه کاری تقسیم می شد، که هر یک روی موضوعی خاص کار می کردند. به منظور هماهنگ کردن فعالیتهای این گروهها، مدیران گروههای کاری نشستهای متعددی برگزار می کردند. سرفصلهایی که گروههای کاری IETF روی آنها کار می کردند، عبارت بودند از: برنامه های کاربردی جدید، گردآوری اطلاعات کاربران، یکپارچه سازی OSI، هدایت و آدرس دهی، امنیت، مدیریت شبکه، و استاندارد سازی. تعداد این گروههای کاری بتدریج افزایش یافت، و اکنون متجاوز از ۷۰ گروه کاری ذیل IETF مشغول به کار هستند.

علاوه بر آن، فرآیند استاندارد سازی رسمی تری (که شبیه ISO بود) پذیرفته شد. برای آن که یک ایده جدید به صورت استاندارد پیشنهادی (Proposed Standard) در آید، بایستی بطور کامل در یک RFC تشریح شود، و در جامعه اینترنت نیز علاقه کافی نسبت به آن وجود داشته باشد. سپس، برای آنکه این استاندارد به مرحله پیش نویس استاندارد (Draft Standard) ارتقاء یابد، بایستی بصورت واقعی پیاده سازی شده و به مدت حداقل ۴ ماه در ۲ سایت اینترنتی بطور همه جانبه تست شود. اگر IAB متقاعد شود که ایده اصلی خوب است و نرم افزار نیز بخوبی کار می کند، می تواند این RFC را رسماً بعنوان یک استاندارد اینترنتی اعلام کند. برخی از این استانداردها حتی بصورت استانداردهای نظامی (MIL-STD) در آمده اند، و رعایت آنها در محصولاتی که به وزارت دفاع فروخته می شوند، اجباریست. دیوید کلارک نقل قول جالبی درباره استانداردهای اینترنت دارد که اکنون بسیار معروف است: «اجماع نظری ناقص، به همراه کدی که کار می کند».

۷. واحدهای اندازه گیری

برای اجتناب از هر گونه سوء تفاهم، لازم است تأکید کنیم که در این کتاب بجای واحدهای اندازه گیری انگلیسی (با آن اندازه ها و ضرایب عجیب و غریب) از سیستم متریک استفاده شده است. در شکل ۱-۳۹ پیشنندهای اصلی سیستم متریک را مشاهده می کنید. اغلب از حروف اختصاری این پیشنندها استفاده خواهیم کرد، و واحد اعداد بزرگتر از 1 با حروف بزرگ می نویسیم، مانند KB و MB (البته با یک استثناء تاریخی: حروف اختصاری

پیشوند	عدداعشاری	توان	پیشوند	عدداعشاری	توان
Kilo	1,000	10^3	milli	0.001	10^{-3}
Mega	1,000,000	10^6	micro	0.000001	10^{-6}
Giga	1,000,000,000	10^9	nano	0.000000001	10^{-9}
Tera	1,000,000,000,000	10^{12}	pico	0.000000000001	10^{-12}
Peta	1,000,000,000,000,000	10^{15}	femto	0.000000000000001	10^{-15}
Exa	1,000,000,000,000,000,000	10^{18}	atto	0.000000000000000001	10^{-18}
Zetta	1,000,000,000,000,000,000,000	10^{21}	zepto	0.000000000000000000001	10^{-21}
Yotta	1,000,000,000,000,000,000,000,000	10^{24}	yocto	0.000000000000000000000001	10^{-24}

شکل ۱-۳۹. پیشوندهای اصلی سیستم متریک.

کیلوبیت/ثانیه بصورت kbps نوشته می‌شود). برای مثال، یک خط انتقال 1-Mbps در هر ثانیه 10^6 بیت اطلاعات را منتقل می‌کند، و 100 psec معادل 10^{-10} ثانیه است. برای تمایز بین پیشوندهای «میلی» و «میکرو» - که هر دو با m شروع می‌شوند -، اولی را با "m" و دومی را با "µ" نمایش می‌دهیم.

همچنین اشاره به این نکته خالی از فایده نیست که، واحدهایی که در صنعت کامپیوتر برای اندازه‌گیری ظرفیت حافظه، دیسک، و فایل بکار برده می‌شوند، کمی با سایر واحدها تفاوت دارند. در اینجا، «کیلو» بجای 10^3 (معادل 1000) معنای 2^{10} (یا 1024) می‌دهد، چون ظرفیت حافظه همیشه توانی از ۲ است - بنابراین، وقتی گفته می‌شود 1-KB، بمعنای ۱۰۲۴ بایت است، نه ۱۰۰۰ بایت. به همین ترتیب، 1-MB معادل 2^{20} (1,048,576 بایت) است، 1-GB معادل 2^{30} (1,073,741,824 بایت)، و 1-TB معادل 2^{40} (1,099,511,627,776 بایت). از طرف دیگر، در صحبت از سرعت انتقال داده‌ها، دیگر اعداد توانی از ۲ نیستند؛ برای مثال، خط 1-kbps دقیقاً 1000 بیت/ثانیه داده منتقل می‌کند، و یک شبکه 1-Mbps دقیقاً با سرعت 1,000,000 بیت/ثانیه کار می‌کند. متأسفانه، بسیاری از افراد قادر به تفکیک این دو نیستند (بویژه در مورد ظرفیت دیسکها). برای اجتناب از هر گونه ابهام، در این کتاب واحدهای KB، MB، و GB بترتیب معادل 2^{10} ، 2^{20} ، و 2^{30} بایت، و واحدهای kbps، Mbps، و Gbps بترتیب معادل 10^3 ، 10^6 ، و 10^9 بیت/ثانیه در نظر گرفته خواهند شد.

۸-۱ طرح کلی مباحث کتاب

در این کتاب شبکه‌های کامپیوتری از جنبه نظری و عملی مورد بررسی قرار گرفته‌اند. اکثر فصول کتاب با بحث در کلیات موضوع مورد نظر شروع شده، و با بررسی مثالها و نمونه‌های عملی ادامه می‌یابد. این نمونه‌ها از اینترنت و شبکه‌های بیسیم انتخاب شده‌اند، چون در عین اهمیت بسیار با یکدیگر تفاوت‌های اساسی نیز دارند. هر گاه لازم بوده به مثالهای دیگر نیز استناد کرده‌ایم.

در این کتاب کار خود را بر اساس مدل ترکیبی شکل ۱-۲۴ بنا نهاده‌ایم، و از فصل آینده بحث درباره سلسله مراتب پروتکل‌های این مدل را (از پائین به بالا) شروع خواهیم کرد. در فصل ۲ ابتدا پیش‌زمینه‌ای درباره سیستمهای مخابرات داده (که شامل سیستمهای کابلی، بیسیم و ماهواره‌ای می‌شود) بدست خواهیم داد. این سیستمها که به لایه فیزیکی مربوط می‌شوند، بیشتر از جنبه معماری مورد بحث قرار گرفته‌اند تا جنبه سخت‌افزاری. در این فصل نمونه‌های متعددی را مورد بررسی قرار داده‌ایم، از شبکه‌های سوئیچینگ تلفن معمولی و تلفن همراه گرفته، تا

شبکه‌های تلویزیون کابلی.

در فصل ۳ لایه پیوند داده و پروتکل‌های آن (بهمراه تحلیل این پروتکلها) به کمک مثالهای متعدد مورد بحث قرار گرفته است. پس از آن، تعدادی از پروتکل‌های مهم این لایه، از جمله HDLC (که در شبکه‌های با سرعت کم تا متوسط مورد استفاده قرار می‌گیرد) و PPP (که در اینترنت کاربرد دارد) را مورد بررسی قرار خواهیم داد.

فصل ۴ با زیرلایه دسترسی داده (که بخشی از لایه پیوند داده است) سروکار دارد. بحث اصلی در اینجا نحوه دسترسی به شبکه در جاهانیست که کانالهای فیزیکی شبکه به صورت اشتراکی مورد استفاده قرار می‌گیرند (مانند شبکه‌های LAN و برخی از شبکه‌های ماهواره‌ای). در این فصل نیز نمونه‌های متعددی (از جمله اترنت، LAN بیسیم، MAN بیسیم، بلوتوث، و شبکه‌های ماهواره‌ای) مورد بررسی قرار گرفته است. درباره پل (bridge) و سوئیچ (switch) نیز در همین فصل صحبت کرده‌ایم.

فصل ۵ به لایه شبکه، بویژه مسیریابی (routing) و الگوریتم‌های آن (استاتیک و دینامیک)، اختصاص دارد. حتی با بهترین الگوریتم‌های مسیریابی، اگر بار بیش از حد به یک شبکه تحمیل شود، امکان بروز ازدحام (congestion) در آن وجود دارد، بنابراین یکی از مباحث مهم این فصل ازدحام و راههای جلوگیری از آن (و از آن هم فراتر، تضمین نوعی کیفیت سرویس) است. نحوه اتصال شبکه‌های غیرمتجانس و رفع مشکلات آن از دیگر مباحث این فصل است. لایه شبکه از اهمیت زیادی در اینترنت برخوردار است.

در فصل ۶ درباره لایه انتقال (و بویژه پروتکل‌های اتصال-گرا) مفصلاً صحبت خواهیم کرد. حتی یکی از سرویسهای ساده این لایه بصورت عملی (بهمراه کد آن) ارائه شده است، تا با نحوه پیاده‌سازی آن بیشتر آشنا شوید. پروتکل‌های اینترنتی TCP و UDP (و کارایی آنها) نیز مفصلاً در این فصل مورد بررسی قرار گرفته‌اند. از دیگر مباحثی که در فصل ۶ مطرح خواهد شد، شبکه‌های بیسیم است.

فصل ۷ با لایه کاربرد (و پروتکلها و برنامه‌های آن) سروکار دارد. اولین مبحث این فصل، دفترچه تلفن اینترنت یعنی DNS است. ایمیل و پروتکل‌های آن مبحث بعدی فصل ۷ است. پس از آن سراغ وب می‌رویم، و مباحثی مانند محتویات استاتیک و دینامیک، فعل و انفعالات سمت مشتری و سرویس دهنده، پروتکل‌های آن، کارایی وب، و وب بیسیم را مورد بررسی قرار خواهیم داد. در پایان نیز، درباره شبکه‌های چندرسانه‌ای (صدا و تصویر جویباری - streaming -، و رادیوی اینترنتی) صحبت خواهیم کرد.

فصل ۸ درباره امنیت شبکه است. از آنجائیکه مباحث این فصل با تمام لایه‌ها سروکار دارد، آنرا در آخر (بعد از پایان بحث لایه‌های شبکه) آورده‌ایم. این فصل را با معرفی تکنولوژیهای رمزنگاری (cryptography) آغاز کرده‌ایم، و نشان داده‌ایم که چگونه می‌توان به کمک این تکنولوژیها امنیت مخابرات داده، ایمیل و وب را تأمین کرد. این فصل را با بحثی درباره تقابل بین امنیت و حریم خصوصی افراد، آزادی بیان، سانسور (و سایر موضوعاتی که با امنیت شاخ به شاخ می‌شوند) به پایان رسانده‌ایم.

در فصل ۹ مقالات و کتابهایی را که می‌توانید برای مطالعه بیشتر درباره موضوعات هر فصل به آنها مراجعه کنید، به همان ترتیب فصول کتاب آورده‌ایم. در این فصل یک کتابنامه مفصل در زمینه موضوعات مطرح شده در کتاب نیز آورده شده است.

در سایت وب مؤلف کتاب

<http://www.prenhall.com/tanenbaum>

می‌توانید لینکهای متعددی به درسته‌ها، صفحات پرسش و پاسخ، شرکتها، کنسرسیومهای صنعتی، سازمانهای حرفه‌ای، سازمانهای استاندارد، تکنولوژیها، و مقالات تخصصی پیدا کنید.

۹-۱ خلاصه

از شبکه های کامپیوتری می توان برای مقاصد مختلفی (در شرکتها، یا برای افراد عادی) استفاده کرد. در شرکتها، شبکه می تواند دسترسی به منابع اطلاعاتی را برای تمام کارکنان فراهم آورد. در این شبکه ها معمولاً از مدل مشتری-سرویس دهنده (که در آن منابع مشترک روی کامپیوترهای قدرتمندی موسوم به سرویس دهنده - server - قرار می گیرند) استفاده می شود. شبکه برای افراد عادی امکان دسترسی به منابع اطلاعاتی یا تفریحی را فراهم می آورد. امروزه افراد بسیاری با یک مودم از منزل خود به شرکتهای خدمات اینترنتی (ISP) متصل شده، و از امکانات آن استفاده می کنند. یکی از زمینه هایی که امروزه سرعت رو به گسترش است، شبکه های بیسیم است، که امکان دسترسی به اینترنت را حتی از تلفنهای همراه به افراد می دهد.

در یک تقسیم بندی بسیار کلی، شبکه ها را می توان به شبکه های محلی (LAN)، شبکه های شهری (MAN)، شبکه های گسترده (WAN) و شبکه های (internetwork) تقسیم کرد، که هر کدام برای خود دارای ویژگیها، سرعت، تکنولوژی و جایگاه خاصی می باشد. شبکه محلی (LAN) سرعت بالایی دارد، ولی معمولاً به یک ساختمان منفرد محدود می شود. محدوده جغرافیایی شبکه شهری (MAN) - همانطور که از نام آن برمی آید - یک شهر است؛ تلویزیون کابلی نمونه ای از شبکه های شهری است. شبکه های گسترده (WAN) یک کشور یا قاره را در بر می گیرند. شبکه های LAN و MAN سوئیچ نشده هستند (یعنی، مسیریاب - router - ندارند)، در حالیکه شبکه های WAN سوئیچ شده اند. امروزه شبکه های بیسیم (بویژه LAN بیسیم) از محبوبیت روزافزونی برخوردارند. از اتصال چند شبکه به یکدیگر نیز یک شبکه شبکه ها شکل می گیرد.

نرم افزار شبکه از پروتکلها (که قواعد حاکم بر ارتباط پروسس ها را تعیین می کنند) تشکیل می شود. پروتکلها یا اتصال-گرا (connection-oriented) هستند یا غیرمتصل (connectionless). اغلب شبکه ها از مدل سلسله مراتبی پروتکلها استفاده می کنند، که در آن هر لایه سرویسهایی را در اختیار لایه های بالاتر قرار می دهد، و آنها از جزئیات کار در لایه های پایینتر بیخبر می کند. مجموعه پروتکل (protocol stack) های مهم امروزی عمدتاً بر مبنای دو مدل OSI و TCP/IP بنا شده اند. هر دوی این مدلها دارای لایه های شبکه، انتقال و کاربرد هستند، ولی در لایه های دیگر با هم فرق دارند. هنگام طراحی پروتکلها باید به مسائلی از قبیل مالتی پلکس کردن (multiplexing)، کنترل جریان (flow control)، کنترل خطا (error control) و مانند آنها توجه ویژه مبذول داشت. (بخش عمده ای از این کتاب به پروتکلها و طراحی آنها اختصاص دارد).

وظیفه شبکه ارائه سرویس به کاربران است، که این سرویسها نیز می توانند اتصال-گرا یا غیرمتصل باشند. در برخی شبکه ها، یک لایه سرویس غیرمتصل در اختیار می گذارد، در حالیکه لایه بالاتر سرویس اتصال-گرا ارائه می کند.

معروفترین شبکه های موجود عبارتند از: اینترنت، ATM، اترنت و LAN بیسیم (IEEE 802.11). اینترنت محصول تکامل شبکه آرپانت (ARPANET - شبکه سوئیچ بسته وزارت دفاع ایالات متحده آمریکا) بود اینترنت در واقع امروزه شبکه ایست از هزاران شبکه دیگر، نه یک شبکه واحد. مشخصه اصلی اینترنت استفاده از مجموعه پروتکل TCP/IP است. شبکه های ATM بیشتر در سیستمهای تلفن، و برای مخابرات راه دور، بکار می روند. اترنت (Ethernet) نیز محبوبترین تکنولوژی LAN است، که در اغلب شرکتها و دانشگاهها از آن استفاده می شود. و بالاخره، شبکه های LAN بیسیم که (با سرعت خیره کننده 54 Mbps) سرعت در حال گسترش هستند.

برای آنکه چند کامپیوتر بتوانند با هم ارتباط برقرار کنند، مهمترین موضوع وجود استانداردهای متعدد (سخت افزاری و نرم افزاری) است. وظیفه تدوین استاندارد برای شبکه های کامپیوتری بر عهده سازمانهایی از قبیل ITU-T، ISO، IEEE و IAB گذاشته شده است.

مسائل

۱. فرض کنید سگی از نژاد سنت برنارد (نوعی سگ قوی هیکل که در کشور کوهستانی سوئیس برای نجات کوهنوردان ساخته دیده تربیت می شود) بنام پرنی دارید، و او را برای حمل جعبه ای حاوی سه نوار 8mm آموزش داده اید (به هر حل پُر شدن دیسک هم نوعی موقعیت اورژانس است). هر نوار 7 GB ظرفیت دارد، و پرنی می تواند تحت هر شرایطی با سرعت 18 km/h حرکت کند. تا چه فاصله ای نرخ انتقال اطلاعات پرنی همچنان از خط انتقالی با سرعت 150 Mbps (بدون در نظر گرفتن سرآیند) بیشتر است؟
۲. یکی از گزینه های رقیب شبکه های LAN سیستم های تسهیم زمانی (time sharing) بزرگ (بهمراه ترمینالهایی که کاربران از آنها استفاده می کنند) است. دو مزیت شبکه های LAN مبتنی بر مدل مشتری سرویس دهنده را نسبت به سیستم های فوق بیان کنید.
۳. کارایی سیستم های مشتری سرویس دهنده به دو ویژگی مهم شبکه وابسته است: پهنای باند (bandwidth) - حداکثر تعداد بیت هایی که شبکه می تواند در هر ثانیه منتقل کند، و زمان تأخیر (latency time) - مدت زمانی که طول می کشد تا اولین بیت از مشتری به سرویس دهنده - یا بالعکس - برسد. دو مثال از شبکه ای با پهنای باند زیاد و زمان تأخیر زیاد، و پهنای باند کم و زمان تأخیر کم بزنید.
۴. غیر از پهنای باند زیاد و زمان تأخیر کم، چه شرایط دیگری باید فراهم باشد تا یک شبکه کیفیت مناسبی برای سرویس صدای دیجیتال داشته باشد؟
۵. یک از عوامل مهم در میزان تأخیر سیستم های سوئیچینگ بسته مبتنی بر مدل ذخیره-هدایت، مدت زمان نیست که صرف ذخیره و هدایت بسته در یک سوئیچ می شود. آیا زمان سوئیچینگ 10 μsec برای یک سیستم مشتری سرویس دهنده که بین کالیفرنیا و نیویورک کار می کند، عامل مهمی محسوب می شود. سرعت انتشار امواج الکترومغناطیس در سیم مسی را 0.667 سرعت نور در نظر بگیرید.
۶. یک سیستم مشتری سرویس دهنده از شبکه ماهواره ای، که ماهواره آن در ارتفاع 40,000 km سطح زمین قرار گرفته، استفاده می کند. زمان تأخیر پاسخ در بهترین حالت چقدر است؟
۷. در آینده ای نه چندان دور مردم می توانند مستقیماً از طریق ترمینالهای متصل به شبکه های کامپیوتری به لوايح و طرحهای قانونی رأی بدهند، و دیگر نیازی به مجالس قانونگذاری وجود نخواهد داشت. جنبه های مثبت چنین دموکراسی مستقیمی نسبتاً روشن است؛ کمی درباره نقاط منفی آن بحث کنید.
۸. می خواهیم پنج مسیریاب را در یک زیرشبکه نقطه-به-نقطه به هم وصل کنیم. هر زوج از این مسیریاب ها را می توان با یک خط پُر سرعت، خطی با سرعت متوسط، و یا یک خط کم سرعت به هم متصل کرد، و یا اینکه اصلاً آنها را به هم وصل نکرد. اگر یک کامپیوتر بتواند هر توپولوژی را در 100 ms طراحی و بررسی کند، چه مدت طول می کشد تا تمام توپولوژیهای ممکن را بررسی کند؟
۹. تعداد 1 - 2ⁿ مسیریاب در یک درخت باینری متقارن (یک مسیریاب در هر گره) به هم متصل شده اند. مسیریاب i برای ارتباط با مسیریاب j، باید ابتدا پیام خود را به ریشه این درخت بفرستد، تا از آنجا بدست مسیریاب j برسد. (با فرض اینکه تمام مسیریاب یکسان هستند) عبارتی بنویسید که تعداد متوسط پرشهای لازم برای رسیدن پیام یک مسیریاب به مسیریاب دیگر را (برای n های بزرگ) بدست دهد.
۱۰. یکی از نقاط منفی زیرشبکه های پخش ظرفیتی است که در اثر اقدام به پخش همزمان توسط چند کامپیوتر از دست می رود (تصادم - collision). اجازه دهید مسئله را ساده کرده، و فرض کنیم زمان به پرشهای مساوی تقسیم شده، و در هر برش زمانی n کامپیوتر با احتمال p اقدام به پخش روی شبکه می کنند. چه

- کسری از برشهای زمانی در اثر بروز حالت تصادم تلف خواهند شد؟
۱۱. دو دلیل برای استفاده از پروتکل‌های لایه‌ای ارائه کنید.
۱۲. رئیس یک شرکت رنگسازی ایده جدیدی برای تولید یک محصول بدیع و نوظهور دارد. موضوع را با اداره حقوقی شرکت در میان می‌گذارد، و آنها هم از اداره مهندسی درخواست کمک می‌کنند. سرپرست اداره مهندسی درباره زوایای مختلف این طرح با یکی از هم‌تایان خود در شرکتی دیگر مشورت می‌کند، که وی نیز موضوع را به اداره حقوقی شرکت خود منتقل می‌کند. در پایان نیز، رؤسای دو شرکت بر سر جزئیات مالی معامله با یکدیگر به مذاکره می‌پردازند. آیا می‌توان این سناریو را مطابق با پروتکل‌های چند لایه مدل OSI دانست؟
۱۳. دو تفاوت عمده بین ارتباطات اتصال-گرا و غیرمتصل چیست؟
۱۴. دو شبکه سرویس‌های اتصال-گرای قابل اعتماد ارائه می‌کنند: یکی از آنها بصورت استریم بایت، و دیگری بصورت استریم پیام. آیا این دو یکسان هستند؟ اگر پاسخ مثبت است، چه دلیلی برای نامگذاری جداگانه آنها وجود دارد؟ اگر خیر، تفاوت آنها در چیست؟
۱۵. در بحث پروتکل‌های شبکه، «مذاکره» (negotiation) چه معنایی دارد؟ یک مثال بزنید.
۱۶. در شکل ۱-۱۹ یک سرویس نشان داده شده است. آیا سرویس دیگری در این شکل وجود دارد؟ اگر آری، کجا؟ اگر خیر، چرا؟
۱۷. در برخی از شبکه‌ها، لایه پیوند داده با درخواست ارسال مجدد فریم‌هایی که بدرستی دریافت نشده‌اند، نوعی کنترل خطا انجام می‌دهد. اگر احتمال خراب شدن یک فریم p باشد، متوسط تعداد دفعاتی که یک فریم باید فرستاده شود، چقدر است؟ فرض کنید فریم‌های تصدیق دریافت (acknowledgement) هرگز خراب نمی‌شوند.
۱۸. کدامیک از لایه‌های OSI وظایف ذیل را بر عهده دارند:
(الف) تقسیم استریم بیت‌ها به فریم
(ب) تعیین مسیری در زیرشبکه، که باید از آن استفاده شود
۱۹. اگر واحد تبادل داده در لایه پیوند داده فریم، و در لایه انتقال بسته نام داشته باشد، فریم‌ها در بسته پیچیده می‌شوند، یا بسته‌ها در فریم؟ توضیح دهید.
۲۰. ساختار سلسله مراتبی پروتکل‌های یک سیستم n لایه دارد. برنامه‌های کاربردی در این سیستم پیام‌هایی بطول M بایت تولید می‌کنند، و در هر لایه یک سرآیند h بایتی به پیام لایه بالاتر اضافه می‌شود. چه کسری از پهنای باند شبکه را این سرآیندها اشغال می‌کنند؟
۲۱. دو شباهت و دو تفاوت مدل‌های مرجع را OSI و TCP/IP را بیان کنید.
۲۲. تفاوت اصلی TCP و UDP چیست؟
۲۳. زیرشبکه شکل ۱-۲۵ (ب) طوری طراحی شده که در مقابل حملات هسته‌ای دوام بیاورد. چند بمب اتمی لازم است تا این زیرشبکه به دو قسمت کاملاً مجزا تقسیم شود؟ فرض کنید برای نابود کردن هر گره و تمام لینک‌های متصل به آن یک بمب اتمی کافیت.
۲۴. تخمین زده شده است که اینترنت هر ۱۸ ماه دو برابر می‌شود. (و با اینکه کسی واقعاً تعداد آنها را نمی‌داند) طبق برآوردهای تخمینی در سال ۲۰۰۱ تعداد کامپیوترهای اینترنت ۱۰۰ میلیون بوده است. با استفاده از این

مفروضات، تعداد کامپیوترهای اینترنت در سال ۲۰۱۰ را محاسبه کنید. آیا این عدد را باور می‌کنید؟ توضیح دهید.

۲۵. هنگام تبادل فایل بین دو کامپیوتر، دو استراتژی تصدیق دریافت ممکن است. در روش اول، فایل به قطعاتی تقسیم شده، و برای هر قطعه تصدیق دریافت جداگانه مطالبه می‌شود، اما برای کل فایل، خیر. در روش دوم، برای تک تک قطعات تصدیق دریافت مطالبه نمی‌شود، ولی در پایان دریافت کل فایل باید به تأیید طرف مقابل برسد. درباره این دو روش بحث کنید.

۲۶. چرا ATM از سلولهای کوچک و با اندازه ثابت استفاده می‌کند؟

۲۷. هر بیت در استاندارد اولیه 802.3 چند متر طول داشت؟ سرعت کار این شبکه را 10 Mbps، و سرعت سیگنالهای الکتریکی در کابل کواکسیال را 2/3 سرعت نور در خلاء فرض کنید.

۲۸. ابعاد یک تصویر 1024 × 768 پیکسل است، و هر پیکسل آن 3 بایت جا می‌گیرد (و فرض کنید این تصویر فشرده نشده است). انتقال این تصویر روی یک خط 56-kbps چقدر طول می‌کشد؟ روی یک خط 1-Mbps چقدر؟ روی شبکه اینترنت 10-Mbps چقدر؟ و روی شبکه اینترنت 100-Mbps چقدر؟

۲۹. اینترنت و شبکه‌های بیسیم شباهتها و تفاوتهایی دارند. یکی از ویژگیهای اینترنت اینست که در هر لحظه فقط یک فریم می‌تواند روی شبکه منتقل شود. آیا در شبکه‌های 802.11 نیز چنین است؟ توضیح دهید.

۳۰. نصب شبکه‌های بیسیم بسیار ساده است، که همین باعث ارزانی آنها شده است، چون معمولاً هزینه‌های نصب در مقابل هزینه تجهیزات رقم قابل توجهی را تشکیل می‌دهد. با این حال، این نوع شبکه معایبی نیز دارد. دو تا از این معایب را نام ببرید.

۳۱. دو مزیت و دو عیب برای استاندارد کردن پروتکل‌های شبکه بشمارید.

۳۲. وقتی یک سیستم از دو قسمت ثابت و متحرک تشکیل می‌شود (مانند درایو CD، و دیسک CD-ROM)، استاندارد کردن آن از اهمیت زیادی برخوردار است، تا محصولات شرکتهای مختلفی که این قطعات را تولید می‌کنند، با هم سازگار باشد. سه مثال از خارج صنعت کامپیوتر بزنید، که چنین استانداردی در آنها وجود دارد. حال سه مثال از خارج صنعت کامپیوتر بزنید، که چنین استانداردی در آنها وجود نداشته باشد.

۳۳. تمام فعالیتهای خود را که در طول شبانه روز بنحوی با شبکه سروکار دارد، فهرست وار بنویسید. اگر این شبکه‌ها به یکباره از کار بیفتند، چه تأثیری روی زندگی شما خواهد گذاشت؟

۳۴. شبکه‌هایی را که در محل کار یا تحصیل خود می‌شناسید (بهمراه نوع، توپولوژی، و روش سونچینگ آنها)، مشخص کنید.

۳۵. برنامه ping به شما اجازه می‌دهد تا بسته‌ای را به یک مقصد مشخص فرستاده، و زمان رفت و برگشت آنرا محاسبه کنید. از این برنامه برای تعیین زمان رفت و برگشت بسته‌ها به چند نقطه مختلف استفاده کنید. با استفاده از این اطلاعات، نمودار زمان انتقال بسته‌ها روی اینترنت (در جهت رفت) را بر حسب فاصله رسم کنید. بهتر است برای این منظور از دانشگاههای مهم و سرشناس (که محل آنها دقیقاً مشخص است) استفاده کنید. برای مثال، سایت berkeley.edu در برکلی - کالیفرنیا است، سایت mit.edu در کمبریج - ماساچوست،

سایت vu.nl در آمستردام-هلند، سایت www.usyd.edu.au در سیدنی-استرالیا، و سایت www.uct.ac.za در کیپ‌تاون-آفریقای جنوبی.

۳۶. سری به سایت IETF (به آدرس www.ietf.org) بزنید، و ببینید چکار می‌کنند. یکی از پروژه‌هایی را که به آن علاقه دارید، انتخاب کرده، و درباره آن مقاله‌ای نیم صفحه‌ای بنویسید.

۳۷. استانداردهای در دنیای شبکه‌های کامپیوتری بسیار مهم است، و همانطور که دیدید سازمانهای ITU و ISO این وظیفه را بر عهده دارند. به سایت وب این سازمانها (www.iso.org و www.iyu.org) رفته، و با کارهای استانداردهای آنها آشنا شوید. گزارش کوتاهی درباره انواع چیزهایی که این دو سازمان استاندارد کرده‌اند، بنویسید.

۳۸. اینترنت از تعداد زیادی شبکه‌های مختلف تشکیل شده است، که آرایش آنها توپولوژی اینترنت را مشخص می‌کند. در خود اینترنت اطلاعات زیادی در زمینه توپولوژی آن وجود دارد. با استفاده از یک موتور جستجو (search engine) در باره این موضوع تحقیق کرده، و خلاصه‌ای از یافته‌های خود را در یک گزارش بنویسید.

لایه فیزیکی



در این فصل پانین ترین لایه سلسله مراتب شکل ۱-۲۴ را مورد بررسی قرار خواهیم داد. مشخصات مکانیکی، الکتریکی و تایمینگ (همزمانی) شبکه در این لایه تعریف می شود. برای شروع کمی درباره تئوری مخابرات صحبت می کنیم، فقط برای اینکه نشان دهیم دست طبیعت چه محدودیتهایی را در زمینه انتقال داده ها به ما تحمیل کرده است.

سیس با رسانه های فیزیکی که برای انتقال داده ها از آنها استفاده می شود، آشنا می شوید: رسانه های هدایت پذیر (مانند سیم مسی و فیبر نوری)، بیسیم (امواج رادیویی زمینی)، و ماهواره. آشنایی با این رسانه ها برای شناخت تکنولوژیهای مدرن مخابراتی اهمیت اساسی دارد.

در ادامه سه نمونه از سیستمهای مخابراتی که در شبکه های کامپیوتری کاربرد گسترده ای دارند، را مفصلاً مورد بررسی قرار خواهیم داد: سیستم تلفن ثابت، شبکه تلفن همراه، و تلویزیون کابلی. در تمام این سیستمها از فیبر نوری بعنوان ستون فقرات (backbone) استفاده می شود، ولی تکنولوژی بکار رفته در آخرین قطعه اتصال (از مرکز تلفن یا ایستگاه توزیع تا مصرف کننده) در آنها متفاوت است.

۱-۲ مبانی نظری مخابرات داده

برای انتقال اطلاعات روی سیم می توان از متغیرهای الکتریکی مانند ولتاژ یا جریان استفاده کرد. با نمایش تغییرات این ولتاژ یا جریان بصورت تابعی از زمان، می توان رفتار سیگنال را مدل سازی و تحلیل ریاضی کرد. موضوع این بخش تحلیل ریاضی سیگنالهای الکتریکی است.

۱-۱-۲ آنالیز فوریه

در اوایل قرن نوزدهم میلادی، ریاضیدان فرانسوی ژان بابتیست فوریه ثابت کرد که هر تابع متناوب، $g(t)$ ، با دوره تناوب T را می توان به صورت مجموع بینهایت جمله سینوسی و کسینوسی نوشت:

$$g(t) = \frac{1}{T} c + \sum_{n=1}^{\infty} a_n \sin(\gamma \pi n f t) + \sum_{n=1}^{\infty} b_n \cos(\gamma \pi n f t) \quad (1-2)$$

که در آن $f = 1/T$ فرکانس اصلی، a_n و b_n ضرایب هارمونی (جمله) n ام تابع سینوس و کسینوس، و c یک عدد ثابت است. به تجزیه یک تابع به مجموع بینهایت جمله سینوسی و کسینوسی سری فوریه (Fourier series) گفته می شود. اگر ضرایب جملات سری فوریه یک تابع و دوره تناوب آن (T) معلوم باشد، می توان این تابع را با

استفاده از معادله (۱-۲) ایجاد کرد.

اگر یک سیگنال داده دارای دوره محدودی باشد (که همیشه هم چنین است)، می‌توان آنرا تکرار بینهایت یک الگوی مشخص فرض کرد (بعبارت دیگر، این موج در فاصله زمانی T تا $2T$ با فاصله 0 تا T کاملاً یکسان است). برای محاسبه ضریب a_n کافیه دو طرف معادله (2-1) را در $\sin(2\pi kft)$ ضرب کرده، و سپس در فاصله 0 تا T از آن انتگرال بگیریم. از آنجائیکه

$$\int_0^T \sin(\gamma \pi kft) \sin(\gamma \pi nft) dt = \begin{cases} 0 & \text{for } k \neq n \\ T/\gamma & \text{for } k = n \end{cases}$$

فقط جمله‌ای که ضریب a_n دارد، باقی می‌ماند، و جملات b_n بکلی از بین خواهند رفت. به همین ترتیب، اگر دو طرف معادله (2-1) را در $\cos(2\pi kft)$ ضرب کرده، و سپس در فاصله 0 تا T از آن انتگرال بگیریم، ضریب b_n بدست خواهد آمد. برای بدست آوردن ثابت c هم کافیه از تابع $g(t)$ در همین فاصله زمانی انتگرال بگیریم. خلاصه عملیات فوق را در فرمولهای ذیل مشاهده می‌کنید:

$$a_n = \frac{\gamma}{T} \int_0^T g(t) \sin(\gamma \pi nft) dt \quad b_n = \frac{\gamma}{T} \int_0^T g(t) \cos(\gamma \pi nft) dt \quad c = \frac{\gamma}{T} \int_0^T g(t) dt$$

۲-۱-۲ محدودیت پهنای باند

اما برای اینکه ببینید اینها چه ربطی به مخابرات داده دارد، اجازه دهید یک مثال بزنیم: مخابره کاراکتر آسکی حرف "b" که بصورت ۸ بیتی کُد شده است. حرف "b" در استاندارد آسکی ۸ بیتی بصورت 01100010 کُد می‌شود، و اینها بیتهایی هستند که باید مخابره شوند. در نمودار سمت چپ شکل ۱-۲ (الف) ولتاژهای خروجی کامپیوتر برای حرف "b" را می‌بینید. با استفاده از آنالیز فوریه، ضرایب ثابت این سیگنال بصورت زیر بدست می‌آیند:

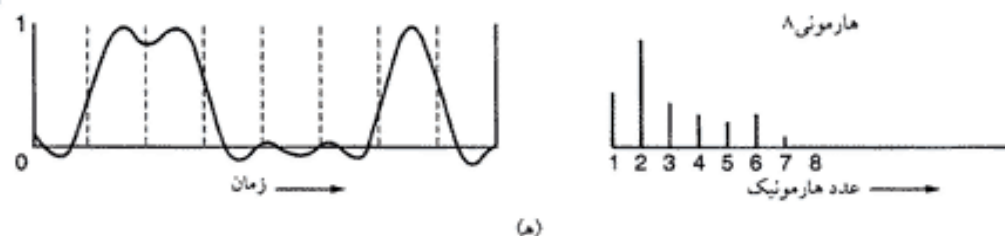
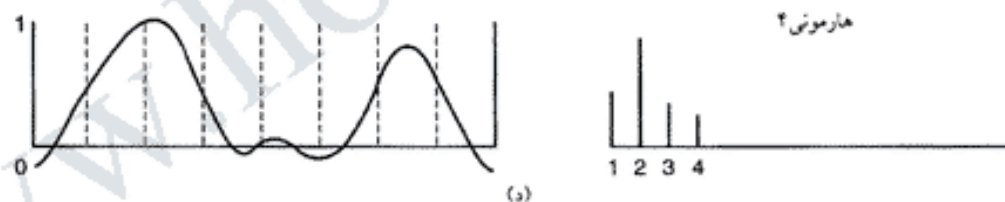
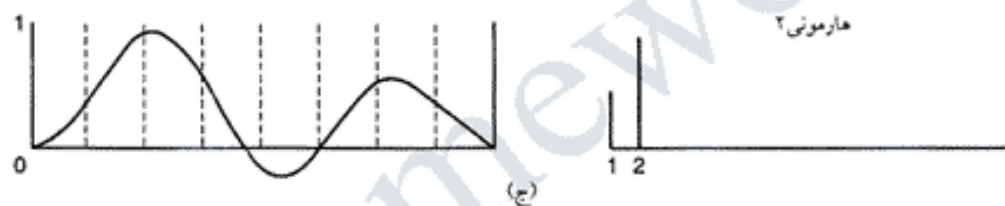
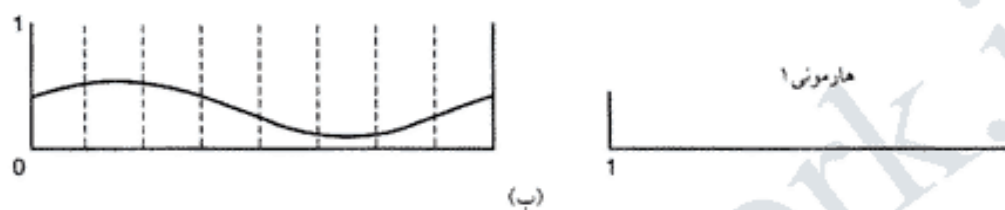
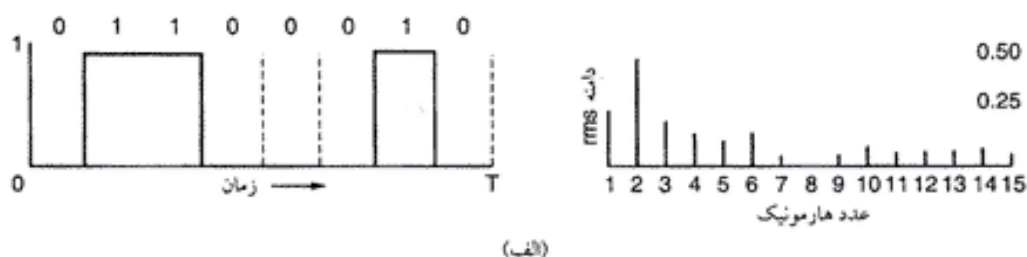
$$a_n = \frac{1}{\pi n} [\cos(\pi n/4) - \cos(3\pi n/4) + \cos(6\pi n/4) - \cos(7\pi n/4)]$$

$$b_n = \frac{1}{\pi n} [\sin(3\pi n/4) - \sin(\pi n/4) + \sin(7\pi n/4) - \sin(6\pi n/4)]$$

$$c = 3/4$$

در نمودار سمت راست شکل ۱-۲ (الف) مقدار rms ضرایب فوریه، یعنی $\sqrt{a_n^2 + b_n^2}$ ، چند هارمونی اول این سیگنال را مشاهده می‌کنید. علت اهمیت این ضرایب آنست که انرژی انتشار سیگنال با مجذور آنها (در یک فرکانس خاص) متناسب است.

از طرف دیگر می‌دانیم که هیچ سیستم پخشی بدون اتلاف انرژی نیست، ولی اگر تمام ضرایب فوریه به یکسان تضعیف شوند، سیگنال حاصله فقط ضعیفتر می‌شود، و بهیچوجه دچار اعوجاج (تغییر شکل) نخواهد شد. متأسفانه، در سیستمهای پخش مختلف ضرایب فوریه به میزانهای متفاوت تضعیف می‌شوند، و همین باعث بروز اعوجاج در شکل موج خواهد شد. معمولاً، تضعیف موج فقط از فرکانس خاصی به بالا (که به فرکانس قطع - f_c معروف است و با هرتز - Hz - یا سیکل بر ثانیه - cycles/sec - سنجیده می‌شود) روی می‌دهد، و زیر این فرکانس تضعیف موج وجود ندارد. طیف فرکانسی که موج می‌تواند بدون تضعیف منتشر شود، به پهنای باند (bandwidth) معروف است. از آنجائیکه در عمل فرکانس قطع عددی دقیق و قطعی نیست، پهنای باند بعنوان فرکانسی که در آن فقط نیمی از انرژی موج عبور می‌کند، تعریف می‌شود.



شکل ۲-۱. (الف) یک سیگنال بسایتری و ضرایب فوریه rms آن. (ب)-(ه) تقریب‌های متوالی سیگنال اولیه.

پهنای باند یکی از خواص فیزیکی رسانه انتقال است، و معمولاً به نوع، شکل، ضخامت و طول آن بستگی دارد. در برخی موارد با قرار دادن یک فیلتر پهنای باندی را که در اختیار مشتری است، عملاً محدود می‌کنند. برای مثال، پهنای باند یک خط تلفن در فواصل کوتاه می‌تواند به 1 MHz برسد، ولی شرکت‌های تلفن با قرار دادن

فیلترهای ویژه آنرا به حدود 3100 Hz محدود می کنند. این پهنای باند برای انتقال واضح و مفهوم صدا کافیتست، و از اتلاف منابع نیز جلوگیری می کند.

حال اجازه دهید ببینیم اگر پهنای باند خط انتقال آنقدر کم باشد که فقط فرکانسهای پائین بتوانند منتقل شوند (بعبارت دیگر فقط هارمونی های اول معادله (2-1) اجازه عبور داشته باشند)، سیگنال شکل ۱-۲ (الف) به چه صورتی در می آید. در شکل ۱-۲ (ب) شکل موج را در حالتی که فقط هارمونی اول (هارمونی اصلی، f) اجازه عبور دارد، می بینید. در شکل های ۱-۲ (ج) - (ه) نیز شکل موج برای پهنای باندهای بالاتر نشان داده شده اند. اگر نرخ انتقال اطلاعات b bits/sec باشد، زمان لازم برای ارسال (مثلاً) ۸ بیت معادل $8/b$ sec است، بنابراین فرکانس هارمونی اصلی $b/8$ Hz خواهد بود. همانطور که گفتیم، فرکانس یک خط تلفن معمولی (که به خط رده صوتی - voice-grade - معروف است) بصورت مصنوعی کمی بالای 3000 Hz نگه داشته می شود. این محدودیت باعث می شود تا بالاترین هارمونی که می تواند از این خط عبور کند، تقریباً $3000/(b/8)$ یا $24000/b$ باشد.

در شکل ۲-۲ ۲ اعداد محاسبه شده برای برخی از مهمترین نرخهای انتقال را می بینید. همانطور که از این شکل برمی آید، اگر بخواهیم داده ها را با سرعت 9600 bps روی یک خط رده صوتی بفرستیم، موج ما از شکل ۱-۲ (الف) به شکل ۱-۲ (ه) تبدیل خواهد شد - که باعث می شود تشخیص بیت های اولیه قدری مشکل شود. از همینجا پیداست که برای سرعت های بالای 38.4 kbps هیچ شانسی برای ارسال داده های باینری وجود ندارد (حتی اگر خط ما کاملاً بدون نویز باشد). بعبارت دیگر، محدود کردن پهنای باند باعث محدود شدن نرخ انتقال اطلاعات (حتی در بهترین کانالها) خواهد شد. با این حال، راههایی وجود دارد (مثلاً، استفاده از چند سطح ولتاژ بجای در سطح ولتاژ) که می توان به نرخهای انتقال بالاتری دست یافت، و در ادامه همین فصل درباره آنها صحبت خواهیم کرد.

۳-۱-۲ حداکثر نرخ داده در یک کانال

در سال ۱۹۲۴ یکی از مهندسان AT&T بنام هنری نایکوئیست دریافت که حتی بهترین کانال انتقال هم ظرفیت محدودی دارد، و توانست معادله حداکثر نرخ انتقال داده یک کانال بدون نویز را برای یک پهنای باند مشخص بدست آورد. در سال ۱۹۴۸ کلود شانون کار نایکوئیست را تکمیل کرد، و معادله وی را برای حالتی که کانال در معرض نویز تصادفی (نویز ترمودینامیک یا حرارتی) است، توسعه دهد (Shannon, 1948). در این قسمت کارهای این دو دانشمند را بطور مختصر بررسی خواهیم کرد.

Bps	T (msec)	اولین هارمونی (Hz)	تعداد هارمونی فرستاده شده
300	26.67	37.5	80
600	13.33	75	40
1200	6.67	150	20
2400	3.33	300	10
4800	1.67	600	5
9600	0.83	1200	2
19200	0.42	2400	1
38400	0.21	4800	0

شکل ۲-۲. رابطه بین نرخ انتقال داده و هارمونی ها.

نایکونیست ثابت کرد که اگر سیگنالی از یک فیلتر پائین‌گذر (low-pass filter) با پهنای باند H عبور داده شود، سیگنال خروجی (فیلتر شده) را می‌توان با داشتن فقط $2H$ نمونه در ثانیه بازسازی کرد. نمونه‌برداری از خط با نرخ بالاتر از $2H$ دقت بیشتری به همراه نخواهد داشت، چون فرکانسهای بالاتری که بدین ترتیب می‌توان بازیابی کرد، قبلاً توسط فیلتر حذف شده‌اند. اگر این سیگنال دارای V سطح مجزا باشد، طبق قضیه نایکونیست:

$$\text{maximum data rate} = 2H \log_2 V \text{ bits/sec}$$

برای مثال، یک کانال 3-kHz بدون نویز نمی‌تواند سیگنالهای باینری (دو سطح ولتاژ) بالاتر از 6000 bps را انتقال دهد.

تا اینجا کانالها را بدون نویز فرض کردیم، که در واقع چنین نیست. اگر در کانال انتقال نویز تصادفی (حرارتی) وجود داشته باشد، اوضاع بسرعت رو به وخامت خواهد گذاشت (و می‌دانیم که هیچ گریزی از نویز حرارتی، که حاصل حرکت مولکولهاست، نیست). مقدار نویز حرارتی با نسبت توان سیگنال به توان نویز سنجیده می‌شود، و به نسبت سیگنال به نویز (signal-to-noise ratio) معروف است. اگر توان سیگنال را با S و توان نویز را با N نشان دهیم، نسبت سیگنال به نویز S/N خواهد بود. معمولاً آنچه که به عنوان نسبت سیگنال به نویز داده می‌شود، نه خود S/N بلکه $10 \log_{10} S/N$ است، که به دسی‌بل (decibel) معروف است و با dB نشان داده می‌شود. اگر نسبت سیگنال به نویز 10 باشد، معادل 10 دسی‌بل خواهد بود، نسبت 100 معادل 20 دسی‌بل است، نسبت 1000 معادل 30 دسی‌بل، و الی آخر. سازندگان تقویت‌کننده‌های استریو اغلب پهنای باندی را که دستگاه آنها بصورت خطی عمل می‌کند، بصورت فرکانس 3 دسی‌بل در هر انتهای طیف مشخص می‌کنند (فرکانس 3 دسی‌بل فرکانسی است که ضریب تقویت دستگاه نصف می‌شود، چون $\log_{10} 3 = 0.5$).

شانون معادله نایکونیست را برای کانالهای نویزدار بصورت زیر تصحیح کرد:

$$\text{maximum number of bits/sec} = H \log_2 (1 + S/N)$$

که در آن H پهنای باند کانال (بر حسب هرتز)، و S/N نسبت سیگنال به نویز است. برای مثال، در کانالی با پهنای باند 3000 Hz که دارای نویز حرارتی 30 dB است (نویز معمول در بخش آنالوگ سیستمهای تلفن)، هرگز نمی‌توان بیش از 30,000 bps داده ارسال کرد (بدون توجه به اینکه تعداد سطوح ولتاژ این سیگنال یا تعداد نمونه‌برداری‌ها چقدر باشد).

معادله شانون از تئوری اطلاعات (information theory) استنتاج شده، و برای هر کانالی که تحت تأثیر نویز حرارتی باشد صادق است. کانالی که این قاعده را نقض کند، باید معادل یک ماشین حرکت دائم فرض کرد (ماشینی که طبق اصول ترمودینامیک نمی‌تواند وجود خارجی داشته باشد). البته باید توجه داشته باشید که این عدد فقط حد بالای ظرفیت کانال را مشخص می‌کند، و سیستمهای واقعی بندرت به این حد دست پیدا می‌کنند.

۲-۲ رسانه انتقال هدایت‌پذیر

وظیفه لایه فیزیکی انتقال بیت‌های خام از یک ماشین به ماشین دیگر است. برای اینکار از رسانه‌های فیزیکی مختلفی می‌توان استفاده کرد، که هر کدام پهنای باند، تأخیر انتشار، هزینه، و سهولت نصب و نگهداری خاص خود را دارند. این رسانه‌ها را می‌توان به دو دسته کلی هدایت‌پذیر (مانند سیم مسی یا فیبر نوری) و هدایت‌ناپذیر (مانند امواج رادیویی و لیزری) تقسیم‌بندی کرد. در قسمتهای آینده این رسانه‌ها را مورد بررسی قرار خواهیم داد.

۱-۲-۲ رسانه مغناطیسی

یکی از متداولترین راههای انتقال اطلاعات از کامپیوتری به کامپیوتر دیگر، نوشتن آنها روی نوار مغناطیسی، دیسک، CD یا DVD و سپس خواندن آنها در کامپیوتر مقصد است. اگرچه این روش بخوبی استفاده از مخابرات

ماهوره‌ای نیست، ولی هزینه آن بسیار کمتر است، بویژه در مواردی که پهنای باند زیاد و قیمت کم جزء عوامل کلیدی باشند.

اجازه دهید با یک مثال ساده مطلب را روشن‌تر کنیم. امروزه نوارهای مغناطیسی با ظرفیت 200 GB بوفور در بازار یافت می‌شوند، و یک جعبه مقوایی به ابعاد $60 \times 60 \times 60$ cm می‌تواند 1000 تا از این نوارها را در خود جای دهد، که بدین ترتیب ظرفیت آن به 200 TB (تراپایت، معادل هزار گیگابایت) یا 1600 Tbit (که معادل 1.6 Pbit - پتابایت - است) می‌رسد. این بسته را می‌توان با استفاده از پست سریع‌السير در کمتر از ۲۴ ساعت به هر نقطه‌ای در ایالات متحده آمریکا تحویل داد. بدین ترتیب، پهنای باند مؤثر این سیستم انتقال 1600 terabits/86,400 sec یا 19 Gbps است - و اگر فاصله این دو نقطه فقط یک ساعت باشد، پهنای باند تا 400 Gbps نیز افزایش خواهد یافت. هیچ شبکه کامپیوتری حتی نمی‌تواند به این سرعت نزدیک شود.

برای شبکه‌های بانکی که در هر روز باید چندین گیگابایت اطلاعات را از نقطه‌ای به نقطه دیگر منتقل کنند (تا در صورت بروز فجایع طبیعی نیز بتوانند به کار خود ادامه دهند) هیچ چیز نمی‌تواند جای نوار مغناطیسی را بگیرد. البته شبکه‌ها هر روز سریعتر می‌شوند، ولی ظرفیت نوارهای مغناطیسی نیز رو به افزایش است.

در مورد هزینه نیز وضعیت مشابهی وجود دارد. قیمت هر نوار مغناطیسی 200 GB چیزی حدود 40 دلار است (البته در خریدهای عمده)، و از هر نوار می‌توان حداقل ۱۰ بار استفاده کرد. بنابراین، هزینه هر بار استفاده از جعبه نوار ۱۰۰۰ تا ۴۰۰۰ دلار خواهد شد. اگر 1000 دلار نیز برای حمل و نقل به آن اضافه کنیم (که البته معمولاً بسیار کمتر است)، هزینه کل به 5000 دلار برای انتقال 200 TB اطلاعات می‌رسد - و این یعنی ۳ سنت (معادل 0.01 دلار) برای هر گیگابایت، که هیچ شبکه‌ای نمی‌تواند با آن رقابت کند. نتیجه اخلاقی داستان:

هرگز پهنای باند کامیونی پُر از نوارهای مغناطیسی را که سرعت در بزرگراه در حال حرکت است، دست کم نگیرید.

۲-۲-۲ زوج تابیده

با اینکه پهنای باند نوار مغناطیسی بسیار عالیست، تأخیر انتشار آن (زمانی که طول می‌کشد تا اولین بیت اطلاعات به مقصد برسد) نامیدکننده است - در شبکه‌ها معمولاً با میلی‌ثانیه سروکار داریم نه روز و ساعت! در بسیاری از کاربردها برقرار بودن دائمی ارتباط یک امر حیاتی است. یکی از قدیمی‌ترین (و همچنان متداولترین) رسانه‌های انتقال زوج تابیده (twisted pair) است. زوج تابیده عبارتست از یک زوج سیم مسی عایق‌دار (ضخامت 1 mm)، که صورت مارپیچ به دور یکدیگر تابیده‌اند (مانند زنجیره مولکول DNA). علت تابیدن سیمها آنست که دو سیم مسی معمولی مانند یک آنتن عمل کرده، و انرژی تلف می‌کنند. در حالت تابیده امواج سیمها یکدیگر را خنثی کرده، و تشعشع به حداقل می‌رسد.

بیشترین کاربرد زوج تابیده در شبکه‌های تلفن است: تقریباً تمام تلفن‌ها با استفاده از زوج تابیده به مرکز تلفن وصل می‌شوند. از زوجهای تابیده می‌توان بطول چندین کیلومتر بدون نیاز به تقویت‌کننده استفاده کرد، ولی برای مسافت‌های طولانیتر به تکرارکننده (repeater) نیاز هست. تعداد زیادی زوج تابیده که در یک غلاف محافظ جمع شده باشند، تشکیل یک کابل زوج تابیده را می‌دهند. اگر سیمهای این کابل بصورت دو به دو به یکدیگر نتابیده باشند، تداخل شدیدی بین آنها بوجود خواهد آمد.

از زوجهای تابیده برای انتقال سیگنالهای آنالوگ و دیجیتال می‌توان استفاده کرد. پهنای باند این خطوط به ضخامت سیمها و مسافت بستگی دارد، و در فواصل کوتاه (دو تا سه کیلومتر) می‌توان به پهنای باند چندین مگابیت بر ثانیه دست یافت. بدلیل کارایی کافی و هزینه پائین، از زوج تابیده بنحو گسترده‌ای استفاده شده است، و بنظر می‌رسد تا سالها نیز این وضعیت ادامه یابد.



(الف)

(ب)

شکل ۲-۳. (الف) زوج تاییده Cat 3. (ب) زوج تاییده Cat 5.

انواع مختلفی از کابلهای زوج تاییده وجود دارد، که دو تا از آنها در شبکه‌های کامپیوتری از اهمیت بیشتری برخوردار هستند. در کابلهای Category 3 (که به Cat 3 نیز معروف است) سیمها با شدت کمتری به هم تاییده‌اند. کابل Cat 3 از چهار زوج تاییده، که در یک غلاف پلاستیکی قرار می‌گیرند، تشکیل می‌شود. تا قبل از سال ۱۹۸۸ تقریباً در همه جا از این نوع کابل استفاده می‌شد. کابل Cat 3 می‌توانست چهار خط تلفن معمولی (یا دو خط تلفن مرکب) را از ایستگاه تلفن به نقطه مورد نظر برساند.

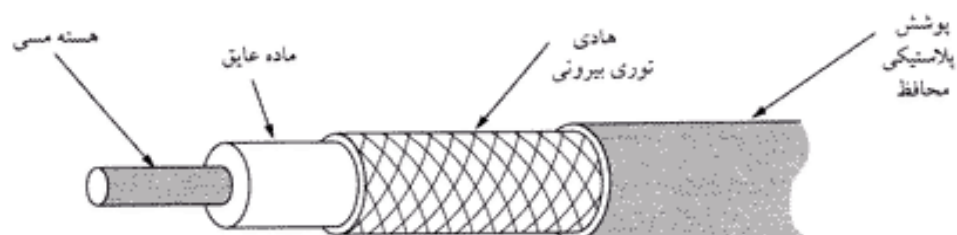
در سال ۱۹۸۸ کابل پیشرفته‌تر Category 5 (یا همان Cat 5) وارد بازار شد. این کابل شبیه Cat 3 است، با این تفاوت که تعداد دورهای آن در واحد طول بیشتر بوده، و به همین دلیل تداخل سیگنال در آن کاهش یافته و برای شبکه‌های پرسرعت مناسبتر است. کابلهای Cat 6 و Cat 7 نیز در حال آمدن به بازار هستند، که سرعت آنها پرتیب به 250 MHz و 600 MHz می‌رسد (برای مقایسه، سرعت کابلهای Cat 3 و Cat 5 پرتیب 16 MHz و 100 MHz است).

تمام این کابلها در بازار بنام UTP (زوج تاییده بدون زره - Unshielded Twisted Pair) شناخته می‌شوند، تا از کابلهای STP (زوج تاییده زره‌دار - Shielded Twisted Pair) که IBM در اوایل دهه ۱۹۸۰ معرفی کرد (و بدلیل گرانی و سختی کار با آن، استقبال چندانی از این نوع کابل نشد)، متمایز باشند. در شکل ۲-۳ دو نوع زوج تاییده Cat 3 و Cat 5 (و تفاوت آنها) را ملاحظه می‌کنید.

۳-۲-۲ کابل کواکسیال

یکی دیگر از رسانه‌های رایج کابل کواکسیال (هم محور - coaxial) است (که اغلب به آن کابل کواکس می‌گویند). این کابل بدلیل داشتن زره (غلاف فلزی) کارایی بهتری نسبت به زوج تاییده (هم از نظر سرعت، هم از نظر مسافت) دارد. کابل کواکسیال دو نوع دارد. اولی کابل 50-ohm است، که از ابتدا برای مخابرات دیجیتال در نظر گرفته شده؛ و دومی کابل 75-ohm، که ابتدا برای مخابرات آنالوگ و تلویزیون کابلی بکار می‌رفت، ولی امروزه با گسترش اینترنت کابلی از اهمیت روزافزونی برخوردار شده است. این تمایز بیش از آن که فنی باشد، جنبه تاریخی دارد: امپدانس آنتنهای دوقطبی قدیمی 300-ohm بود، و ترانسفورماتورهای تطبیق امپدانس 4:1 بوفور در بازار یافت می‌شد.

کابل کواکسیال تشکیل می‌شود از یک سیم مسی سخت بعنوان هسته (core)، یک لایه عایق استوانه‌ای به دور این هسته، یک لایه توری فلزی که به دور عایق بافته شده، و لایه پلاستیکی محافظ خارجی (شکل ۲-۴).



شکل ۲-۴. ساختمان کابل کواکسیال.

ساختمان و نحوه عایق‌بندی کابل کواکسیال باعث شده تا این نوع کابل از نظر سرعت و مصونیت در مقابل نویز کارایی بسیار خوبی داشته باشد. پهنای باند کابل‌های کواکسیال به کیفیت مواد آن، طول کابل و نسبت سیگنال به نویز امواج ارسالی بستگی دارد، و در کابل‌های جدید به 1 GHz نیز می‌رسد. از کابل کواکسیال بیشتر در سیستم‌های تلفن راه دور استفاده می‌شد، که امروزه بتدریج جای خود را به فیبرهای نوری می‌دهد. با این حال، در شبکه‌های شهری و تلویزیون کابلی هنوز هیچ رقیبی برای کابل کواکس وجود ندارد.

۴-۲-۲ فیبرنوری

بسیاری از افرادی که در صنعت کامپیوتر کار می‌کنند، از شتاب غیره‌کننده رشد تکنولوژی آن به خود می‌بالند. اولین کامپیوتر شخصی (PC) که IBM در سال ۱۹۸۱ به بازار عرضه کرد، با سرعت ساعت 4.77 MHz کار می‌کرد؛ بعد از گذشت بیست سال، امروزه PC ها می‌توانند با سرعتی متجاوز از 2 GHz کار کنند، و این یعنی ۲۰ برابر شدن سرعت در هر ۱۰ سال - چندان هم بد نیست!

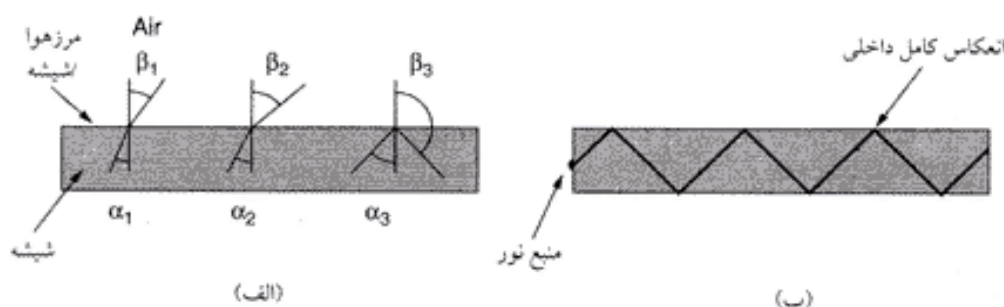
در همین دوره، سرعت مخابرات راه دور از 56-kbps (آرپانت) به 1 Gbps (مخابرات نوری جدید) رسیده است - این یعنی رشدی معادل ۱۲۵ برابر در هر دهه (در حالیکه میزان خطا هم از یک بیت در هر 10^3 بیت به تقریباً صفر رسیده).

اما پیشرفت هم مرزهای فیزیکی خاص خود را دارد؛ CPU ها مدتیست با موانعی از قبیل محدودیت سرعت نور و مشکل اتلاف گرما روبرو هستند. در حالیکه این مسائل در مخابرات وجود ندارد، و تکنولوژی فیبر نوری می‌تواند براحتی به پهنای باندی فراتر از 50,000 Gbps (یا 50 Tbps) دست یابد (و از هم اکنون بسیاری افراد با جدیت بدنبال تکنولوژی‌های بهتر نیز هستند). سرعت فعلی مخابرات فیبر نوری به چیزی در حدود 10 Gbps محدود می‌شود، ولی علت اصلی آن محدودیت سرعت تبدیل سیگنال‌های الکتریکی به پالسهای نوری است، نه محدودیت ذاتی فیبرهای نوری (البته در آزمایشگاهها به سرعت 100 Gbps نیز دست یافته‌اند).

در مسابقه بین کامپیوتر و مخابرات، مسلماً مخابرات برنده است. البته هنوز عده زیادی از دانشمندان و مهندسان کامپیوتر (که به قانون ناپکوئیست و شانون فکر می‌کنند، و نمی‌دانند که این قانونها مربوط به سیماهای مسی است) نمی‌توانند تصور پهنای باند نامحدود را به ذهن خود راه دهند. قانون آینده باید این باشد که سرعت کامپیوترها هرگز به گرد شبکه‌ها نیز نخواهد رسید، و باید از هر نوع پردازشی روی اطلاعات شبکه خودداری کرد (حتی اگر به معنای تلف شدن مقداری از پهنای باند باشد). اما اجازه دهید ببینیم فیبر نوری چگونه کار می‌کند.

یک سیستم انتقال نوری سه مولفه کلیدی دارد: منبع نور، رسانه انتقال، آشکارساز. طبق قرارداد، یک پالس نوری معادل بیت 1 و فقدان نور معادل بیت 0 است. رسانه انتقال یک رشته (فیبر) فوق‌العاده نازک شیشه است. آشکارساز (detector) نیز دستگاهیست که با برخورد نور به آن یک پالس الکتریکی تولید می‌کند. با قرار دادن یک منبع نور در یک سر فیبر نوری و یک آشکارساز در سمت دیگر، یک سیستم انتقال نوری یکطرفه خواهیم داشت که سیگنال الکتریکی را گرفته، آنرا به پالسهای نوری تبدیل کرده، و در طرف دیگر پالس دریافتی را به سیگنال‌های الکتریکی تبدیل می‌کند.

اما می‌دانیم که نور از شیشه خارج می‌شود، و چنین سیستمی بکلی بی‌فایده خواهد بود. اینجاست که یکی از قوانین جالب فیزیک نور به کمک ما می‌آید. این قانون (که به قانون شکست نور معروف است) می‌گوید که وقتی پرتو نور از یک محیط (مثلاً، شیشه) وارد محیط دیگر (مثلاً، هوا) می‌شود، در مرز این دو محیط دچار خمیدگی یا شکست (refraction) می‌شود. به شکل ۲-۵ (الف) نگاه کنید؛ در این شکل یک پرتو نور می‌بینید که با زاویه α_1 به مرز شیشه و هوا برخورد کرده، و با زاویه β_1 از شیشه خارج و وارد هوا می‌شود. مقدار خمیدگی یا شکست پرتو نور به خواص فیزیکی دو محیط (بویژه ضریب شکست آنها) بستگی دارد. اگر زاویه برخورد نور از یک مقدار



شکل ۲-۵. (الف) سه مثال از پرتوهای نوری که به مرز شیشه-هوا برخورد کرده، و شکسته می‌شوند. (ب) پرتو نور بدلیل شکست کلی در داخل شیشه گرفتار شده است.

بحرانی بیشتر باشد، پرتو نور دچار شکست کلی شده و دوباره به داخل شیشه برمی‌گردد، و هرگز وارد هوا نخواهد شد. نوری که با این زاویه (یا بیشتر از آن) به داخل شیشه تابانده شود، برای همیشه در آن محبوس می‌شود، و می‌تواند مسافت‌های طولانی را بدون اتلاف انرژی در فیبر نوری بپیماید - شکل ۲-۵ (ب) را ببینید. در شکل ۲-۵ (ب) فقط یک پرتو نور نشان داده شده است، ولی از آنجائیکه هر پرتو نوری که با زاویه بالاتر از زاویه بحرانی به مرز شیشه و هوا برخورد کند به داخل شیشه برمی‌گردد، در هر لحظه پرتوهای متعددی با زاویه‌های مختلف در داخل فیبر نوری به بالا و پائین حرکت می‌کنند، و اصطلاحاً گفته می‌شود که هر یک از این پرتوها دارای حالت (mode) خاص خود است. به فیبری که چنین ویژگی داشته باشد، فیبر چندحالتی (multimode fiber) گفته می‌شود.

ولی اگر قطر فیبر فقط چند برابر طول موج نور باشد، پرتو نور فقط می‌تواند در جهت مستقیم (بدون جهش به بالا و پائین) حرکت کند. به چنین فیبری که بصورت یک موج‌بر (wave guide) عمل می‌کند، فیبر تک‌حالتی (single-mode fiber) گفته می‌شود. فیبرهای تک‌حالتی گرانتر از فیبرهای چندحالتی هستند، ولی در مسافت‌های طولانی اغلب از این نوع فیبرها استفاده می‌شود. فیبرهای تک‌حالتی امروزی می‌توانند با ظرفیتهایی تا 50 Gbps و بطول 100 کیلومتر بدون نیاز به تقویت‌کننده کار کنند. (در آزمایشگاهها برای مسافت‌های کوتاهتر به پهنای باند بالاتری نیز دست یافته‌اند.)

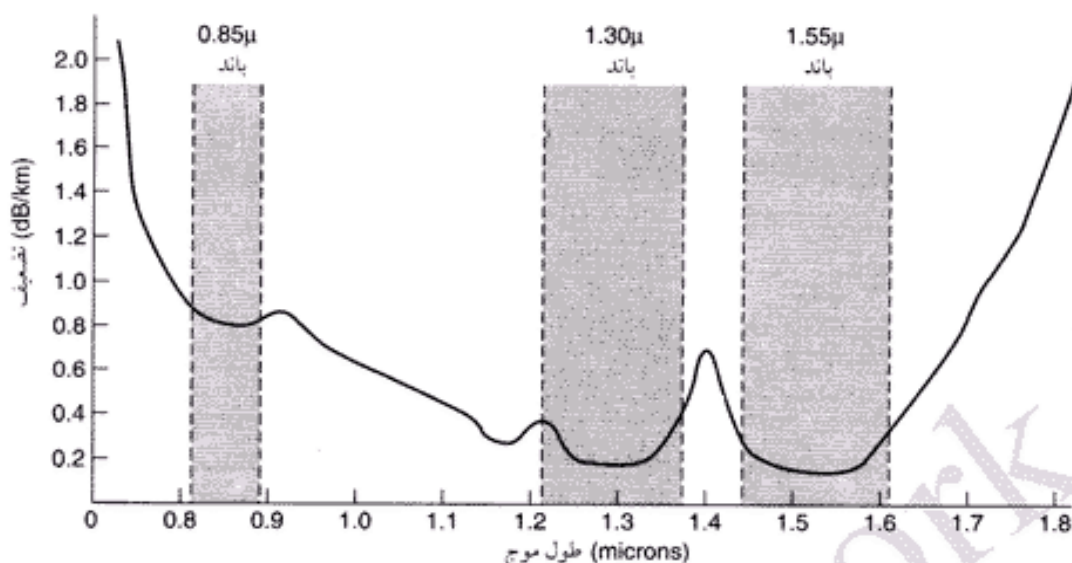
عبور نور در دوران فیبر نوری

فیبرهای نوری از شیشه ساخته می‌شوند و خود شیشه هم از شن، ماده‌ای که بوفور در طبیعت یافت می‌شود. شیشه برای اولین بار در مصر باستان ساخته شد، ولی این شیشه‌ها بقدری کدر بودند که در ضخامت‌های بیشتر از 1 mm نور را از خود عبور نمی‌دادند. شیشه‌ای که برای استفاده در پنجره‌ها مناسب باشد، در دوره رنسانس ساخته شد. شیشه‌ای که در ساخت فیبرهای نوری بکار می‌رود آنقدر شفاف است که اگر اقیانوس را با آن پُر کنیم، کف آن بوضوح دیده خواهد شد.

تضعیف نور (attenuation) در شیشه به طول موج آن و برخی از خواص فیزیکی شیشه بستگی دارد. در شکل ۲-۶ میزان تضعیف نور در فیبرهای نوری بر حسب دسی‌بل بر کیلومتر (خطی) نشان داده شده است. معادله تضعیف نور چنین است:

$$\text{attenuation (dB)} = 10 \log_{10} \frac{\text{transmitted power}}{\text{received power}}$$

برای مثال، اگر قدرت دریافتی در خروجی نصف قدرت ورودی باشد، میزان تضعیف سیگنال 3 dB است. در شکل ۲-۶ نمودار تضعیف نور را در ناحیه مادون قرمز طیف، که در عمل نیز از آن



شکل ۲-۶. تضعیف نور عبوری از فیبر نوری در ناحیه مادون قرمز.

استفاده می شود، می بینید. طول موج نور مرئی (0.4-0.7 میکرون، یا 400-700 نانومتر) قدری کوتاهتر از نور مادون قرمز است ($1 \mu\text{m} = 10^{-6} \text{m}$ و $1 \text{nm} = 10^{-9} \text{m}$).

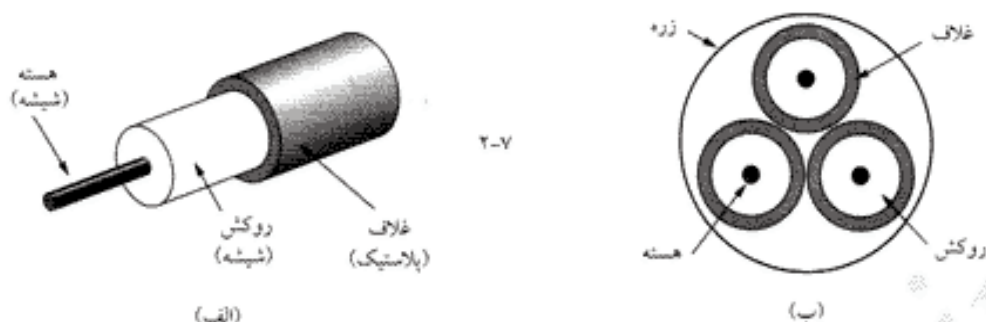
برای مخابرات نوری سه باند از طول موجهای نور مادون قرمز مورد استفاده قرار می گیرد، که بترتیب حول طول موجهای 0.85، 1.30 و 1.55 میکرون متمرکز شده اند. دو باند آخر دارای مشخصات تضعیف خوبی (کمتر از ۵ درصد در هر کیلومتر) هستند. باند 0.85 میکرون میزان تضعیف بالاتری دارد، ولی در این طول موج می توان تجهیزات نوری و الکترونیکی را از یک ماده نیمه هادی واحد (گالیوم آرسنید) ساخت. پهنای این باندها همگی بین 25,000 GHz تا 30,000 GHz است.

وقتی پالسهای نور از فیبر نوری عبور می کنند، پهنای آنها زیاد می شود، که این پدیده به پراکنش کروماتیک (chromatic dispersion) معروف است، و مقدار آن به طول موج نور بستگی دارد. یکی از راههای جلوگیری از تداخل این پالسهای پهن شده، افزایش فاصله آنهاست، و این کار نیز فقط با کاهش نرخ ارسال سیگنال ممکن است. خوشبختانه با کشف این موضوع که می توان با تولید پالسهایی که تقارن کسینوس هذلولی دارند، اثر پراکنش کروماتیک را بکلی حذف کرد، امکان آن بوجود آمده که پالسهای نوری را بدون تغییر شکل محسوس تا مسافتیهای هزاران کیلومتری مخابره کرد. امروزه تلاشهای زیادی در دست انجام است تا این پالسها را، که به آنها سولیتون (soliton) گفته می شود، از حالت آزمایشگاهی خارج و در عمل بکار بگیرند.

کابل فیبر نوری

کابل فیبر نوری بسیار شبیه کابل کواکسیال است، با این تفاوت که غلاف توری فلزی بیرونی را ندارد (به شکل ۲-۷ الف) نگاه کنید). رشته شیشه ای که نور از آن عبور می کند، در مرکز کابل قرار دارد. در فیبرهای چندحالتی قطر این رشته معمولاً $50 \mu\text{m}$ (تقریباً معادل ضخامت موی سر انسان) است، و در فیبرهای تک حالت قطر آن به 8 تا 10 میکرون می رسد.

هسته فیبر نوری با یک روکش شیشه ای (با ضریب شکست کمتر) پوشانده می شود، تا تمام نور در رشته مرکزی باقی بماند. پوشش بیرونی نیز (که اغلب پلاستیکی است) نقش محافظ را بازی می کند. معمولاً چند کابل تک رشته را در یک غلاف گرد هم می آورند (شکل ۲-۷ ب) را ببینید).



شکل ۲-۷. (الف) نمای کناری یک فیبر منفرد. (ب) سطح مقطع کابلی با سه فیبر.

کابلهای زمینی معمولاً در عمق یک متری سطح زمین دفن می‌شوند (که در آنجا از آسیب بیلهای مکانیکی یا موشهای جونده در امان نیستند). در قسمتهای کم عمق ساحل، کابلهای زیر دریایی را در کانالهای مخصوص پنهان می‌کنند، ولی در آبهای عمیق (که کندن کانال عملی نیست) آنها را آزاد در کف دریا رها می‌کنند (و با این کار آنها را در معرض آسیب از طرف کشتی‌های ماهیگیری و اسکونیدهای غول پیکر قرار می‌دهند).

سه روش برای متصل کردن فیبرهای نوری وجود دارد. در روش اول، به انتهای کابل پایانه‌های مخصوص وصل کرده، و آنها را به سوکت فیبر نوری متصل می‌کنند. این پایانه‌ها ۱۰ تا ۲۰ درصد نور را تلف می‌کنند، ولی کار با آنها بسیار ساده است.

دوم اینکه، می‌توان رشته‌ها را بطور مکانیکی بهم متصل کرد. در این روش دو سر رشته‌ها را که با دقت بریده شده‌اند، در یک غلاف روبروی هم قرار می‌دهند، و آنها را در جای خود محکم می‌کنند. در این روش با عبور دادن نور و تنظیم رشته‌ها می‌توان به حداکثر سیگنال عبوری دست یافت. اتلاف نور در این روش فقط ۱۰ درصد است، و یک فرد تعلیم دیده می‌تواند در عرض ۵ دقیقه چنین اتصالی را بوجود آورد.

در روش سوم، دو سر رشته‌ها ذوب و در هم فرو برده می‌شود، تا یک اتصال یکپارچه بوجود آید. این بهترین نوع اتصال است، چون رشته‌ها در واقع یکی می‌شوند، ولی حتی در این روش هم مقداری افت توان وجود دارد. در هر سه روش فوق، محل اتصال می‌تواند مقدارن از نور را بازتابش کند، که این نور با سیگنال اصلی تداخل خواهد کرد.

تبدیل سیگنالهای الکتریکی به پالسهای نوری معمولاً به دو روش صورت می‌گیرد: لیزرهای نیمه‌هادی و LED (Light Emitting Device). هر کدام از این منابع نوری ویژگیهای خاص خود را دارند (به شکل ۲-۸ نگاه کنید). طول موج یک منبع نور را می‌توان با قرار دادن تداخل‌سنج (interferometer) هایی از نوع فابری-پروت (Fabry-Perot) یا ماخ-زئندر (Mach-Zehnder) بین منبع نور و فیبر نوری تنظیم کرد. تداخل‌سنج فابری-پروت یک حفره تشدید (resonant cavity) ساده است، که از دو آینه موازی تشکیل می‌شود. نور بصورت عمود به این آینه‌ها تابانده می‌شود، و فقط طول موجهایی که مضرب صحیحی از طول حفره باشند، می‌توانند از آن خارج شوند. در تداخل‌سنج ماخ-زئندر نور به دو پرتو جداگانه تقسیم می‌شود، که پس از طی مسافتی کوتاه دوباره با هم ترکیب می‌شوند؛ این دو پرتو فقط در طول موجهای خاصی با یکدیگر همگام هستند. در انتهای دیگر فیبر نوری یک فتودیود (photodiode) قرار می‌گیرد، که با هر پالس نوری یک سیگنال الکتریکی تولید می‌کند. زمان پاسخ این نوع دیودها معمولاً 1 nsec است ($1 \text{ nsec} = 10^{-9} \text{ sec}$)، که باعث می‌شود نرخ داده‌ها به 1 Gbps محدود شود. از آنجائیکه نویز حرارتی در اینجا هم می‌تواند باعث بروز مشکل می‌شود، پالس نوری باید آنقدر انرژی داشته باشد که بتوان آنرا از نویز تشخیص داد (در فیبرهای نوری می‌توان ضریب خطا را با افزایش انرژی پالسهای نوری به میزان دلخواه پائین آورد).

لیزر نیمه هادی	LED	آیتم
زیاد	کم	سرعت داده
چند حالت یا تک حالت	چند حالت	نوع فیبر
بلند	کوتاه	فاصله
کم	زیاد	طول عمر
قابل توجه	کم	حساسیت به دما
زیاد	کم	قیمت

شکل ۲-۸ مقایسه ای بین لیزرهای نیمه هادی و LED بعنوان منبع نور.

شبکه های فیبر نوری

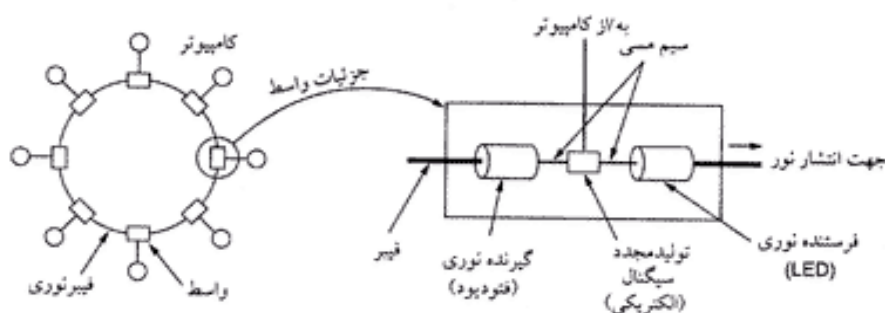
از فیبر نوری، علاوه بر مخابرات راه دور، در شبکه های محلی نیز می توان استفاده کرد - اگرچه این کار نسبت به اینترنت با مشکلات بیشتری همراه است. یک شبکه حلقوی را در نظر بگیرید (شکل ۲-۹) - این شبکه در واقع مجموعه ایست از چند اتصال نقطه-به-نقطه. هر کامپیوتر یک اتصال T (T junction) به شبکه دارد، که می تواند پالسهای نور را از خود عبور دهد، یا آنها را دریافت کند.

دو نوع اتصال وجود دارد: غیرفعال (passive) و فعال (active). در اتصال غیرفعال دو تویی به فیبر اصلی متصل می شود، که یکی LED یا دیود لیزری دارد (برای ارسال)، و دیگری فتودیود (برای دریافت). خود تویی هیچ عنصر فعالی ندارد و به همین دلیل فوق العاده قابل اعتماد است، چون خراب شدن LED یا فتودیود باعث قطعی شبکه نخواهد شد و فقط همان یک کامپیوتر از شبکه قطع می شود.

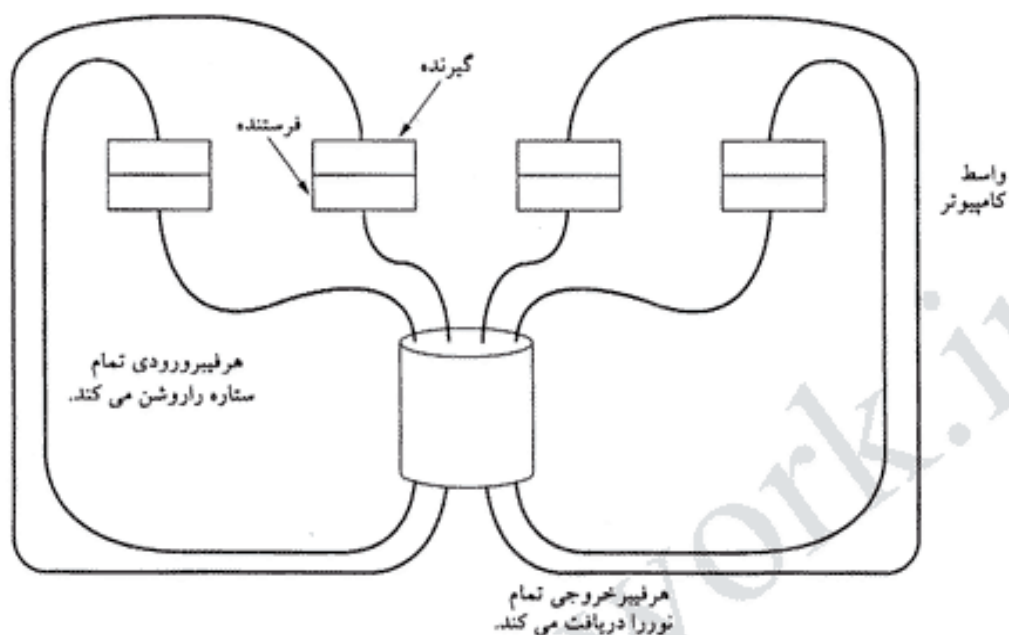
اتصال دیگری که در شکل ۲-۹ می بینید، تکرارکننده فعال (active repeater) است. در اینجا، نور ورودی ابتدا به سیگنال الکتریکی تبدیل شده، در صورت نیاز تقویت و دوباره بصورت نور منتشر می شود. اتصال به کامپیوتر توسط یک سیم مسی ساده که از تقویت کننده سیگنال منشعب شده، برقرار می شود. امروزه تقویت کننده های تمام نوری (که مرحله تبدیل به سیگنال الکتریکی - و بالعکس - در آنها حذف شده) نیز به بازار آمده اند؛ پهنای باند این تکرارکننده ها بسیار بالاست.

اگر یک تکرارکننده فعال خراب شود، حلقه شکسته شده و کل شبکه از کار می افتد. از طرف دیگر به علت تقویت سیگنال، فاصله کامپیوترها در این شبکه می تواند به چندین کیلومتر برسد، و از نظر تعداد کامپیوترها هم هیچ محدودیتی وجود ندارد. (در نوع غیرفعال، هر اتصال مقداری از توان نوری را هدر می دهد، بنابراین تعداد کامپیوترها و طول کل حلقه بشدت محدود است.)

توپولوژی حلقه تنها راه برای ایجاد شبکه های محلی با فیبر نوری نیست، و می توان به کمک سخت افزارهای خاص (انتشاردهنده های نوری) ساختاری موسوم به ستاره غیرفعال (passive star) بوجود آورد (شکل ۲-۱۰).



شکل ۲-۹. یک شبکه حلقوی فیبر نوری، با تکرارکننده های فعال.



شکل ۲-۱۰. شبکه‌ای با ساختار ستاره غیرفعال.

انتشاردهنده نوری یک استوانه شیشه‌ای خاص است، که فیبرهای ورودی (گیرنده) به یک طرف و فیبرهای خروجی (فرستنده) به طرف دیگر آن جوش خورده‌اند. وقتی یک کامپیوتر سیگنالی می‌فرستد، پالس نوری حاصله از طریق این استوانه در تمام فیبرهای متصل به گیرنده‌ها پخش می‌شود. در واقع، ستاره غیرفعال پالسهای ورودی را ترکیب کرده، و نور حاصله را روی تمام فیبرها می‌فرستد. از آنجائیکه انرژی نوری دریافتی در ستاره غیرفعال بین تمام رشته‌ها پخش می‌شود، تعداد کامپیوترهای متصل به این شبکه محدود (و وابسته به حساسیت فتودیودها) است.

مقایسه فیبر نوری و سیم مسی

در اینجا بد نیست مقایسه‌ای بین فیبر نوری و سیم مسی داشته باشیم؛ حتماً آموزنده خواهد بود. فیبر نوری مزایای متعددی نسبت به سیم مسی دارد. از همه مهمتر اینکه پهنای باند آن بسیار بیشتر از سیم مسی است، و همین دلیل کفایت تا در شبکه‌های پرسرعت انتخاب اول باشد. دیگر اینکه بدلیل اتلاف قدرت ناچیز فقط در هر ۵۰ کیلومتر به تکرارکننده نیاز دارد، در حالیکه این مسافت برای سیم مسی ۵ کیلومتر بیشتر نیست، که این باعث صرفه‌جویی زیادی در هزینه‌ها خواهد شد. مزیت دیگر فیبر نوری عدم تأثیرپذیری آن نسبت به نویز و تداخل‌های الکترومغناطیسی است. مقاومت عالی شیشه در مقابل خوردگی‌های شیمیایی نیز یکی دیگر از نقاط قوت فیبر نوری محسوب می‌شود.

جالب است بدانید که علت علاقه شرکت‌های تلفن به فیبر نوری اساساً چیز دیگریست: نازکی و سبکی. کانالهای زیرزمینی که شرکت‌های تلفن به نقاط مختلف حفر کرده‌اند، اکنون تا حد اشباع پر شده‌اند، و دیگر جای خالی برای انشعابات جدید ندارند. جایگزین کردن کابلهای مسی با فیبر نوری، علاوه بر خالی کردن این کانالها، درآمد جدیدی نیز برای این شرکتها به‌مراه دارد: فروش مس خالص به پالایشگاههای مس. علاوه بر آن، فیبرهای نوری بسیار سبکتر از سیمهای مسی هستند: هر کیلومتر از کابل مسی با هزار زوج ناپییده نزدیک به ۸۰۰۰ کیلوگرم وزن

دارد، در حالیکه وزن کابل فیبر نوری با همان ظرفیت فقط ۱۰۰ کیلوگرم است. کاهش وزن نیز معادل است با کاهش نیاز به وسایل مکانیکی سنگین برای نصب و پشتیبانی سیستم، بگونه ای که امروزه دیگر در مسیرهای جدید فقط از فیبر نوری استفاده می شود.

آخر اینکه، نور از فیبرهای نوری نشت نمی کند، و گرفتن اشعاب غیرمجاز از آن بسیار مشکل است. این ویژگی باعث شده تا فیبر نوری از ایمنی بسیار بالایی در مقابل سارقان اطلاعات برخوردار باشد.

البته فیبر نوری چندان بی عیب و نقص هم نیست. اول اینکه، این تکنولوژی هنوز بسیار جدید است و حتی بسیاری از مهندسين نیز توانایی کار با آن را ندارند. دیگر اینکه فیبر نوری بسیار آسیب پذیرتر از سیم مسی است، و حتی خم کردن بیش از حد باعث خرابی آن می شود. از آنجائیکه انتقال نوری ذاتاً یک طرفه است، برای ارتباط دوطرفه باید از دو رشته فیبر (یا یک رشته فیبر با دو باند فرکانسی) استفاده کرد. دست آخر اینکه، تجهیزات ارتباطی نوری گرانتر از انواع الکتریکی آن است. با این وجود، آینده مخابرات در فواصلی حتی بیش از چند متر از آن فیبرهای نوری است. برای بحثی جامع درباره جنبه های مختلف فیبر نوری و شبکه های آن، به (Hecht, 2001) مراجعه کنید.

۳-۲ انتقال بیسیم

یکی از پدیده های عصر ما معتادان اینترنتی است: کسانی که می خواهد بیست و چهار ساعته بر خط (on-line) باشند. برای این قبیل افراد (که دائماً در حال جابجا شدن هستند) دیگر زوج ناپایده، کابل کوکس و فیبر نوری کاربرد ندارد. آنها می خواهند بدون مقید شدن به هیچیک از سیستمهای مخابراتی دست و پا گیر، اطلاعات مورد نیازشان را روی کامپیوتر کیفی، جیبی و حتی مچی (!) دریافت کنند؛ و مخابرات بیسیم تنها چیزیست که می تواند توقعات اینها را برآورده کند. در این قسمت نگاهی کلی به مخابرات بیسیم خواهیم انداخت، نه فقط بخاطر اینکه عده ای دوست دارند کنار دریا هم از دوستان اینترنتی شان جدا نشوند، بلکه به دلیل آنکه این تکنولوژی نقش مهمی در زندگی امروزی ما بازی می کند.

برخی افراد معتقدند که در آینده فقط دو نوع مخابرات وجود خواهد داشت: فیبر نوری و بیسیم. تمام تجهیزات ثابت (کامپیوترهای رومیزی، تلفن و فکس) از فیبر نوری، و تجهیزات متحرک از بیسیم بهره خواهند گرفت. حتی در مواردی می توان از بیسیم برای تجهیزات ثابت استفاده کرد. مثلاً، اگر کابل کشی به یک ساختمان (بعلت وجود موانع طبیعی مانند کوه، جنگل و باتلاق) مشکل باشد، استفاده از بیسیم ترجیح دارد. همانطور که قبلاً هم گفتیم، مخابرات بیسیم دیجیتال برای اولین بار در مجمع الجزایر هاوایی بکار گرفته شد.

۱-۳-۲ طیف الکترومغناطیس

امواج الکترومغناطیس حاصل حرکت الکترونها هستند، که می توانند در فضا (حتی در خلأ) منتشر شوند. وجود چنین امواجی اولین بار بصورت تئوری در سال ۱۸۶۵ توسط فیزیکدان انگلیسی جیمز کلرک ماکسول پیش بینی شد، و در سال ۱۸۸۷ فیزیکدان آلمانی هاینریش هرتز موفق شد آنها را مشاهده کند. تعداد نوسانهای یک موج در ثانیه فرکانس، f ، نامیده می شود، و واحد آن (به افتخار هاینریش هرتز) هرتز (Hz) نامگذاری شده است. فاصله بین دو قله (یا حضیض) متوالی موج را نیز طول موج می گویند، و آنرا با حرف یونانی λ نشان می دهند.

اگر آنتنی با اندازه مناسب به یک مدار الکتریکی وصل شود، می تواند امواج الکترومغناطیسی منتشر کند، که این امواج را می توان در فواصل مناسب دریافت کرد. تمام سیستمهای مخابرات بیسیم بر این اساس کار می کنند.

امواج الکترومغناطیسی (صرفنظر از فرکانس آنها) در خلأ با سرعت ثابتی (سرعت نور) حرکت می کنند، که مقدار تقریبی آن 3×10^8 m/sec است، و با c نشان داده می شود (سرعت نور سرعت حد در طبیعت است، یعنی

هیچ چیز نمی تواند سریعتر از آن حرکت کند. سرعت این امواج در سیم مسی یا فیبر نوری به حدود $\frac{2}{3}c$ کاهش می یابد، که تا حدی نیز به فرکانس وابسته است.

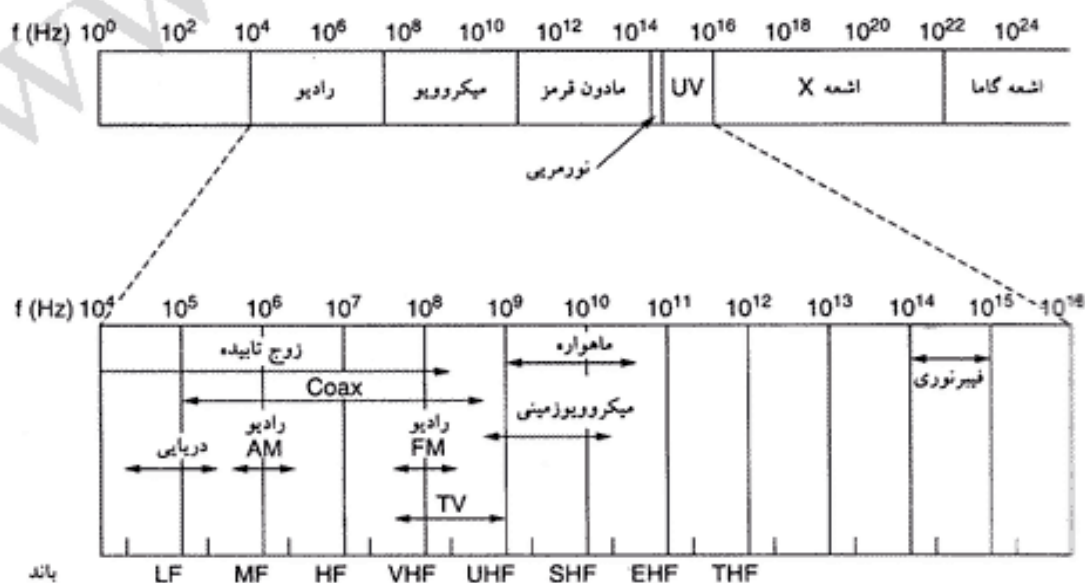
رابطه اساسی بین f ، λ و c (در خلأ) چنین است:

$$\lambda f = c \quad (2-2)$$

از آنجائیکه c ثابت است، با داشتن f می توان λ را محاسبه کرد، و بالعکس. اگر λ را برحسب متر و f را برحسب MHz داشته باشیم، رابطه بالا بصورت $\lambda f \approx 300$ در می آید. برای مثال، در فرکانس 100-MHz طول موج 3 m است، که در فرکانس 1000-MHz طول موج آن به 0.3 m کاهش می یابد؛ موجی با طول موج 0.1 m فرکانسی معادل 3000-MHz (در خلأ) دارد.

در شکل ۱۱-۲ طیف الکترومغناطیس را ملاحظه می کنید. از ناحیه های رادیویی، مایکروویو، مادون قرمز، و مرئی این طیف می توان (با مدولاسیون دامنه، فرکانس، و فاز) برای انتقال اطلاعات استفاده کرد. نور ماوراء بنفش، اشعه X و اشعه گاما بدلیل فرکانس بالاتر برای منظور بهتر هستند، ولی از آنجائیکه تولید و مدولاسیون آنها مشکل است، از اجسام غیرشفاف عبور نمی کنند، و برای سلامتی موجودات زنده خطرناکند، کاربرد مخابراتی ندارند. باندهایی که در پائین شکل ۱۱-۲ می بینید، رسماً توسط ITU (برحسب طول موج) تقسیم بندی و نامگذاری شده اند. اصطلاحات LF، MF و HF بترتیب معادل فرکانس پائین، فرکانس متوسط و فرکانس بالا هستند. پیداست وقتی این نامگذاریها انجام می شد، کسی تصور فرکانسهای بالاتر از 10 MHz را هم نمی کرد، و بهمین دلیل امروزه برای باندهای بالاتر از آن از نامهای VHF (فرکانس خیلی بالا)، UHF (فرکانس بسیار بالا)، SHF (فرکانس فوق العاده بالا)، EHF (فرکانس شدیداً بالا)، و THF (فرکانس خارق العاده بالا) استفاده می شود - خوشبختانه برای فرکانسهای بالاتر هنوز نامی انتخاب نشده، و شما هم می توانید ذوق خود را در این زمینه امتحان کنید.

مقدار اطلاعاتی که یک موج الکترومغناطیس می تواند حمل کند، به پهنای باند آن بستگی دارد. با تکنولوژی امروزی، می توان در فرکانسهای پائین بازای هر هرتز از پهنای باند ۲ تا ۳ بیت، و در فرکانسهای بالا تا ۸ بیت



شکل ۱۱-۲. طیف الکترومغناطیس و کاربردهای مخابراتی آن.

اطلاعات منتقل کرد، بنابراین یک کابل کوآکسیال با پهنای باند 750 MHz می‌تواند در هر ثانیه تا چندین گیگابیت اطلاعات منتقل کند. با یک نگاه به شکل ۲-۱۱ می‌توان براحتی دریافت که چرا فیبر نوری این همه طرفدار دارد.

اگر معادله (2-2) را برای f حل کرده و سپس از آن نسبت به λ دیفرانسیل بگیریم، خواهیم داشت

$$\frac{df}{d\lambda} = -\frac{c}{\lambda^2}$$

حال اگر بجای دیفرانسیل تفاضل محدود و قدر مطلق مقادیر را در نظر بگیریم، داریم

$$\Delta f = \frac{c\Delta\lambda}{\lambda^2}$$

بنابراین با داشتن پهنای یک باند، $\Delta\lambda$ ، می‌توان فرکانس متناظر با آن باند، Δf ، و از آنجا نرخ داده آن، را محاسبه کرد. هر چه پهنای یک باند بزرگتر باشد، نرخ انتقال داده آن باند بیشتر است. برای مثال، باند $1.30 \mu\text{m}$ را در شکل ۲-۶ در نظر بگیرید. در اینجا، $\lambda = 1.3 \times 10^{-6} \text{ m}$ و $\Delta\lambda = 0.17 \times 10^{-6} \text{ m}$ است، بنابراین Δf تقریباً 30 THz خواهد شد. با چنین پهنای باندی (و با فرض 8 bits/Hz) می‌توان به نرخ انتقال داده 240 Tbps دست یافت. در اکثر سیستمهای انتقال از باندهای فرکانس باریک استفاده می‌شود (یعنی، $\frac{\Delta f}{f} \ll 1$)، تا گیرنده بتواند بیشترین توان را دریافت کند. اما گاهی نیز استفاده از باند فرکانسی وسیع ضروریست، که بر دو نوع است. در نوع طیف گسترده با پرش فرکانسی (frequency hopping spread spectrum)، فرستنده در هر ثانیه صدها بار فرکانس خود را عوض می‌کند (بعبارت دیگر از فرکانسی به فرکانس دیگر می‌پرد). این تکنیک بویژه در مخابرات نظامی کاربرد دارد، چون تشخیص فرکانس فرستنده بسیار مشکل است و پارازیت انداختن روی آن نیز تقریباً غیرممکن است. در این روش تضعیف موج در اثر انعکاس نیز وجود ندارد، چون وقتی موج انعکاسی (کمی بعد از موج اصلی) به گیرنده می‌رسد، فرکانس آن تغییر کرده است و فرکانسهای قبلی را دیگر قبول نمی‌کند. در سالهای اخیر این تکنیک در محصولات تجاری هم کاربرد پیدا کرده است - برای مثال، بلوتوث و 802.11 هر دو از این تکنیک استفاده می‌کنند.

اختراع این تکنیک نیز داستانی جالب دارد، که شنیدن آن خالی از لطف نیست. یکی از دو مخترع تکنیک طیف گسترده با پرش فرکانسی، خانم هدی لامار هنرپیشه اتریشی‌الاصیل است. شوهر اول این خانم، که سازنده تسلیحات نظامی بود، برای وی توضیح داد که چگونه می‌توان با استفاده از سیگنالهای رادیویی از درها و موشک‌ها را کنترل کرد. وقتی خانم لامار فهمید که شوهرش به هیتلر اسلحه می‌فروشد، دچار وحشت شد، با لباس مبدل گریخت، و به هالیوود رفت تا به حرفه موردعلاقه‌اش یعنی هنرپیشگی ادامه دهد. لامار به کمک دوستش جورج آنتیل (که یک آهنگساز بود) و برای کمک به متفقین، تکنیک پرش فرکانسی را اختراع کرد. طرح اولیه آنها دارای ۸۸ فرکانس (به تعداد کلیدهای پیانو) بود، که به شماره 2,292,387 در اداره ثبت اختراعات ایالات متحده آمریکا به ثبت رسید. با این حال، آنها نتوانستند مفید و عملی بودن این اختراع را به نیروی دریایی آمریکا بقبولانند، و هرگز پولی بابت آن دریافت نکردند. تنها سالها پس از سپری شدن از حق‌الاختراع آن بود که این تکنیک مورد توجه محافل علمی قرار گرفت.

نوع دوم باندهای فرکانسی وسیع، طیف گسترده با توالی مستقیم (direct sequence spread spectrum) نام دارد، که در آن سیگنال روی طیف وسیعی از فرکانسها پخش می‌شود. این تکنیک نیز امروزه، بویژه در تلفنهای همراه نسل دوم، کاربردهای تجاری پیدا کرده است، و با آمدن نسل سوم تلفنهای همراه تسلط آن بر بازار کامل خواهد شد، زیرا دارای کارایی طیفی خوب، مصنویت در برابر نویز عالی و بسیاری ویژگیهای دیگر است. در برخی از شبکه‌های محلی بیسیم نیز از تکنیک طیف گسترده با توالی مستقیم استفاده شده است. در قسمتهای آینده همین

فصل باز هم به مبحث طیف گسترده برمی گردیم؛ اگر به تکنیکهای مخابرات طیف گسترده و تاریخچه آن علاقمند هستید، کتاب (Scholtz, 1982) را جالب خواهید یافت.

فعلاً فرض را بر این می گذاریم که همه جا از باند فرکانسی باریک استفاده می شود، و بحث خود را با کاربرد بخشهای مختلف طیف الکترومغناطیس شکل ۲-۱۱ ادامه می دهیم - از رادیو شروع می کنیم.

۲-۳-۲ مخابرات رادیویی

امواج رادیویی کاربرد گسترده ای در مخابرات (در فضاهای سرپوشیده یا باز) دارند، چون باسانی می توان آنها را تولید کرد، بُرد زیادی دارند، و از ساختمانها و موانع عبور می کنند. امواج رادیویی همه-طرفه هستند، یعنی در تمام جهات منتشر می شوند، بنابراین نیازی به تنظیم دقیق موقعیت گیرنده و فرستنده نسبت بیکدیگر نیست.

البته همه-طرفه بودن امواج رادیویی همیشه هم خوب نیست. در دهه ۱۹۷۰، شرکت جنرال موتورز تصمیم گرفت اتومبیلهای کادیلاک جدید خود را به ترمز ضد-قفل کامپیوتری (ABS) مجهز کند. در این سیستم وقتی راننده ترمز می گیرد، کامپیوتر مرکزی پالسهای به ترمزها می فرستد، که آنها را چندین بار در ثانیه باز و بسته می کند، و ترمزها دیگر قفل نمی کنند. مدتی بعد در یک روز آفتابی، یکی از پلیسهای گشت بزرگراه ایالت اوهایو تصمیم گرفت رادیوی بیسیم جدید خود را امتحان کرده و با مرکز فرماندهی تماس بگیرد. با این کار، اتومبیل کادیلاکی که در نزدیکی وی حرکت می کرد، بلافاصله مثل یک اسب چموش شروع به حرکات عجیبی کرد. وقتی افسر پلیس اتومبیل خاطی را متوقف کرد، راننده ادعا کرد که هیچ کاری انجام نداده و اتومبیل بیکپاره دیوانه شده است.

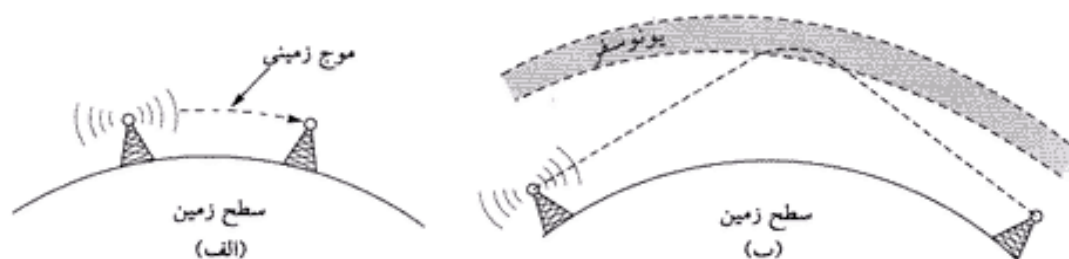
کمی بعد ماجرا روشتتر شد: اتومبیلهای کادیلاک همه جا خوب کار می کردند، و فقط در بزرگراههای ایالت اوهایو، آن هم وقتی پلیس می دیدند، دیوانه می شدند. تا مدتها جنرال موتورز سرگردان بود و نمی توانست بفهمد چرا کادیلاک در همه جا جز بزرگراههای ایالت اوهایو خوب کار می کند (کادیلاک حتی در خیابانها و جاده های معمولی اوهایو هم خوب کار می کرد). فقط بعد از تحقیقات گسترده بود که معلوم شد، مدارهای سیمکشی کادیلاک برای فرکانس سیستم رادیویی جدید پلیس بزرگراه اوهایو تبدیل به یک آنتن خوب می شود، و پالسهای این رادیو سیستم ABS را فعال می کند.

ویژگیهای امواج رادیویی به فرکانس آنها وابسته است. در فرکانسهای پائین، امواج براحتی از موانع عبور می کنند، ولی توان آنها بر حسب فاصله بسرعت افت می کند (با ضریب $1/r^2$). در فرکانسهای بالا، امواج رادیویی به خط مستقیم حرکت می کنند، و در برخورد با موانع منعکس می شوند. حتی قطرات باران هم امواج با فرکانس بالا را جذب می کند. امواج رادیویی در تمام فرکانسها در معرض تداخل ناشی از کارکرد وسایل الکتریکی (مانند موتور) هستند.

بدلیل بُرد زیاد امواج رادیویی، خطر تداخل آنها در سرتاسر دنیا وجود دارد، بهمین دلیل دولتها نحوه استفاده از فرستنده های رادیویی را (البته با یک استثنا) بشدت کنترل می کنند.

امواج رادیویی در باندهای VLF، LF و MF از انحناى زمین تبعیت می کنند - به شکل ۲-۱۲ (الف) نگاه کنید. این امواج نزدیک به ۱۰۰۰ کیلومتر بُرد دارند. ایستگاههای رادیویی AM در باند MF کار می کنند، و بهمین دلیل بُرد آنها چنین زیاد است. امواج رادیویی در این باند براحتی از موانع و ساختمانها عبور می کنند، و علت دریافت آنها در داخل ساختمانها نیز همین است. تنها مشکل این امواج پهنای باند کم آنهاست (معادله 2-3 را ببینید).

در باندهای HF و VHF، آن قسمت از امواج رادیویی که نزدیک سطح زمین حرکت می کنند، جذب زمین می شوند. ولی بخشی از امواج که به سمت فضا می روند، پس از برخورد با یونوسفر (ionosphere - لایه ای از جو زمین به ارتفاع ۱۰۰ تا ۵۰۰ کیلومتر، که حاوی ذرات باردار است) به طرف زمین برمی گردند؛ شکل ۲-۱۲ (ب) را



شکل ۲-۱۲. (الف) در باندهای VLF، LF و MF امواج رادیویی از انحناى زمین تبعیت می کنند. (ب) در باندهای HF و VHF آنها بین زمین و یونوسفر رفت و برگشت می کنند.

نگاه کنید. در شرایط خاصی این امواج می توانند چندین بار بین زمین و یونوسفر رفت و برگشت کرده، و صدها کیلومتر دورتر دریافت شوند. طرفداران رادیو آماتور و مخابرات نظامی از باندهای HF و VHF استفاده می کنند.

۳-۳-۲ مخابرات مایکروویو

در فرکانسهای بالای 100 MHz امواج تقریباً به خط مستقیم حرکت می کنند، و می توان آنها را دقیقاً روی یک نقطه متمرکز کرد. متمرکز کردن تمام انرژی موج در یک پرتو باریک (با استفاده از آنتنهای بشقابی) باعث بالا رفتن نسبت سیگنال به نویز می شود، ولی از طرف دیگر لازم است تا گیرنده و فرستنده بدقت تنظیم شوند. علاوه بر آن، باریک بودن پرتوها باعث می شود تا تداخل فرستنده ها به حداقل برسد، مشروط بر اینکه کمی با هم فاصله داشته باشند. تا قبل از اختراع فیبر نوری، این امواج (که به مایکروویو معروفند) ستون فقرات مخابرات راه دور محسوب می شدند. در واقع، یکی از شرکت های رقیب AT&T بنام MCI شبکه گسترده ای از برج های مایکروویو با فواصل دهها کیلومتر ایجاد کرده بود (حتی نام این شرکت - Microwave Communications, Inc. - هم بنوعی گویای استفاده از امواج مایکروویو است). MCI مدتهاست به فیبر نوری روی آورده، و اکنون در WorldCom ادغام شده است.

از آنجائیکه امواج مایکروویو به خط مستقیم حرکت می کنند، اگر فاصله ایستگاه مبدأ و مقصد زیاد باشد، انحناى زمین مانع از رسیدن امواج خواهد شد. به همین دلیل لازم است تا در فاصله بین آنها از برج های تکرارکننده استفاده شود. هر چه ارتفاع این برجها بیشتر باشد، فاصله آنها می تواند زیادتر باشد. فاصله دو ایستگاه تقریباً با توان دوم ارتفاع آنها رابطه دارد؛ برای مثال، دو برج ۱۰۰ متری می توانند چیزی در حدود ۸۰ کیلومتر فاصله داشته باشند. امواج مایکروویو، برخلاف امواج رادیویی فرکانس پائین، نمی توانند بخوبی از موانع عبور کنند. علاوه بر این، با وجود متمرکز شدن موج در فرستنده، امواج مایکروویو در طول مسیر دچار پراکندگی جزئی می شوند. قسمتی از موج مایکروویو نیز که توسط لایه های پائین جو منعکس می شود، فاصله بیشتری را طی کرده و هنگام رسیدن به گیرنده دیگر با موج اصلی هم فاز نیست، و باعث خنثی شدن آن می شود. این پدیده که به محوشدگی چندمسیره (multipath fading) معروف است، یکی از مشکلات جدی در مخابرات رادیویی محسوب می شود، و به وضع هوا و فرکانس موج بستگی دارد. در برخی از موارد، ۱۰ درصد ظرفیت کانال برای مقابله با این وضعیت کنار گذاشته می شود، تا در صورت بروز محوشدگی چندمسیره بتوانند موقتاً از فرکانسهای جایگزین استفاده کنند. تقاضا برای پهنای باند بیشتر باعث شده تا فرکانسها هر روز بالا و بالاتر برود. امروزه باندهای 10 GHz نیز در حال کار هستند، ولی از فرکانس 4 GHz به بالا مشکل جدیدی خودنمایی می کند: جذب شدن انرژی امواج توسط

آب. طول موج این قبیل امواج فقط چند سانتی متر است، و براحتی جذب قطرات باران می شوند. این پدیده شاید برای کسانی که می خواهند یک اجاق مایکروویو بزرگ بسازند و پرندگان را در هوا کباب کنند خوب باشد، ولی برای مخابرات مایکروویو یک مشکل جدی است. تنها راه حل این مشکل (مانند محوشدگی چندمسیره) قطع کردن کانالهای باران زده، و استفاده از کانالهای جایگزین است.

خلاصه اینکه، استفاده از امواج مایکروویو در مخابرات راه دور، تلفنهای همراه، و تلویزیون چنان گسترش یافته، که امروزه دیگر در طیف مایکروویو جایی برای استفاده بیشتر نمانده است. مخابرات مایکروویو مزایای زیادی نسبت به فیبر نوری دارد. اول اینکه نیازی به تصرف زمین برای حفر کانال و کشیدن کابل نیست، و با خرید چند تکه زمین کوچک در فواصل ۵۰ کیلومتری و نصب چند برج مایکروویو، می توان یک سیستم مخابراتی مستقل ایجاد کرد. بهمین خاطر بود که MCI توانست سرعت بعنوان یکی از غولهای صنعت تلفن راه دور قد علم کند. (یکی دیگر از شرکتهای بزرگ مخابراتی یعنی Sprint، مسیر کاملاً متفاوتی در پیش گرفت: این شرکت که متعلق به راه آهن پاسیفیک جنوبی بود، از زمینهای حاشیه خطوط آهن استفاده کرده، و کابلهای خود را در آنجا دفن کرد.)

دیگر اینکه، مایکروویو نسبتاً ارزان است. نصب چند برج ساده برای نصب آنتنها (که می تواند فقط یک دکل ساده یا چهار سیم مهارکننده باشد)، بسیار ارزاتر از خرید زمین در مناطق شلوغ شهری یا صعب العبور کوهستانی (و یا حتی اجاره خطوط تلفنی از شرکتهای تلفن) است.

طیف الکترومغناطیس و سیاست

برای جلوگیری از هرج و مرج و آشوب، توافقی در سطح ملی و بین المللی برای نحوه استفاده از فرکانسها شکل گرفته است. از آنجائیکه هر کس پهنای باند بیشتری می خواهد، دولتها مجبور به دخالت هستند، و پهنای باند لازم برای رادیوی AM و FM، تلویزیون، تلفنهای همراه، شرکتهای تلفن، پلیس، ناوبری هوایی و دریایی، مصارف نظامی و دولتی (و خلاصه، همه آنها) که فرکانس می خواهند) را به آنها اختصاص می دهند. در سطح بین المللی نیز یکی از شعبات ITU-R (بنام WARC) به هماهنگی این اقدامات اختصاص یافته، تا کارکرد وسایل مخابراتی را در کشورهای مختلف تضمین کند. با این همه، توصیه های ITU-R الزام آور نیست، و گاهی پیش می آید که FCC (Federal Communication Commission - سازمانی که مسئول تخصیص باندهای فرکانسی در ایالات متحده آمریکا است) توصیه های آنرا نادیده بگیرد، تا (مثلاً) یک طیف فرکانسی را به یک گروه قدرتمند سیاسی اختصاص دهد.

علاوه بر اختصاص قسمتی از طیف فرکانسی به یک کاربرد خاص (مثلاً، تلفن همراه)، باید مشخص شود که کدام حامل (carrier) ها مجاز به استفاده از این فرکانسها هستند. در گذشته، برای این منظور سه روش کاربرد گسترده تری داشت. در قدیمی ترین آنها، که به مسابقه زیبایی معروف بود، هر حامل بایستی توضیح می داد که چرا پیشنهاد آنها بیشترین نفع را برای جامعه دارد - و این مقامات دولتی بودند که بهترین پیشنهاد را انتخاب می کردند. از آنجائیکه انتخاب یک حامل منافع میلیاردی برای آن در بر داشت، در این روش رشوه، فساد و پارتی بازی بیداد می کرد. علاوه بر آن، حتی وقتی یک کارمند وظیفه شناس تشخیص می داد که پیشنهاد یک شرکت خارجی بهتر از شرکتهای داخلی است، قبولاندن آن به مقامات بالاتر کار ساده ای نبود.

مشکلات روش انتخاب دولتی، به روش دوم منجر شد: قرعه کشی بین حامل های متقاضی یک طیف فرکانسی. مشکل این روش آن بود که حتی شرکتهایی که هیچ نفعی در استفاده از طیف نداشتند، می توانستند در قرعه کشی شرکت کنند. بدین ترتیب حتی صاحب یک رستوران یا کفش فروشی هم می توانست برنده قرعه کشی شده، و بعد با خیال راحت و بدون هیچ خطری امتیاز خود را به شرکتهای ذینفع بفروشد و سود کلانی به جیب بزند.

بخشیدن چنین ثروتهای بادآورده ای به افراد زرنگ و هوشیار، باعث انتقادات زیادی شد، و منجر به اتخاذ روش سوم اعطای امتیاز حامل طیفهای فرکانسی گردید: حراج کردن پهنای باند. وقتی انگلستان در سال ۲۰۰۰ طیف فرکانسی تلفنهای همراه نسل سوم را به حراج گذاشت، پیش بینی می کرد که از این راه چهار میلیارد دلار عایدی داشته باشد. ولی در واقع از این مزایده چهل میلیارد دلار سود برد، چون شرکتهای زیادی (از ترس اینکه از قافله تلفن همراه عقب بمانند) به تکاپوی خرید طیف فرکانسی افتادند. این حادثه طمع سایر دولتها را تحریک کرد، و آنها هم به فکر راه انداختن مزایده فروش طیف فرکانسی افتادند. شرکتهای بسیاری به همین دلیل تا مرز ورشکستگی پیش رفتند، و آنهایی که ماندند باید سالها تلاش کنند تا پول از دست رفته را جبران کنند.

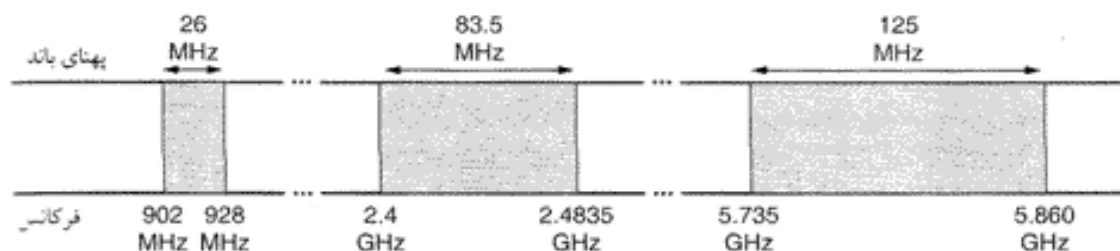
روش دیگر تخصیص فرکانس این است که اصلاً آنرا به کسی اختصاص ندهند: اجازه دهید هر کس با هر فرکانسی که می خواهد کار کند، فقط توان فرستنده های آنها را بگونه ای کنترل کنید که بُرد کمی داشته باشند، و با فرستنده های دیگر تداخل نکنند. بسیاری از دولتها بخشی از طیف فرکانسی خود را (که به باند ISM : صنعتی، علمی، پزشکی معروف است) با همین نیت آزاد می گذارند تا هر کس که مایل است از آنها استفاده کند. در باز کن های برقی، تلفنهای بیسیم خانگی، اسباب بازی های کنترل از راه دور، و بسیاری از وسایل خانگی مجهز به کنترل رادیویی از باندهای ISM استفاده می کنند. برای به حداقل رساندن تداخل بین این دستگاهها، FCC و سایر دولتها سازندگان این قبیل وسایل را مجبور می کنند تا از تکنیکهای طیف گسترده استفاده کنند.

محل طیف ISM از کشوری به کشور دیگر متفاوت است. برای مثال، در ایالات متحده آمریکا وسایلی با توان تشعشعی کمتر از ۱ وات می توانند (بدون نیاز به اخذ مجوز FCC) از باندهای مشخص شده در شکل ۲-۱۳ استفاده کنند. باند 900-MHz بهترین کارایی را دارد، ولی علاوه بر اینکه خیلی شلوغ است، در تمام دنیا هم آزاد نیست. باند 2.4-GHz در اغلب کشورهای دنیا آزاد است، ولی در معرض تداخل امواج اجاقهای مایکروویو و تأسیسات رادار قرار دارد. بلوتوث و برخی از شبکه های محلی 802.11 در این باند کار می کنند. باند 5.7-GHz نسبتاً جدید است و وسایل آن هنوز گران هستند، ولی از آنجائیکه استاندارد 802.11a در این باند کار می کند، بزودی شاهد رواج آن خواهیم بود.

۲-۳-۴ امواج مادون قرمز و میلیمتری

امواج هدایت نشده مادون قرمز و میلیمتری کاربرد زیادی در مخابرات بُرد کوتاه دارند. دستگاههای کنترل از راه دور وسایل صوتی و تصویری همگی از امواج مادون قرمز استفاده می کنند. این دستگاهها جهت دار، ارزان و ساده هستند، ولی یک عیب عمده دارند: امواج مادون قرمز از اجسام صلب عبور نمی کند (بین دستگاه کنترل از راه دور و تلویزیون خود بایستید، و ببینید هنوز کار می کند یا نه). در حالت کلی، هر چه به سمت فرکانسهای طیف مرئی نور برویم، امواج بیشتر شبیه نور (و کمتر شبیه امواج رادیویی) عمل می کنند.

از طرف دیگر، عبور نکردن امواج مادون قرمز از موانع سخت یک مزیت محسوب می شود: شما که نمی خواهید تلویزیونتان با کنترل از راه دور همسایه روشن و خاموش شود. استراق سمع امواج مادون قرمز نیز



شکل ۲-۱۳. باندهای ISM در ایالات متحده آمریکا.

مشکلتر از امواج رادیویی است، و به همین دلیل از ایمنی بالاتری برخوردار است. بر خلاف امواج رادیویی، استفاده از امواج مادون قرمز هیچ نیازی به کسب مجوز از مقامات رسمی ندارد. با وجود برخی کاربردهای امواج مادون قرمز در صنعت کامپیوتر (مانند اتصال ماوس و چاپگر)، این امواج در صحنه مخابرات حرفی برای گفتن ندارند.

۵-۳-۲ مخابرات امواج نوری

ارسال علائم نوری قرنهایست که شناخته شده و بکار برده می شود. امروزه برای متصل کردن دو شبکه که در ساختمانهای جداگانه قرار دارند، از لیزرهایی که روی پشت بام ساختمانها قرار می گیرند، استفاده می شود. پرتوهای لیزر اساساً یکطرفه هستند، بنابراین هر ساختمان باید فرستنده و گیرنده لیزری جداگانه داشته باشد. پهنای باند این روش بسیار بالا و هزینه آن نیز بسیار پایین است، و علاوه بر اینکه نصب ساده ای (نسبت به تجهیزات مایکروویو) دارد، نیازی به مجوز FCC نیز ندارد.

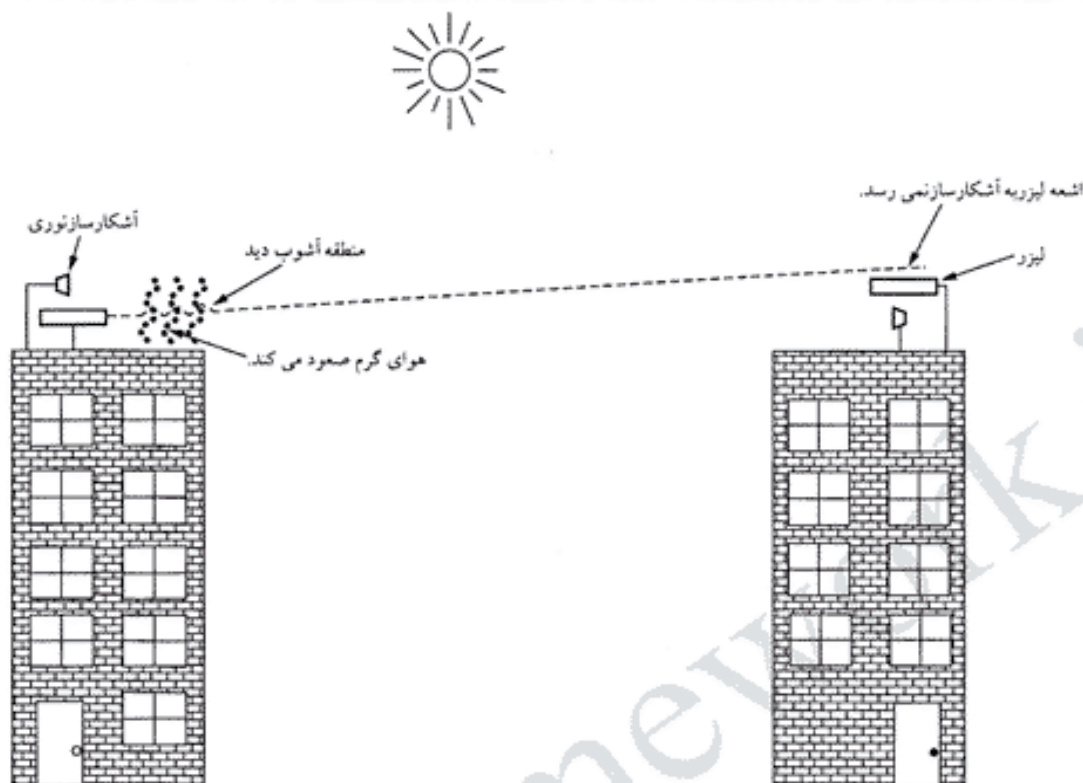
البته مزیت اصلی پرتو لیزر (یعنی باریک بودن آن) در اینجا یک نقطه ضعف محسوب می شود. هدفگیری یک پرتو لیزر به قطر ۱ میلی متر روی هدفی باندازهٔ ته سوزن در فاصلهٔ ۵۰۰ متری حتی برای قهرمان تیراندازی المپیک هم دشوار است. در این سیستمها معمولاً از عدسیهای خاصی برای پراکنده کردن نور لیزر و پهن کردن پرتوهای آن استفاده می شود.

یکی دیگر از معایب لیزر اینست که نمی تواند در روزهای بارانی و مه آلود بخوبی عمل کند، ولی برای هوای صاف و آفتابی ایده آل است. تجربه جالبی که مؤلف این کتاب از چنین سیستمی دارد، می تواند آموزنده باشد. چندی قبل در یک کنفرانس علمی که در یکی از هتلهای مدرن اروپایی ترتیب یافته بود، شرکت داشتم. مسئولین با درایت کنفرانس تعدادی کامپیوتر در یک اتاق نصب کرده بودند، تا مدعوین بتوانند در حین شرکت در سخنرانیها ایمیل خود را نیز چک کنند. از آنجائیکه شرکت مخابرات محلی حاضر نشده بود تعداد زیادی خط تلفن را فقط برای سه روز برگزاری کنفرانس در اختیار برگزارکنندگان آن قرار دهد، آنها یک دستگاه لیزر روی پشت بام هتل قرار داده بودند، تا آنجا را به ساختمان چهار تمان کامپیوتر دانشگاه که در چند کیلومتری آن محل قرار داشت، وصل کنند. سیستم شب قبل از افتتاح کنفرانس تست شده بود، و همه چیز مرتب بنظر می رسید. ساعت ۹ صبح روز بعد (که روزی صاف و آفتابی بود) ارتباط بکلی قطع شد، و تمام آن روز نیز وصل نشد. عصر همان روز، برگزارکنندگان کنفرانس دوباره سیستم را تست کردند، و این بار هم همه چیز درست و عالی بود. اما صبح روز بعد (و روز بعد از آن) دوباره همان اتفاق تکرار شد. فقط بعد از پایان کنفرانس بود که علت کشف شد.

با شروع روز، گرمای خورشید باعث گرم شدن بام ساختمان شده، و هوای گرم به بالا صعود می کرد (شکل ۱۴-۲). این جریان هوای آشفته باعث می شد تا پرتو لیزر به مقدار ناچیز منحرف شده، و تمرکز آن روی گیرنده ساختمان هدف بهم بخورد. این دقیقاً همان پدیده ایست که باعث می شود تا ستارگان در شب چشمک بزنند، و یا در روزهای داغ تابستان در جاده ها همه چیز مواج بنظر برسد.

۴ ماهواره های مخابراتی

در سالهای ۱۹۵۰ و اوایل ۱۹۶۰ برخی افراد تلاش کردند تا با استفاده از بالونهای هوای گرم که پوششی فلزی داشتند، نوعی سیستم انعکاس امواج رادیویی بسازند. متأسفانه سیگنال برگشتی چنان ضعیف بود که به هیچ دردی نمی خورد. سپس نیروی دریایی آمریکا متوجه بالونی شد که همیشه در آسمان است؛ ماه؛ سیستمی که ساخته شد با استفاده از انعکاس امواج از ماه، بین کشتی ها و پایگاههای ساحلی ارتباط برقرار می کرد. پیشرفت بیشتر در مخابرات فضایی فقط زمانی ممکن شد که انسان اولین ماهواره های مخابراتی را به فضا

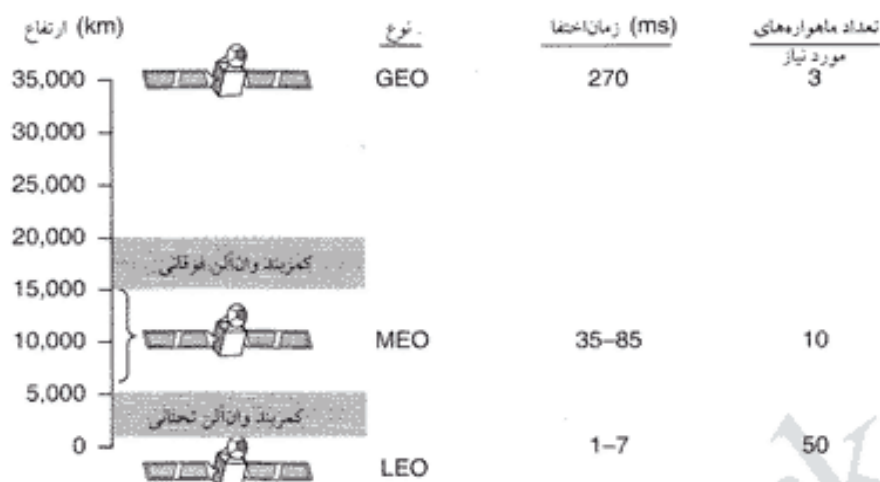


شکل ۲-۱۴. جریانهای همرفتی می تواند باعث انحراف پرتوهای لیزر شود.

پرتاب کرد. تفاوت کلیدی بین ماهواره های مصنوعی و اجسام فضایی این بود که ماهواره های ساخته دست بشر می توانستند قبل از برگشت دادن سیگنال آنها تقویت کنند - بدین ترتیب، یک کنجکاو عجیب تبدیل به یکی از ارکان زندگی بشر شد.

ماهواره های مخابراتی چند ویژگی دارند که آنها را برای کاربردهای مختلف جذاب می کند. یک ماهواره مخابراتی، در ساده ترین شکل خود، یک تکرارکننده مایکروویو بزرگ در آسمان است. چنین ماهواره ای دارای چندین ترانسپاندر (تکرارکننده - transponder) است، که به فرکانسهای خاصی گوش می کنند، سیگنال دریافتی را تقویت کرده، و سپس آنها را با فرکانس متفاوتی برمی گردانند (تا با سیگنال ورودی تداخل نکند). پرتو برگشتی می تواند وسیع باشد و بخش بزرگی از سطح کره زمین را بپوشاند، یا اینکه باریک باشد و فقط منطقه ای به قطر چند صد کیلومتر مربع را پوشش دهد. به این حالت شیپور خمیده (bent pipe) گفته می شود.

طبق قانون کپلر، دوره گردش مداری یک ماهواره با توان $\frac{3}{4}$ شعاع مدار آن متناسب است - هر چه مدار ماهواره بالاتر باشد، دوره گردش آن بیشتر است. در مدارهای نزدیک زمین دوره گردش ماهواره ها معمولاً ۹۰ دقیقه است، و آنها سرعت از دید ایستگاههای زمینی خارج می شوند. برای پوشش دادن تمام سطح کره زمین با چنین ماهواره هایی، به تعداد زیادی از آنها نیاز است. در مداری به ارتفاع 35,800 km دوره گردش ماهواره ۲۴ ساعت (معادل یک دور گردش زمین به دور خود) است. اگر ارتفاع ماهواره را تا 384,000 km بالا ببریم، دوره گردش آن یک ماه خواهد شد - اگر باور ندارید، به ماه نگاه کنید؛ ارتفاع مداری ماه از سطح زمین دقیقاً همین مقدار است. دوره گردش ماهواره اهمیت زیادی دارد، ولی تنها عاملی نیست که تعیین می کند ماهواره کجا باید قرار گیرد. عامل مهم دیگر، وجود کمربند وان آلن (Van Allen belt) است - کمربند وان آلن لایه ای از ذرات باردار پُرانرژی است که توسط میدان مغناطیسی زمین بدام افتاده اند. هر ماهواره ای که در داخل این لایه پرواز کند.



شکل ۲-۱۵. ماهواره‌های مخابراتی و برخی از ویژگی‌های آنها، از جمله ارتفاع از سطح زمین، زمان تأخیر رفت و برگشت سیگنال، و تعداد ماهواره‌های که برای پوشش کل زمین لازم است.

بسرعت توسط ذرات باردار پُرانرژی آن از بین می‌رود. از ترکیب این دو عامل سه ناحیه مشخص می‌شود، که می‌توان ماهواره‌ها را با اطمینان خاطر در آنها قرار داد (شکل ۲-۱۵ را ببینید). در قسمتهای آینده ماهواره‌هایی را که در هر یک از این ناحیه‌ها قرار داده می‌شوند، تشریح خواهیم کرد.

۱-۴-۲ ماهواره‌های زمین ثابت

در سال ۱۹۴۵، نویسنده‌ی داستانهای علمی-تخیلی آر تور سی. کلارک با محاسبات خود نشان داد که اگر ماهواره‌ای در یک مدار استوایی و در ارتفاعی معادل ۳۵,۸۰۰ km قرار گیرد، نسبت به زمین ثابت بنظر خواهد رسید، و نیازی نیست که ایستگاه زمینی آنرا تعقیب کند (Clarke, 1945). وی در این مقاله یک سیستم کامل مخابراتی را با استفاده از این ماهواره‌های زمین ثابت (geostationary satellites) تشریح کرده بود، و از جمله مدار آنها، پانلهای خورشیدی (برای تأمین انرژی مورد نیاز)، فرکانسهای رادیویی، و نحوه پرتاب ماهواره‌ها را بدقت توضیح داده بود. متأسفانه، وی در این مقاله نتیجه گرفته بود که این طرح غیر عملی است، چون قرار دادن تقویت‌کننده‌های لامپ خلا (که بسیار پرمصرف، حجیم و شکننده بودند) در مدار زمین غیر ممکن است، و هرگز این ایده را دنبال نکرد، اگر چه چند داستان علمی-تخیلی در این زمینه نوشت.

اختراع ترانزیستور همه چیز را تغییر داد، و اولین ماهواره مخابراتی بنام تل‌استار (Telstar) در ژوئیه ۱۹۶۲ به فضا پرتاب شد. از آن زمان به بعد، ماهواره‌های مخابراتی صنعتی چندمیلیارد دلاری را بوجود آورده‌اند، که تنها زمینه سودآور تحقیقات فضایی است. به این ماهواره‌های بلندپرواز اغلب ماهواره‌های GEO (Geostationary Earth Orbit) گفته می‌شود.

با تکنولوژی موجود (برای جلوگیری از تداخل امواج) حداقل فاصله دو ماهواره زمین ثابت نمی‌تواند کمتر از ۲ درجه در صفحه استوایی باشد. بدین ترتیب، در هر زمان بیش از $360 / 2 = 180$ نمی‌توانند در مدار باشند. با این حال برای بالا بردن پهنای باند در دسترس، می‌توان در هر ترانسپاندر از فرکانسها و پولاریزاسیونهای مختلف استفاده کرد.

در این مورد هم برای جلوگیری از هرج و مرج در مدار زمین، تخصیص مکانهای مداری توسط ITU انجام می‌شود. پای سیاست به اینجا هم باز شده است: کشورهایی که بزحمت از عصر حجر فاصله گرفته‌اند، سهم خود

را از مکانهای مداری طلب می کنند، تا بعد بتوانند آنرا به قیمت خوب به متقاضیان اجاره بدهند. کشورهایی هم هستند که ادعا می کنند قلمرو خاک آنها تا کره ماه ادامه دارد، و هیچکس حق ندارد در آسمان بالای سر آنها ماهواره داشته باشد. و برای شلوغی بیشتر اوضاع، فقط مخابراتی ها نیستند که بر سر تصاحب مدارها می جنگند: تلویزیونهای ماهواره ای، دولتها، و نظامیها هم می خواهند سهمی از مدارهای زمین داشته باشند.

ماهواره های امروزی بسیار بزرگ هستند، که وزن آنها گاهی به ۴۰۰۰ کیلوگرم می رسد، و دهها کیلووات انرژی الکتریکی مصرف می کنند (انرژی که توسط پانلهای خورشیدی تولید می شود). مدار و موقعیت این ماهواره ها تحت تأثیر جاذبه خورشید، ماه و سایر سیارات منظومه شمسی پیوسته در حال تغییر است، که این تأثیرات توسط موتورهای موشکی کوچکی که در آنها تعبیه شده، خنثی می شود (به این کار نگهداری ایستگاه - station keeping - می گویند). از آنجائیکه سوخت این موتورها محدود است، بعد از مدتی (که معمولاً حدود ۱۰ سال است) ماهواره بکلی از کنترل خارج و بلااستفاده می شود. ارتفاع چنین ماهواره هایی بتدریج کاهش می یابد، و پس از ورود به جو زمین می سوزند، و یا در برخورد با سطح زمین متلاشی می شوند.

مکانهای مداری تنها چیزی نیست که سر آن دعواست. فرکانسها نیز خواهان زیادی دارد، چون احتمال تداخل فرکانسهای ارسالی از ماهواره (downlink) با فرکانسهای زمینی وجود دارد. برای رفع این مشکل، ITU چند باند فرکانسی را به کاربردهای ماهواره ای اختصاص داده است، که آنها را در شکل ۲-۱۶ ملاحظه می کنید. باند C اولین باندی بود که به کاربردهای تجاری اختصاص یافت. این باند دو محدوده دارد: محدوده پائینی که برای ارسال از ماهواره (downlink) استفاده می شود، و محدوده بالایی که به ارسال به ماهواره (uplink) تخصیص یافته است. این باندها بسیار شلوغ هستند، چون در مخابرات مایکروویو زمینی هم از آنها استفاده می شود. باندهای L و S در سال ۲۰۰۰ طبق توافقهای بین المللی اضافه شدند، اما این باندها نیز بسیار باریک و شلوغ هستند.

باند بعدی که در اختیار مخابرات تجاری قرار دارد، باند Ku (K under) است. این باند هنوز پُر نشده، و در فرکانسهای آن می توان ماهواره ها را تا فاصله ۱ درجه مداری به هم نزدیک کرد، ولی یک مشکل عمده دارد: باران. آب جاذب خوبی برای انرژی امواج در این طول موج است. خوشبختانه، طوفانهای بزرگ معمولاً به یک منطقه کوچک محدود هستند، و اگر بجای یک ایستگاه زمینی از چندین ایستگاه با فواصل زیاد استفاده کنیم، این مشکل مرتفع خواهد شد (که البته این راه حل هزینه بسیار بالایی دارد). باند Ka (K above) نیز به مصارف تجاری تخصیص یافته، ولی تجهیزات آن هنوز بسیار گران است. علاوه بر این باندهای تجاری، دهها باند دولتی و نظامی نیز وجود دارد.

یک ماهواره امروزی در حدود ۴۰ ترانسپاندر دارد، که پهنای باند هر کدام از آنها 80-MHz است. معمولاً هر ترانسپاندر بعنوان یک شیپور خمیده عمل می کند، ولی در ماهواره های جدیدتر (که امکانات پردازشی بیشتری دارند) می توان بگونه ای دیگر نیز عمل کرد. در ماهواره های اولیه، تقسیم ترانسپاندرها به کانالهای مختلف استاتیک بود: پهنای باند بصورت ساده به چند باند فرکانسی ثابت تقسیم می شد. امروزه، پرتو ترانسپاندر به بُرشهای زمانی

باند	دریافت از ماهواره	ارسال به ماهواره	پهنای باند	مشکلات
L	1.5 GHz	1.6 GHz	15 MHz	پهنای باند کم به شلوغ
S	1.9 GHz	2.2 GHz	70 MHz	پهنای باند کم به شلوغ
C	4.0 GHz	6.0 GHz	500 MHz	تداخل زمینی
Ku	11 GHz	14 GHz	500 MHz	باران
Ka	20 GHz	30 GHz	3500 MHz	باران با قیمت بالای تجهیزات

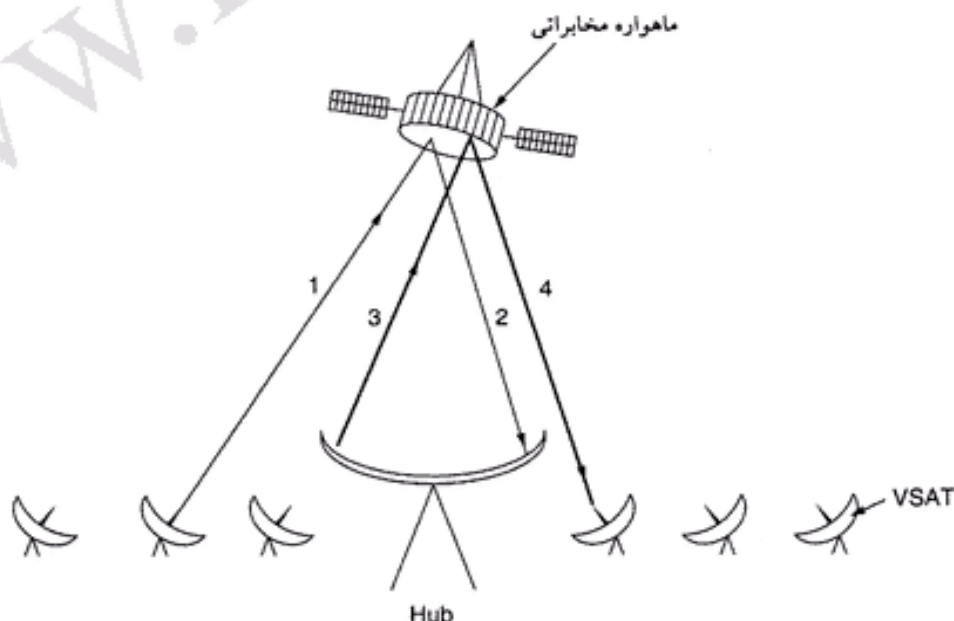
شکل ۲-۱۶. باندهای ماهواره ای.

تقسیم می‌شود، و هر کاربر می‌تواند بنویسند از آن استفاده کند. در قسمتهای آینده همین فصل این دو تکنیک (مالتی پلکس تقسیم فرکانسی و مالتی پلکس تقسیم زمانی) را به تفصیل بررسی خواهیم کرد.

اولین ماهواره‌های زمین ثابت یک پرتو فضایی واحد داشتند، که تقریباً $\frac{1}{3}$ سطح زمین را می‌پوشاند (به این ناحیه جای پای ماهواره - footprint - گفته می‌شود). اما با افت شدید قیمت، اندازه، و توان مصرفی وسایل میکروالکترونیک، استراتژیهای هوشمندانه تری امکان پذیر گردید. هر ماهواره به چندین آنتن و ترانسپاندر مجهز شد، و پرتو ارسالی از ماهواره چنان باریک شد که فقط منطقه کوچکی را در بر می‌گرفت، و بدین ترتیب امکان مخابرات همزمان چندین سیگنال بوجود آمد. این پرتوهای نقطه‌ای (spot beam) معمولاً بیضی شکل هستند، و می‌توان آنها را بقدری باریک کرد که فقط ناحیه‌ای به قطر چند صد کیلومتر را پوشش دهند. برای مثال، یک ماهواره مخابراتی که فقط برای ایالات متحده آمریکا به فضا پرتاب شده، را می‌توان طوری تنظیم کرد که یک پرتو برای ۴۸ ایالت خاک اصلی آمریکا، یک پرتو برای آلاسکا، و یکی برای هاوایی داشته باشد.

یکی از پیشرفتهای جدید در دنیای مخابرات ماهواره‌ای، ایستگاههای ارزان قیمتی هستند که VSAT (ترمینال با باریکه بسیار کوچک - Very Small Aperture Terminal) نامیده می‌شوند (Abramson, 2000). این ترمینالهای بسیار کوچک آنتنهایی بقطر ۱ متر یا کمتر دارند (برای مقایسه، قطر آنتنهای استاندارد GEO ۱۰ متر است)، و قدرت تشعشعی آنها در حدود ۱ وات است. سرعت ارسال این سیستمها در حدود 19.2-kbps است، ولی می‌توانند تا 512-kbps یا بیشتر دریافت کنند. تلویزیونهای ارسال مستقیم ماهواره‌ای (DB-SAT) معمولاً از این تکنولوژی استفاده می‌کنند.

در بسیاری از سیستمهای VSAT، ایستگاههای زمینی بدلیل توان پائین نمی‌توانند مستقیماً با هم ارتباط برقرار کنند (البته از طریق ماهواره). در این موارد، از یک ایستگاه زمینی خاص، با آنتنی بزرگ و قوی بعنوان هاب (hub)، برای رله کردن سیگنالها از یک VSAT به VSAT دیگر استفاده می‌کنند (شکل ۲-۱۷). در این حالت، یکی از طرفین باید آنتنی بزرگ با یک تقویت‌کننده قوی داشته باشد. زمان تأخیر سیگنال در این روش بیشتر از ارتباط



شکل ۲-۱۷. چند سیستم VSAT به‌همراه یک هاب.

مستقیم است، ولی کاربران معمولاً هزینه کمتر را به کمی تأخیر ترجیح می‌دهند. سیستمهای VSAT بیشترین کاربرد را در مناطق روستایی دارد. متأسفانه، هنوز نیمی از جمعیت دنیا به تلفن دسترسی ندارند، و کشیدن خطوط تلفن به هزاران دهکده کوچک و دورافتاده از توان اغلب کشورهای جهان سوم خارج است - ولی نصب یک آنتن VSAT یک متری (که با سلولهای خورشیدی کار کند) کار چندان مشکلی نیست. VSAT تکنولوژی است که می‌تواند همه مردم دنیا را به هم وصل کند.

مخابرات ماهواره‌ای ویژگیهایی دارد که بشدت با ارتباطات نقطه-به-نقطه زمینی متفاوتند. اولین ویژگی اینست که، با آنکه سیگنالهای ماهواره‌ای با سرعت نور (نزدیک $300,000 \text{ km/sec}$) حرکت می‌کنند، (بعلت فاصله زیاد ماهواره تا سطح زمین) ارتباط با ماهواره‌های GEO دارای زمان تأخیر سیگنال قابل ملاحظه‌ای است. بسته به فاصله کاربر از ایستگاه زمینی و زاویه ماهواره نسبت به افق در آن محل، زمان ارتباط نقطه-به-نقطه بین 25° تا 30° میلی‌ثانیه متغیر است. زمان متوسط این تأخیر 27° میلی‌ثانیه است (که برای سیستمهای VSAT دارای هاب به دو برابر، یعنی 54° میلی‌ثانیه، می‌رسد).

برای مقایسه بد نیست بدانید که، زمان تأخیر انتشار در لینکهای میکروویو زمینی حدود $3 \mu\text{sec/km}$ ، و برای کابلهای کواکسیال و فیبر نوری تقریباً $5 \mu\text{sec/km}$ است (امواج الکترومغناطیس در هوا سریعتر از مواد جامد حرکت می‌کنند).

ویژگی مهم دیگر ماهواره‌ها اینست که آنها ذاتاً رسانه‌های پخش هستند؛ فرستادن پیام برای یک نفر هیچ تفاوتی با هزاران نفر ندارد. در برخی از کاربردها (مانند تبلیغات اینترنتی) این ویژگی بسیار مفید است. با اینکه بوسیله ارتباطات نقطه-به-نقطه هم می‌توان چنین وضعیتی را شبیه سازی کرد، ولی این کار با استفاده از ماهواره بسیار ارزاتر تمام می‌شود. از سوی دیگر، از دیدگاه امنیت و حفظ حریم خصوصی افراد، ماهواره یک فاجعه تمام عیار است: هر کسی می‌تواند پیامهای خصوصی دیگران را بشنود. اینجاست که اهمیت رمزنگاری (encryption) روشن می‌شود.

در مخابرات ماهواره‌ای، هزینه انتقال پیام به فاصله فرستنده و گیرنده بستگی ندارد: تماس با آن سوی اقیانوسها از نظر هزینه هیچ فرقی با تماس با خانه و برویی ندارد. مخابرات ماهواره‌ای از نظر نرخ خطا بسیار عالیست، و زمان به بهره‌برداری رسیدن آن نیز بسیار کوتاه است (نکته‌ای که در مخابرات نظامی بسیار اهمیت دارد).

۲-۴-۲ ماهواره‌های مدار متوسط

در مداری بسیار پائینتر از مدار زمین‌ثابت، و بین دو کمربند وان‌الن، ماهواره‌های مدار متوسط که به MEO (Medium-Earth Orbit) معروفند، قرار می‌گیرند. این ماهواره‌ها بطور متوسط هر ۶ ساعت یکبار دور زمین می‌گردند، و بهمین دلیل ایستگاه زمینی باید آنها را تعقیب کند. بعلت ارتفاع پائین، ماهواره‌های MEO جای پای کوچکتری نسبت به ماهواره‌های GEO دارند، ولی توان تشعشعی لازم برای ارسال به آنها نیز بسیار کمتر است. در حال حاضر از این ماهواره‌ها برای مقاصد مخابراتی استفاده نمی‌شود، بنابراین ما هم درباره آنها بیش از این صحبت نخواهیم کرد. ماهواره‌های ۲۴ گانه GPS (سیستم مکان‌یابی جهانی - Global Positioning System) که در ارتفاع 18,000 km پرواز می‌کنند، در این دسته قرار می‌گیرند.

۳-۴-۲ ماهواره‌های مدار پائین

اگر باز هم پائینتر ببینیم، به ماهواره‌های مدار پائین (Low-Earth Orbit) LEO می‌رسیم. بدلیل سرعت زیاد گردش مداری این ماهواره‌ها، برای ایجاد سیستمی با پوشش جهانی به تعداد زیادی از آنها نیاز داریم. از طرف دیگر، بدلیل ارتفاع کم ماهواره‌های LEO، برای ارتباط با آنها به توان کمی نیاز است، و زمان رفت و برگشت

سیگنال نیز بسیار کم (در حدود چند میلی‌ثاذه) خواهد بود. در این قسمت سه سیستم ماهواره‌ای LEO را - که دو نای آنها به سرویس صدا و یکی به سرویسهای اینترنت اختصاص دارند - مورد بررسی قرار خواهیم داد.

ایریدیوم

همانطور که گفتیم بعلافت سرعت زیاد گردش مداری ماهواره‌های LEO - و عبور سریع از مقابل آنتنهای زمینی - تا همین اواخر از این ماهواره‌ها برای مقاصد مخابراتی استفاده نمی‌شد. این وضعیت در سال ۱۹۹۰ تغییر کرد: در این سال موتورولا درخواستی برای دریافت مجوز پرتاب ۷۷ ماهواره مخابراتی LEO به FCC ارائه کرد - این پروژه ایریدیوم (Iridium) نام داشت (ایریدیوم عنصر ۱۷۷ام جدول تناوبی است). بعدها این پروژه تغییر کرد، و قرار شد از فقط ۶۶ ماهواره استفاده شود - بالطبع نام پروژه را هم باید به دیسپروسیوم (عنصر ۶۶ام جدول تناوبی) تغییر می‌دادند، اما نام دیسپروسیوم بیشتر شبیه یک بیماری خطرناک است تا یک سیستم ماهواره‌ای! ایده اصلی در این سیستم آن است که به محض خارج شدن یک ماهواره از دید آنتن زمینی، ماهواره دیگری بلافاصله جای آنرا می‌گیرد. این پیشنهاد به یک هیجان عمومی در میان شرکتهای مخابراتی دامن زد، و همه خواهان آن بودند که زنجیره‌ای از ماهواره‌های LEO به فضا پرتاب کنند.

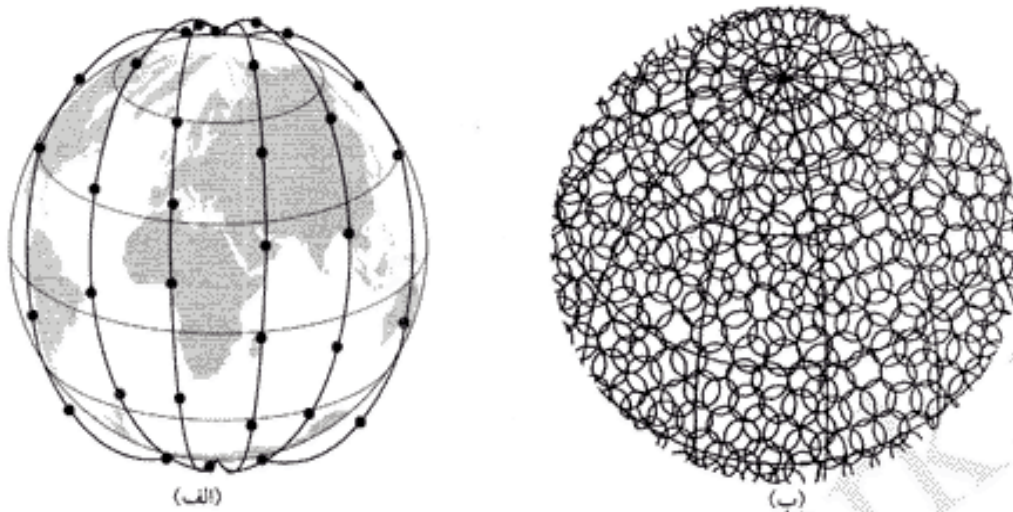
بعد از هفت سال اشتراک مساعی فنی و مالی بین چند شرکت بزرگ، بالاخره ماهواره‌های ایریدیوم در سال ۱۹۹۷ به فضا پرتاب شدند، و از نوامبر ۱۹۹۸ سرویسهای مخابراتی آنها شروع بکار کرد. اما متأسفانه بدلیل گسترش شبکه تلفنهای همراه، تقاضای ناچیزی برای این سرویسها در بازار وجود داشت، و متعاقب آن در آگوست ۱۹۹۹ پروژه ایریدیوم (طی یکی از فاجعه‌بارترین ورشکستگی‌های تاریخ) متوقف شد. ماهواره‌ها و سایر تجهیزات این پروژه (که ۵ میلیارد دلار ارزش داشت) به قیمت ۲۵ میلیون دلار به یک سرمایه‌گذار فروخته شد؛ در واقع، پروژه ایریدیوم را باید بزرگترین بازار اسقاطی فضایی بشمار آورد. سرویسهای ایریدیوم از مارس ۲۰۰۱ دوباره راه‌اندازی شد.

کار اصلی ایریدیوم ارائه سرویسهای مخابراتی از طریق تجهیزاتی که مستقیماً با ماهواره ارتباط برقرار می‌کنند، بود (و هست). ایریدیوم سرویسهای صدا، داده، فکس، پیجر، و هدایت و ناوبری را در تمام نقاط زمین (خشکی، دریا و هوا) در اختیار کاربران خود می‌گذارد. کاربران این سرویسها را عمدتاً دریانوردان، هوانوردان، کارکنان سکوها، اکتشاف نفت، و مسافران و جهانگردانی که به مناطق فاقد زیرساختهای مخابراتی (مانند کوه، صحرا، جنگل، و برخی از کشورهای جهان سوم) سفر می‌کنند، تشکیل می‌دهند.

ماهواره‌های ایریدیوم در مدار قطبی و در ارتفاع ۷۵۰ کیلومتری زمین پرواز می‌کنند. این ماهواره‌ها بصورت کمربندهایی که از قطبهای شمال-جنوب زمین می‌گذرند، آرایش یافته‌اند، و با یکدیگر ۳۲ درجه عرض جغرافیایی فاصله دارند - شکل ۲-۱۸ (الف) را ببینید. برای پوشش دادن تمام سطح زمین شش تا از این حلقه‌ها کفایت می‌کند. آنهایی که کمی با شیمی آشنا باشید، می‌توانند این ماهواره‌ها را الکترونهای یک اتم غول‌آسای دیسپروسیوم فرض کنند که بدور هسته اتم (کره زمین) در گردشند.

هر ماهواره حداکثر ۴۸ سلول (پرتو نقطه‌ای) دارد، که بدین ترتیب تعداد کل سلولها به ۱۶۲۸ می‌رسد، و می‌توانند کل سطح زمین را پوشش دهند - شکل ۲-۱۸ (ب) را ببینید. هر ماهواره ایریدیوم ۳۸۴۰ کانال - و کل سیستم ۲۵۳،۴۴۰ کانال - ظرفیت دارد، و می‌توان از آنها برای سرویسهای مختلف استفاده کرد.

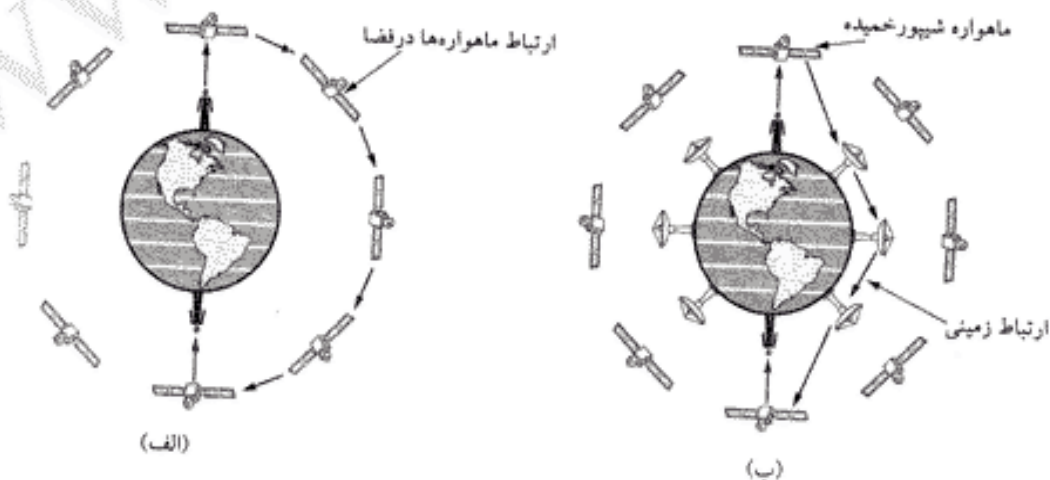
یکی از نکات جالب در مورد ایریدیوم اینست که ارتباط کاربرانی که از یکدیگر فاصله زیادی دارند، در فضا صورت می‌گیرد (بعبارت دیگر، رله کردن اطلاعات بین ماهواره‌ها در فضا انجام می‌شود)؛ شکل ۲-۱۹ (الف) را ببینید. همانطور که در این شکل می‌بینید، وقتی کاربری در قطب شمال بخواهد با قطب جنوب تماس بگیرد، سیگنال را به ماهواره‌ای که بالای سرش قرار دارد، می‌فرستد. سپس این سیگنال از یک ماهواره به ماهواره دیگر رله می‌شود، تا به قطب جنوب برسد.



شکل ۲-۱۸. (الف) ماهواره های ابریدیوم شش حلقه بدور زمین می سازند. (ب) تعداد سلولهای متحرک ابریدیوم به ۱۶۲۸ می رسد.

گلوبال استار

یکی از طرحهای رقیب ابریدیوم، پروژه گلوبال استار (Globalstar) است. این پروژه از ۴۸ ماهواره LEO تشکیل شده، ولی روش سونیچینگ آن با ابریدیوم فرق دارد، و در آن از سونیچینگ زمینی استفاده می شود - شکل ۲-۱۹ (ب) را ببینید. در اینجا سیگنالها از ماهواره مبدأ به یک ایستگاه زمینی بزرگ واقع در سانتا ورک شاپ هدایت شده، و از آنجا به نزدیکترین ایستگاه زمینی نزدیک مقصد فرستاده می شوند، تا بالاخره (بعد از ارسال به ماهواره) به گیرنده برسند. مزیت این روش آنست که (بر خلاف ابریدیوم) قسمت پیچیده کار روی زمین انجام می شود، و مدیریت آن ساده تر است. همچنین، با آنتنهای بزرگ زمینی (که سیگنالهای قویتری می فرستند، و می توانند سیگنالهای ضعیفتری دریافت کنند) از تجهیزات انفرادی ساده تری می توان استفاده کرد (توان تشعشی تلفنها در حد چند میلی وات است، و حتی بعد از تقویت در ماهواره باز هم چندان قوی نیست).



شکل ۲-۱۹. (الف) رله سیگنال در فضا. (ب) رله سیگنال روی زمین.

تله‌دزیک

کاربران اصلی ایریدیوم افرادی هستند که به جاهای پرت و دورافتاده می‌روند. نمونه دیگر سیستمهای ماهواره‌ای تله‌دزیک (Teledesic) است، که کاربران اینترنت پرسرعت را (در تمام دنیا) هدف گرفته است. ایده این سیستم به کریگ مک‌کاو (پیشگام صنعت تلفنهای همراه) و بیل گیتس (بنیانگذار مایکروسافت) تعلق دارد، که از سرعت لاک‌پشتی دسترسی اینترنت ناراضی بودند. هدف سیستم تله‌دزیک فراهم آوردن دسترسی به اینترنت با سرعتهای بالا (تا 100-Mbps ارسال، و تا 720-Mbps دریافت) برای میلیونها کاربر همزمان است. این سیستم از آنتنهای کوچک و ثابت (شبیه VSAT) استفاده می‌کند، و هیچ نیازی به ارتباطات تلفنی معمولی ندارد.

در طرح اولیه سیستم تله‌دزیک قرار بود از ۲۸۸ ماهواره پرتو باریک، که در دوازده گروه در ارتفاع ۱۳۵۰ کیلومتری (درست زیر کمربند وان‌آلن تحتانی) پرواز می‌کنند، استفاده شود. ولی این طرح بعدها به ۳۰ ماهواره با جای پای بزرگتر تغییر کرد. ماهواره‌ها در باند نسبتاً خلوت Ka و بصورت سونیچینگ بسته‌ای کار خواهند کرد، و می‌توانند بسته‌ها را بین خود رد و بدل کنند. وقتی یک کاربر به پهنای باند نیاز پیدا می‌کند، درخواست خود را بصورت یک بسته می‌فرستد، و ۵۰ میلی‌ثانیه بعد پهنای باند موردنیاز بصورت خودکار به وی اختصاص داده می‌شود. اگر همه چیز طبق نقشه پیش برود، این سیستم در سال ۲۰۰۵ عملیاتی خواهد شد.

۲-۴-۴ ماهواره یا فیبر؟

مقایسه‌ای بین مخابرات ماهواره‌ای و زمینی آموزنده خواهد بود. تا همین ۲۰ سال پیش هیچکس شک نداشت که آینده مخابرات متعلق به ماهواره‌هاست، چون سیستمهای تلفن در ۱۰۰ سال گذشته تغییر عمده‌ای نکرده بودند، و بنظر نمی‌رسد در ۱۰۰ سال بعدی هم هیچ اتفاق خاصی بیفتد. شرکت‌های تلفن سرویسهای صوتی خوب (با قیمت مناسب) ارائه می‌کردند، و سرمایه‌گذاری آنها سودی تضمین شده داشت. برای آنهایی هم که سرویس داده می‌خواستند، مودمهای 1200-bps موجود بود - و این تمام بضاعت شرکت‌های تلفن بود.

اما در سال ۱۹۸۴ آتش رقابت ایالات متحده آمریکا و اروپا را در بر گرفت، و اوضاع تغییر کرد. شرکت‌های تلفن در سرویسهای راه‌دور فیبرهای نوری را جایگزین کابلهای مسی کردند، و سرویسهای با پهنای باند زیاد مانند ADSL (خط دیجیتال نامتقارن) در اختیار کاربران خود قرار دادند. تغییر اساسی دیگر کاهش نرخهای مخابرات راه‌دور بود، که تا آن زمان بطور مصنوعی و به نفع کاربران محلی بالا نگه داشته شده بود. بنظر می‌رسید فیبر نوری برنده جنگ باشد، ولی ماهواره هم مزایایی دارد که فیبر نمی‌تواند (و احتمالاً نخواهد توانست) با آنها رقابت کند. اول اینکه، پهنای باند فیبر (که شاید یک رشته آن از تمام ماهواره‌های پرتاب شده بیشتر باشد) در اختیار اغلب کاربران نیست. فیبر نوری بیشتر در مخابرات راه‌دور و برای متصل کردن شبکه‌های تلفن بکار می‌رود، تا رساندن پهنای باند بالا به کاربران منفرد. در حالیکه ماهواره‌ها (به کمک یک آنتن ساده در بالای پشت بام) می‌توانند مستقیماً و بدون هیچ واسطه‌ای پهنای باند زیاد را در اختیار تک تک کاربران قرار دهند. سیستم تله‌دزیک بر اساس همین ایده شکل گرفته است.

مخابرات سیار دیگر قلمرو دست‌نیافتنی ماهواره‌هاست. امروزه بسیاری از افراد مایلند در حال قدم‌زدن، اتومبیل‌سواری، و حتی قایقرانی و پرواز به سرویسهای مخابراتی دسترسی داشته باشند. اینجا دیگر فیبر نوری بکلی بی‌استفاده است، و فقط مخابرات ماهواره‌ای می‌تواند راه چاره باشد. البته برای آنهایی که شعاع حرکتشان محدود است، می‌توان از ترکیب فیبر و بیسیمهای رادیویی استفاده کرد، ولی در هوا و دریا این روش هم دیگر کارایی ندارد.

مزیت دیگر ماهواره در مخابرات پخش (broadcasting) است. وقتی بخواهید یک پیام را در آن واحد به هزاران نفر برسانید (مانند ارسال نرخ کالا و ارز، یا قیمت سهام و اوراق قرضه)، هیچ چیز جای ماهواره را (از نظر

سهولت و هزینه) نمی‌گیرد.

پراکندگی جمعیت و سرزمین از دیگر عواملیست که بکارگیری مخابرات ماهواره‌ای را مقرون بصرفه می‌کند. برای مثال، کشور اندونزی برای هدایت ترافیک تلفن داخلی خود نیز از ماهواره استفاده می‌کند، چون پرتاب یک ماهواره بسیار ساده‌تر است، تا کشیدن کابلهای زیر دریایی به ۱۳۶۷۷ جزیره‌ای که این کشور را تشکیل می‌دهند. در جاهایی که تملک زمین برای کشیدن کابل زمینی دشوار (و یا پرهزینه) است، نیز ماهواره مزیت نسبی دارد. سرعت نصب و راه‌اندازی سیستم نیز یکی از جواهراتیست که ماهواره بر کابل زمینی پیروز می‌شود. وقتی جنگی در می‌گیرد، و نیروهای نظامی به تماس با نقاط جدید نیاز پیدا می‌کنند، پرتاب یک ماهواره معمولاً سریعترین راه‌حل ممکن است.

بطور خلاصه، بنظر می‌رسد که جریان اصلی مخابرات در آینده بر فیبرهای نوری و تلفنهای همراه متکیست، ولی در برخی از کاربردهای خاص ماهواره‌ها برتری دارند. اما همه اینها تابع یک چیز هستند: اقتصاد. با اینکه فیبر نوری پهنای باند بیشتری ارائه می‌کند، اما حرف آخر را قیمت می‌زند. اگر تکنولوژی جدیدی اختراع شود که هزینه پرتاب ماهواره‌ها را بشدت کاهش دهد (مانند ساتلهای فضایی که بتوانند هر بار دهها ماهواره را به مدار ببرند)، یا کاربرد غیرمتظره‌ای برای ماهواره‌های مدار پائین پیدا شود، معلوم نیست که فیبر نوری در تمام زمینه‌ها برنده باشد.

۵-۲ شبکه تلفن عمومی

اگر بخواهیم دو کامپیوتر را که نزدیک به هم و یا در یک ساختمان هستند، به یکدیگر متصل کنیم، کار بسیار ساده است و فقط کفایت یک رشته کابل بین آنها بکشیم - این همان شبکه محلی یا LAN است. اما اگر فاصله کامپیوترها زیاد باشد، یا کابل بایستی از املاک خصوصی یا شهری عبور داده شود، هزینه انجام کابل کشی معمولاً به مانعی بزرگ تبدیل می‌شود (البته اگر این کار غیرقانونی نباشد، که در اغلب کشورها هست). در نتیجه، طراحان شبکه به تأسیسات مخابراتی موجود روی می‌آورند.

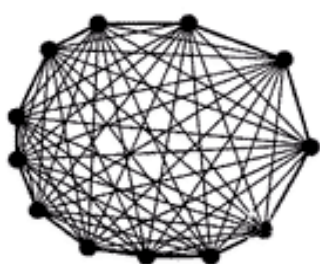
این تأسیسات بویژه شبکه تلفن عمومی - (Public Switched Telephone Network) PSTN - سالها قبل و با هدفی دیگر طراحی شده‌اند: انتقال صدای انسان بگونه‌ای کمابیش قابل تشخیص. کارایی این تجهیزات برای ارتباط کامپیوتر-کامپیوتر در بهترین حالت اغلب حاشیه‌ای است، اما با به بازار آمدن فیبر نوری و تکنولوژیهای دیجیتال این وضعیت سرعت در حال تغییر است. در هر حال، سیستم تلفن عمومی چنان با شبکه‌های (گسترده) کامپیوتری عجین است، که جا دارد وقت بیشتری به بررسی آن اختصاص دهیم.

برای بهتر شکافتن صورت مسئله، اجازه دهید مقایسه‌ای کلی (ولی روشن‌کننده) بین ارتباط دو کامپیوتر از طریق کابل مستقیم شبکه و از طریق خطوط تلفن داشته باشیم. کابل مستقیم شبکه می‌تواند داده‌ها را با سرعتی معادل 10^9 bps (یا حتی بیشتر) منتقل کند، در حالیکه حداکثر سرعت یک خط تلفن 56-kbps بیشتر نیست - تفاوتی در حد ۲۰,۰۰۰ برابر. تفاوت این دو مانند تفاوت سرعت یک مرغابی که سلانه سلانه در علفزار راه می‌رود، با موشکی است که به طرف ماه پرواز می‌کند. اگر به جای خط تلفن از اتصال ADSL استفاده کنیم، تفاوت سرعت به ۱۰۰۰ تا ۲۰۰۰ برابر خواهد رسید.

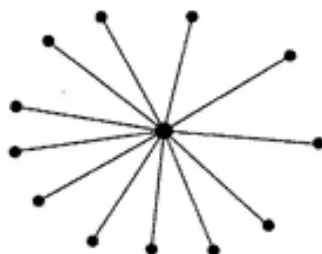
ناگفته پیداست که چنین سطح از تفاوتی برای طراحان سیستمهای کامپیوتری مشکل‌ساز است، و آنان تمام تلاش خود را برای بهینه کردن آن متمرکز کنند. در قسمتهای آینده سیستم تلفن و طرز کار آن را تشریح خواهیم کرد؛ برای اطلاعات بیشتر در این زمینه به (Bellamy, 2000) مراجعه کنید.

۱.۵.۲ ساختار سیستم تلفن

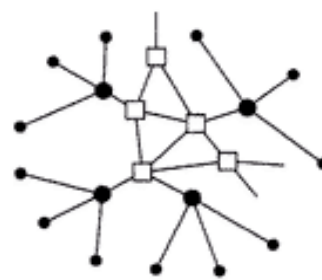
بلافاصله بعد از آن که آلكساندر گراهام بل در سال ۱۸۷۶ (و درست چند ساعت زودتر از رقیبش، الیساگری)



(الف)



(ب)



(ج)

شکل ۲-۲۰. (الف) شبکه‌ای با اتصالات داخلی کامل. (ب) سوئیچ مرکزی. (ج) سلسله مراتب دو سطحی.

اختراع خود را به ثبت رساند، سبیل تقاضا برای آن سرازیر شد. تلفنهای اولیه صورت جفتی کار می‌کردند، و مشتریان مجبور بودند بین خودشان یک رشته سیم بکشند (مسیر برگشت الکترونها از زمین بود). اگر کسی می‌خواست با n نفر تماس تلفنی داشته باشد، مجبور بود به این n نقطه سیم بکشد. در کمتر از یک سال شهرها تبدیل شدند به جنگلی از سیمهای تو در تو که از خانه‌ای به خانه دیگر (از فراز ساختمانها و درختها) کشیده شده بود. خیلی زود معلوم شد که مدل اتصال هر تلفن به تمام تلفنهای دیگر مدلی عملی نیست - شکل ۲-۲۰ (الف). خوشبختانه بل خیلی زود متوجه این مشکل شد، و با تأسیس شرکتی بنام شرکت تلفن بل، اولین مرکز سوئیچینگ را در سال ۱۸۷۸ (در نیویورک، کانتیکات) راه‌اندازی کرد. این شرکت یک رشته کابل به خانه یا دفتر هر مشتری می‌کشید. برای برقراری تماس، مشتری دسته تلفن را می‌چرخاند تا زنگی در مرکز تلفن بصدا در آید و توجه اپراتور جلب شود؛ سپس این اپراتور ارتباط وی را توسط یک رشته کابل با مقصد موردنظر برقرار می‌کرد. این مدل را در شکل ۲-۲۰ (ب) ملاحظه می‌کنید.

بزودی در هر شهر و دهکده‌ای مراکز سوئیچینگ بل دایر شد، و بل مجبور شد برای برقراری تماسهای راه دور این مراکز را نیز به یکدیگر متصل کند. اما در اینجا هم همان مشکل قبلی خود را نشان داد: ارتباط مستقیم هر مرکز سوئیچینگ با تمام مراکز دیگر غیرممکن بود، بنابراین مراکز سوئیچینگ سطح-دوم اختراع شد. پس از مدتی کوتاه تعداد مراکز سطح-دوم نیز بشدت افزایش یافت - شکل ۲-۲۰ (ج) را ببینید. بعدها این سلسله مراتب تا پنج سطح بالا رفت.

تا سال ۱۸۹۰ این سیستم تلفن سه بخش عمده داشت: مراکز سوئیچینگ، سیمهایی که بین مراکز سوئیچینگ و مشترکان کشیده می‌شد (این سیمها دیگر سیمهای لخت با برگشت زمین نبود، بلکه تبدیل به سیمهای زوج تابیده عایق‌دار شده بود)، و اتصالات راه دور بین مراکز سوئیچینگ. با آن که پیشرفتهایی در هر یک از این سه بخش صورت گرفته، اما مدل اولیه بل در ۱۰۰ سال گذشته تقریباً بدون تغییر باقی مانده است. برای دیدن تاریخچه‌ای مختصر از سیستم تلفن بل، به (Hawley, 1991) نگاه کنید.

تا قبل از دو پاره شدن AT&T در سال ۱۹۸۴، سیستم تلفن سیستمی با سلسله مراتب چندسطحی (و با پراکندگی زیاد) بود. در بحث زیر این ساختار تا حد زیادی ساده شده، ولی عصاره اصلی آن همچنان حفظ شده است. هر تلفن با دو رشته سیم مسی به نزدیکترین ایستگاه پایانی (end office یا local central office) وصل می‌شود. فاصله این دو معمولاً بین ۱ تا ۱۰ کیلومتر است (و در شهرها کمتر از مناطق روستائیت). فقط در ایالات متحده آمریکا نزدیک به ۲۲,۰۰۰ ایستگاه پایانی وجود دارد. به اتصال دو-سیمه بین ایستگاه پایانی و مشترک تلفن

اصطلاحاً مدار پایانی (local loop) گفته می شود. اگر مدارهای پایانی موجود در سراسر دنیا را بدنبال هم ردیف کنند، می تواند هزار بار فاصله زمین تا ماه را بپیماید.

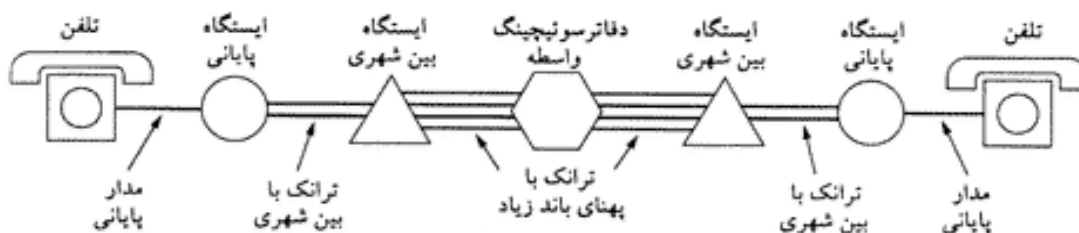
طبق یکی از تخمین ها، ۸۰ درصد ارزش سرمایه AT&T مس موجود در مدارهای پایانی آن است - بدین ترتیب، می توان AT&T را بزرگترین معدن مس دنیا بحساب آورد. خوشبختانه بورس کالا از این موضوع اطلاع چندانی ندارد - اگر اینرا می دانستند، بلافاصله AT&T را می خریدند، تمام سرویسهای تلفن را قطع می کردند، مس ها را در می آوردند، و دوباره به صنایع مس می فروختند.

وقتی یک مشترک به مشترک دیگری که به همان ایستگاه پایانی وصل است تلفن می زند، دستگاههای سوئیچینگ بین این دو مدار پایانی یک ارتباط الکتریکی مستقیم برقرار می کنند - این ارتباط مستقیم در تمام طول تماس برقرار می ماند.

اما اگر مشترک موردنظر متعلق به ایستگاه پایانی دیگری باشد، روش کار فرق می کند. هر ایستگاه پایانی دارای چند خط ارتباطی با یک یا چند مرکز سوئیچینگ، که اصطلاحاً ایستگاه بین شهری (toll office) نامیده می شوند. است - اگر این ایستگاهها در یک منطقه باشند، به آنها ایستگاه شریک (tandem office) نیز گفته می شود. این خطوط ارتباطی را ترانک های مرتبط کننده بین شهری (toll connecting trunks) می نامند. اگر ایستگاههای پایانی تلفن کننده و تلفن شونده به یک ترانک وصل باشند (که در صورت نزدیکی آنها بسیار محتمل است)، امکان دارد ارتباط در همان ایستگاه بین شهری برقرار شود. در شکل ۲-۲۰ (ج) یک شبکه تلفن ساده را می بینید، که در آن تلفنها با نقاط سیاه کوچک، ایستگاههای پایانی با نقاط سیاه بزرگ، و ایستگاههای بین شهری با مربع نشان داده شده اند.

اگر تلفن کننده و تلفن شونده دارای ایستگاه بین شهری مشترک نباشند، مسیر ارتباطی بایستی در سطح بالاتری برقرار شود. ایستگاههای بین شهری از طریق ایستگاههای اولیه (primary office)، ناحیه ای (sectional office) و منطقه ای (regional office) به هم متصل می شوند. اتصال این ایستگاهها از طریق ترانک بین شهری (intertoll trunk یا interoffice trunk) برقرار می شود. نوع مراکز سوئیچینگ و توپولوژی آنها (مثلاً، اینکه دو ایستگاه ناحیه ای می توانند مستقیماً به یکدیگر وصل شوند، یا باید این ارتباط از طریق یک ایستگاه منطقه ای باشد؟) از کشوری به کشور دیگر فرق می کند، و به تراکم تلفن در آن کشور بستگی دارد. در شکل ۲-۲۱ نحوه هدایت یک تماس راه دور را ملاحظه می کنید.

در مخابرات راه دور از رسانه های مختلفی استفاده می شود. امروزه در مدارهای پایانی از کابل Cat 3 استفاده می شود، در حالیکه در سالهای اولیه اختراع تلفن از سیمهای بدون پوششی که با فاصله ۲۵ سانتیمتر بر فراز تیرهای تلفن کشیده می شد، استفاده می کردند. بین مراکز سوئیچینگ معمولاً کابل کواکسیال، مایکروویو یا فیبر نوری بکار برده می شود.



شکل ۲-۲۱. مدار هدایت یک تماس راه دور.

در گذشته انتقال سیگنالهای تلفن بصورت آنالوگ بود: ابتدا صدا به ولتاژ الکتریکی تبدیل شده، و سپس همین سیگنال روی خط تلفن ارسال می‌شد. با اختراع فیبر نوری، الکترونیک دیجیتال و کامپیوتر، امروزه دیگر تمام ترانک‌ها و سوئیچها دیجیتال هستند، و تنها قسمتی که هنوز از تکنولوژی آنالوگ استفاده می‌کند، همان مدار پایانی است. مزیت انتقال دیجیتال در اینست که دیگر مشکل اعوجاج سیگنال در تقویت‌کننده‌های مختلف وجود ندارد، و فقط کفایت بتوانیم 0 را از 1 تشخیص دهیم. انتقال دیجیتال مطمئنتر، ارزاتر، و نگهداری آن ساده‌تر است.

بطور خلاصه، سیستم تلفن از سه قسمت عمده تشکیل شده است:

۱. مدارهای پایانی (زوجهای تابیده آنالوگ که به خانه‌ها و دفاتر کشیده می‌شوند)
۲. ترانک‌ها (فیبرهای نوری دیجیتال که مراکز سوئیچینگ را به یکدیگر متصل می‌کنند)
۳. مراکز سوئیچینگ (مراکزی که تماسهای تلفنی را از یک خط اصلی به خط دیگر هدایت می‌کنند)

اجازه دهید قبل از پرداختن به ادامه این بحث، کمی هم درباره تلفن و سیاست صحبت کنیم. بعد از آن خواهیم دید که مدارهای پایانی و ترانک‌ها چگونه عمل می‌کنند، و سوئیچینگ چگونه انجام می‌شود.

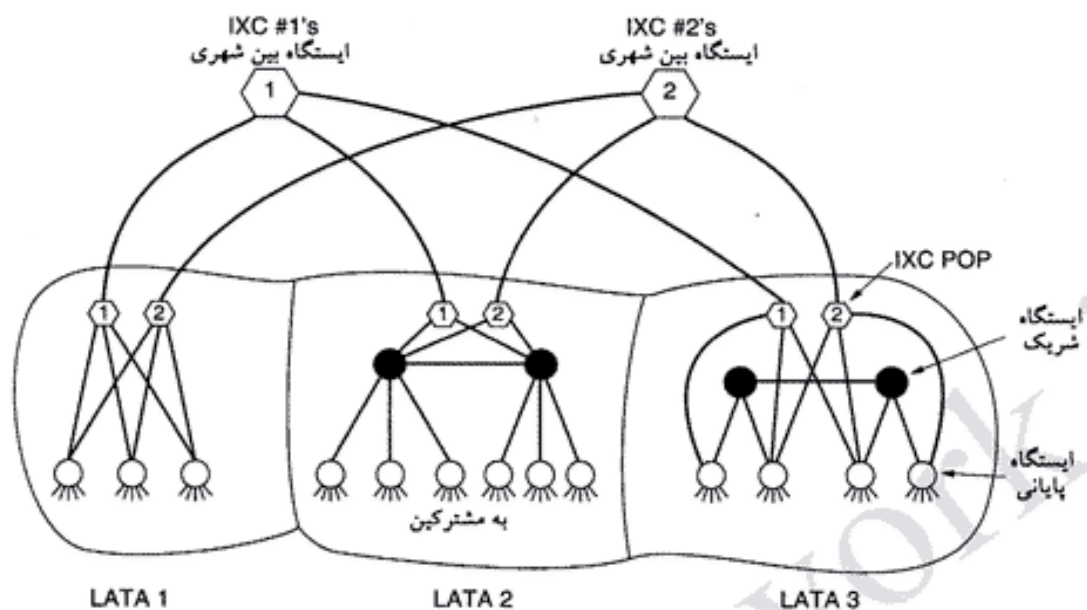
۲-۵-۲ تلفن و سیاست

تا قبل از سال ۱۹۸۴ برای چندین دهه انحصار تلفن شهری و راه دور ایالات متحده آمریکا در انحصار شرکت بل سیستم بود. در دهه ۱۹۷۰ دولت فدرال به این نتیجه رسید که این انحصار غیرقانونی است، و برای شکستن آن به دادگاه شکایت کرد. در اول ژانویه ۱۹۸۴ رأی دادگاه به نفع دولت فدرال صادر شد، و شرکت AT&T به شرکتهای AT&T Long Lines، BOC ۲۳ (شرکتهای Bell Operating Company)، و چند شعبه کوچک دیگر شکسته شد. شرکتهای BOC برای آن که بتوانند در بازار رقابت کنند، هفت گروه منطقه‌ای (RBOC) تشکیل دادند. این رأی (که به Modified Final Judgment - MFJ - معروف است) مخابرات راه دور ایالات متحده را یک شبه دچار دگرگونی اساسی کرد.

دستور MFJ منجر به تغییرات مهمی در سرویسهای مخابراتی شد: افزایش رقابت بین شرکتهای مخابراتی، بهبود سرویسها، و کاهش قیمت مخابرات راه دور. از طرف دیگر قیمت سرویسهای شهری بشدت بالا رفت، چون این شرکتهای بایستی بنوعی کاهش درآمد خود را (که قبلاً از درآمد مخابرات راه دور تأمین می‌شد) جبران می‌کردند. بسیاری از کشورهای دیگر نیز در حال پیروی از این مدل هستند.

برای آن که مشخص شود چه کسی مجاز به انجام چه کاریست، ایالات متحده به ۱۶۴ LATA (مناطق محلی دسترسی و انتقال - Local Access and Transport Area) تقسیم شد. تقسیم‌بندی LATA ها تا حد زیادی (اما نه کاملاً) بر کدهای منطقه (area code) منطبق است. هر LATA دارای یک LEC (کاربر تبادل محلی - Local Exchange Carrier) است که انحصار کامل سرویسهای تلفن در آن منطقه را در اختیار دارد. مهمترین این LEC ها همان BOC ها هستند، ولی در برخی از مناطق شرکتهای مستقل نیز (که تعداد آنها به ۱۵۰۰ می‌رسد) فعالیت دارند.

کلیه ترافیک بین LATA ها توسط شرکتهایی بنام IXC (کاربر تبادل ملی - InterExchange Carrier) انجام می‌شود. تا مدتی قبل AT&T Long Lines تنها IXC فعال در بازار بود، ولی اکنون شرکتهای بزرگی مانند Sprint و WorldCom نیز در این زمینه فعال هستند. یکی از دغدغه‌هایی که بعد از شکسته شدن AT&T وجود داشت آن بود که IXC ها در زمینه کیفیت خط، تعرفه‌ها، و تعداد رقمهای پیش‌شماره یکسان باشند. روش این در شکل ۲-۲۲ نشان داده شده است؛ در این شکل سه LATA با تعدادی ایستگاه پایانی می‌بینید. LATA های ۲ و ۳ با ایستگاه‌های بین‌شهری نیز ارتباط دارند.



شکل ۲-۲۲. ارتباط بین LATA ها و LEC ها و IXC ها

هر IXC که بخواهد مجری تماسهای یک LATA باشد، یک ایستگاه سونیچینگ بنام POP (نقطه تماس - Point Of Presence) در آنجا تأسیس می کند. وظیفه برقراری تماس IXC با ایستگاههای پایانی بر عهده LEC آن منطقه است (خواه بصورت مستقیم مانند LATA های ۱ و ۳، یا بصورت غیرمستقیم مانند LATA ۲). تماس (خواه فنی یا مالی) باید برای تمام IXC ها یکسان باشد. بدین ترتیب، مشتری که مثلاً در LATA ۱ است، می تواند تصمیم بگیرد توسط کدام IXC با مشتری در LATA ۳ تماس بگیرد.

در MFJ تصریح شده است که IXC ها نباید وارد بازار تلفن محلی شوند، و LEC ها هم از حضور در سرویسهای راه دور منع شده اند (البته آنها می توانند هر کار دیگری انجام دهند، مثلاً چیس و پفک بفروشند). در سال ۱۹۸۴ این دستور نسبتاً واضح بود، ولی تکنولوژی همواره راههای جالبی برای منسوخ کردن قوانین پیدا می کند. از آنجائیکه تلفنهای همراه و تلویزیون کابلی مشمول دستور MFJ نمی شوند، LEC ها و IXC ها بسمت خرید این شرکتها (با ادغام با آنها) روی آوردند.

در سال ۱۹۹۵ کنگره متوجه شد که جدا نگه داشتن حوزه فعالیت شرکتهای مختلف دیگر عملی نیست، و بهمین دلیل با تصویب یک لایحه به شرکتهای تلفن شهری، راه دور و تلویزیون کابلی اجازه داد تا وارد بازارهای یکدیگر شوند. ایده اصلی این لایحه اینست که سرویسهای صدا، داده و تلویزیون کابلی را یکپارچه کرده، و باعث رقابت بین شرکتهای مختلف (برای سرویس بهتر و قیمت کمتر) شود. این قانون از فوریه ۱۹۹۶ به اجرا گذاشته شد، و باعث شد تا تعدادی از BOC ها وارد حوزه IXC شوند، و از طرف دیگر شرکتهای تلویزیون کابلی نیز به ارائه سرویسهای تلفن بپردازند (و با LEC ها رقابت کنند).

یکی از نکات جالب قانون ۱۹۹۶ آنست که LEC ها بایستی شماره های تلفن را بگونه ای تنظیم کنند که قابل انتقال باشند. این بدان معناست که یک مشترک می تواند از این منطقه به منطقه دیگر برود، بدون آنکه نیازی باشد شماره تلفنش را عوض کند. این ویژگی باعث می شود تا افراد آزادی بیشتری در عوض کردن LEC خود داشته باشند، و در نتیجه تنوع رقابت داغتر شود. با قانون ۱۹۹۶، مخابرات ایالات متحده دستخوش تغییرات ساختاری شدیدی شده است. کشورهای بسیاری نیز ترغیب شده اند تا دست به چنین اقداماتی بزنند. البته در اغلب

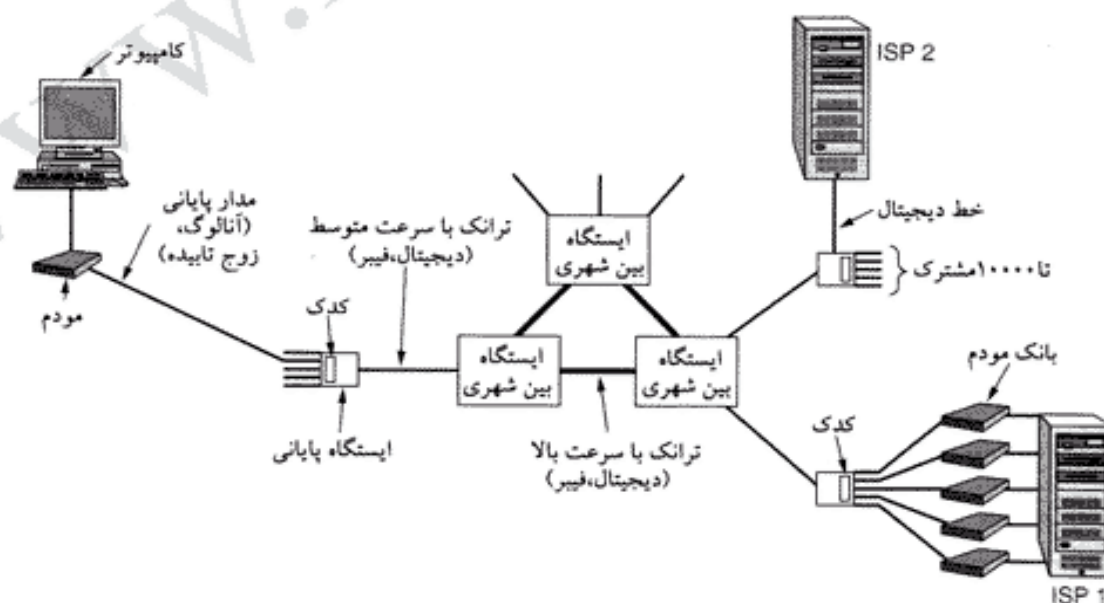
موارد کشورهای دیگر صبر می کنند تا نتیجه کار در ایالات متحده مشخص شود: اگر نتیجه مثبت بود، همان راه را می روند؛ اگر نتیجه منفی بود، راه دیگری را امتحان می کنند.

۳-۵-۲ مدارهای پایانی: مودم، ADSL، و بیسیم

اکنون زمان آن است که تا به بررسی نحوه کار سیستم تلفن پردازیم. بخشهای اصلی این سیستم در شکل ۲-۲۳ نشان داده شده است. در این شکل مدارهای پایانی، ترانکها، ایستگاههای بین شهری و ایستگاههای پایانی را می بینید. هر ایستگاه پایانی (در ایالات متحده و کشورهای بزرگ) تا ۱۰,۰۰۰ مدار پایانی دارد. در حقیقت تا همین اواخر، کد منطقه + شماره ناحیه مشخص کننده ایستگاه پایانی بود - برای مثال، 601-xxxx (212) یک ایستگاه پایانی با ۱۰,۰۰۰ مشترک (از 0000 تا 9999) است. با ایجاد شدن امکان رقابت بر سر سرویسهای محلی دیگر این وضعیت عملی نیست، چون شرکتهای زیادی طالب بدست آوردن کدهای محلی هستند. روشهای شماره گذاری نیز بایستی بکلی عوض شود، چون اغلب کدهای منطقه ای مصرف شده اند.

اجازه دهید با بخشی که بیشتر مردم با آن آشنا نیستند، شروع کنیم: دو رشته سیمی که از ایستگاه پایانی شرکت تلفن به محل مشترک (خانه یا محل کار) کشیده می شود. به این دو رشته سیم اغلب «کیلومتر آخر» گفته می شود، اگرچه طول آن ممکنست به چندین کیلومتر هم برسد. در طول یکصد سال گذشته در این بخش از سیگنالهای آنالوگ استفاده شده است، و (بدلیل هزینه زیاد تجهیزات دیجیتال) احتمالاً در چند سال آینده نیز وضع به همین منوال خواهد بود. با این حال، در این آخرین سنگز آنالوگ نیز تغییرات شروع شده است. در این قسمت مدارهای پایانی و آخرین تحولات آنرا (با تأکید بر مخابرات داده بین کامپیوترها) بتفصیل بررسی خواهیم کرد.

وقتی یک کامپیوتر می خواهد داده های دیجیتال را روی یک خط تلفن آنالوگ ارسال کند، ابتدا باید آنها را به سیگنالهای آنالوگ تبدیل کند - کاری که با دستگاهی بنام مودم (modem) انجام می شود. در ایستگاه پایانی این اطلاعات مجدداً به سیگنالهای دیجیتال تبدیل شده، و برای ارسال روی ترانک راه دور (trunk) آماده می شود. اگر مقصد اطلاعات یک کامپیوتر باشد، سیگنال مجدداً به آنالوگ تبدیل می شود تا بتوان آنرا روی مدار پایانی



شکل ۲-۲۳. ترکیبی از انتقال دیجیتال و آنالوگ برای تماس کامپیوتر با کامپیوتر. تبدیل سیگنال بوسیله مودمها و کدکها انجام می شود.

آن ارسال کرد؛ و در مقصد یک مودم دیگر اطلاعات را از آنالوگ به دیجیتال تبدیل می‌کند. در شکل ۲-۲۳، ISP 1 (ارائه‌دهنده سرویسهای اینترنتی) یک بانک مودم دارد، که هر کدام از آنها به یک خط تلفن (مدار پایانی) مستقل متصلند. این ISP می‌تواند در آن واحد به تعداد مودمهای خود به افراد مختلف سرویس بدهد (البته با این فرض که کامپیوترهایش به اندازه کافی قوی باشند). این آرایش تا وقتی که مودمهای 56-kbps وارد بازار شدند، آرایش متداولی بود (بزودی علت آنرا خواهید فهمید).

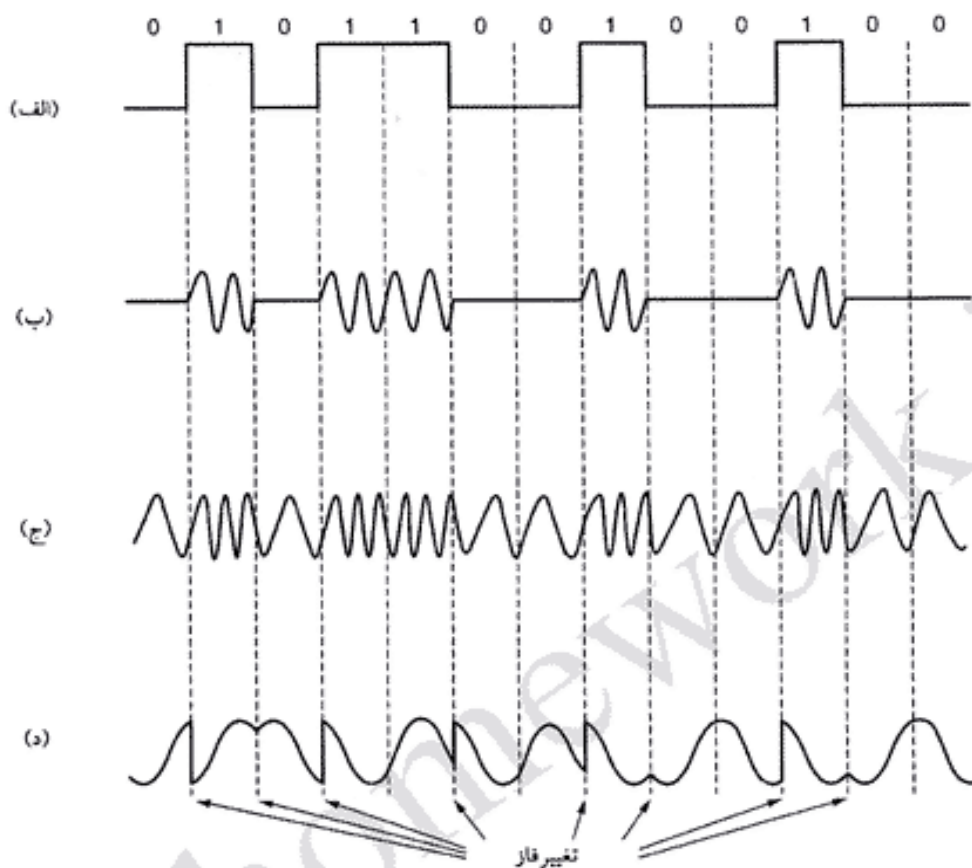
در مخابرات آنالوگ برای انتقال اطلاعات از ولتاژی که با زمان تغییر می‌کند، استفاده می‌شود. اگر رسانه انتقال کامل و بدون نقص باشد، گیرنده دقیقاً همان سیگنالی را دریافت خواهد کرد که فرستنده ارسال کرده است. اما متأسفانه چنین نیست، و گیرنده همان سیگنال ارسالی را دریافت نمی‌کند - و در مخابرات دیجیتال این یعنی خطا. انتقال سیگنالهای الکتریکی روی خطوط انتقال با سه مشکل عمده روبروست: تضعیف سیگنال، اعوجاج تأخیری، و نویز. وقتی یک سیگنال در رسانه انتقال منتشر می‌شود، انرژی خود را از دست می‌دهد که به آن تضعیف (attenuation) می‌گویند، و بر حسب دسی بل بر کیلومتر اندازه‌گیری می‌شود. میزان تضعیف یک سیگنال به فرکانس آن بستگی دارد. شاید فکر کنید که این وابستگی به فرکانس اشکال زیادی ایجاد نمی‌کند، ولی وقتی یک موج را بصورت مجموعه‌ای از مؤلفه‌های فوریه در نظر بگیرید، تأثیر وابستگی به فرکانس خود را نشان خواهد داد. در واقع هر یک از مؤلفه‌های فوریه بگونه‌ای متفاوت تضعیف می‌شوند، و ترکیب مجدد آنها در گیرنده موج کاملاً متفاوتی ایجاد می‌کند.

اما وضع از این هم بدتر است، چون مؤلفه‌های فوریه با سرعتهای متفاوتی در رسانه انتقال (سیم) منتشر می‌شوند. این تفاوت سرعتها باعث اعوجاج تأخیری (delay distortion) سیگنال در گیرنده می‌شود. مشکل دیگر نویز (noise) - انرژی ناخواسته از منابعی غیر از فرستنده - است. نویز حرارتی حاصل حرکات تصادفی الکترونها در سیم است، و بکلی نمی‌توان از آن اجتناب کرد. نویز القایی نیز حاصل القای ولتاژ در اثر عبور جریان از سیمهای مجاور است. گاهی پیش آمده که وقتی تلفنی با کسی صحبت می‌کنید، مکالمه دیگری را نیز در زمینه می‌شنوید؛ علت این پدیده (که به هم‌شنوایی - crosstalk - مشهور است) نویز القایی می‌باشد. در اثر قطع و وصل خطوط قدرت نویز دیگری بنام نویز ضربه‌ای روی خطوط مخابراتی القا می‌شود، که می‌تواند چندین بیت از اطلاعات را از بین ببرد.

مودم

بدلیل مشکلاتی که در بالا گفته شد (بویژه وابستگی تضعیف سیگنال و سرعت انتشار آن به فرکانس)، سعی می‌شود از سیگنالهایی با محدوده فرکانسی پائین استفاده شود. متأسفانه، شکل موج مربعی سیگنالهای دیجیتال دارای طیف فرکانسی وسیعی است، و بشدت در معرض تضعیف و اعوجاج تأخیری قرار دارد. این تأثیرات باعث شده تا سیگنالهای بیس باند (DC) فقط برای سرعتهای پائین و مسافتهای کوتاه مناسب باشد.

برای حل مشکل سیگنالهای DC بویژه در خطوط تلفن، از سیگنال AC استفاده می‌شود. در این جا یک تون ۱۰۰۰ تا ۲۰۰۰ هرتزی بعنوان موج حامل سینوسی (sine wave carrier) بکار برده می‌شود. برای انتقال اطلاعات می‌توان دامنه، فرکانس یا فاز این موج حامل را مدوله کرد. در مدولاسیون دامنه (amplitude modulation) از دو دامنه متفاوت بعنوان 0 و 1 استفاده می‌شود. در مدولاسیون فرکانس (frequency modulation) - که به کُدگذاری با شیفت فرکانس (frequency shift keying) نیز معروف است - از دو تون متفاوت برای 0 و 1 استفاده می‌شود. (در صنعت مخابرات، اصطلاحات مدولاسیون و کُدگذاری معادل یکدیگرند.) در مدولاسیون فاز (phase modulation) موج حامل در فواصل یکنواخت 0 یا 180 درجه شیفت پیدا می‌کند. با استفاده از شیفتهای 45، 135، 225 یا 315 درجه‌ای می‌توان در هر فاصله زمانی بجای یک بیت، ۲



شکل ۲-۲۴. (الف) سیگنال باینری. (ب) مدولاسیون دامنه. (ج) مدولاسیون فرکانس. (د) مدولاسیون فاز.

بیت اطلاعات منتقل کرد. همچنین، وجود تغییر فاز در انتهای هر فاصله زمانی تشخیص مرزهای آنها را برای گیرنده آسانتر می‌کند.

شکل ۲-۲۴ این سه نوع مدولاسیون را نشان می‌دهد. در شکل ۲-۲۴ (ب) یکی از دامنه‌ها صفر و دیگری غیرصفر است. در شکل ۲-۲۴ (ج) از دو فرکانس متفاوت برای نمایش 0 و 1 استفاده شده است. در شکل ۲-۲۴ (د) در هر فاصله زمانی وجود (یا عدم وجود) تغییر فاز نشان‌دهنده تغییر مقدار بیت (یا عدم تغییر آن) است.

دستگاهی که جریان بیت‌ها را بعنوان ورودی گرفته، و با ایجاد یک موج حامل و اعمال یکی از انواع مدولاسیون (یا ترکیبی از آنها) یک خروجی آنالوگ تولید می‌کند (و یا بر عکس، با گرفتن موج آنالوگ اطلاعات دیجیتال را از آن استخراج می‌کند)، مودم (modulator-demodulator) نامیده می‌شود. مودم بین کامپیوتر (دیجیتال) و سیستم تلفن (آنالوگ) قرار می‌گیرد.

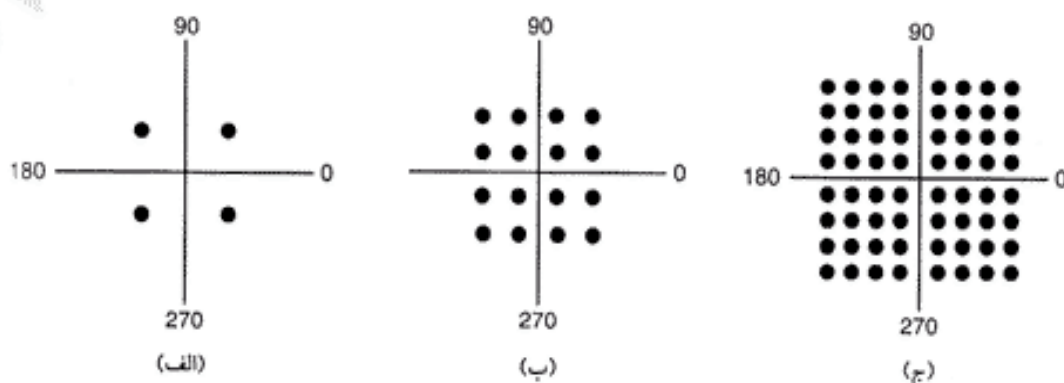
بالا بردن سرعت بسادگی و فقط با زیاد کردن نرخ نمونه‌برداری (sampling rate) ممکن نیست. قضیه ناپکوئیست می‌گوید که برای یک خط کامل 3000 Hz (که خطوط تلفن مسلماً چنین نیستند)، حداکثر نرخ نمونه‌برداری 6000 Hz است. در عمل، اکثر مودمها با نرخ 2400 times/sec نمونه‌برداری می‌کنند، ولی سعی می‌کنند در هر نمونه‌برداری بیت‌های بیشتر را بخوانند.

به تعداد نمونه ها در ثانیه باد (baud) گفته می شود، و در هر باد یک سمبل (symbol) فرستاده می شود. بنابراین، یک خط n -baud در هر ثانیه n سمبل ارسال می کند (برای مثال، یک خط 2400-baud در هر 416.667 μ sec یک سمبل می فرستد). اگر این سمبل فقط حاوی ولتاژ 0 v (برای نمایش 0 منطقی) یا 1 v (برای نمایش 1 منطقی) باشد، سرعت مودم 2400 bps خواهد بود. اما اگر از ولتاژهای 0، 1، 2، و 3 ولت استفاده کنیم، هر سمبل می تواند حاوی 2 بیت باشد، و سرعت انتقال اطلاعات مودم به 4800 bps می رسد. در مدولاسیون فاز چهار درجه ای نیز می توان در هر سمبل 2 بیت ارسال کرد، و بدین ترتیب سرعت ارسال داده دو برابر سرعت باد خواهد بود. این تکنیک که کاربرد زیادی گسترده ای نیز دارد، QPSK (گذگذاری با شیفت فاز چهارگانه - Quadrature Phase Shift Keying) نامیده می شود.

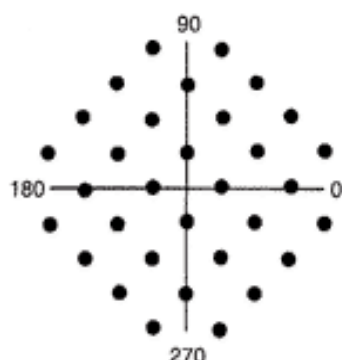
مفاهیم پهنای باند، باد، سمبل، و نرخ بیت بسیار با هم اشتباه می شوند، بنابراین اجازه دهید یک بار دیگر آنها را بیان کنیم. پهنای باند یک رسانه محدوده فرکانسی است که می تواند سیگنال را با کمترین تضعیف عبور دهد. این یکی از ویژگیهای فیزیکی رسانه انتقال است، و با هرتز (Hz) سنجیده می شود. به تعداد نمونه برداری در هر ثانیه باد گفته می شود، و هر نمونه حاوی یک قطعه از اطلاعات (یا سمبل) است. بنابراین، نرخ باد (baud rate) و نرخ سمبل (symbol rate) در واقع یکی هستند. تعداد بیت بر سمبل توسط نوع مدولاسیون (مثلاً، QPSK) تعیین می شود. نرخ بیت (bit rate) مقدار اطلاعاتیست که روی یک کانال فرستاده می شود، و برابر است با نرخ سمبل (symbol/sec) \times تعداد بیت در هر سمبل (bits/symbol).

تمام مودمهای پیشرفته برای ارسال بیشترین بیتهای ممکن در هر باد، از مدولاسیونهای ترکیبی (چند دامنه ای و چند فازی) استفاده می کنند. در شکل ۲-۲۵ (الف) نقاطی را می بینید که در فواصل یکسان از مبدأ مختصات (دامنه یکسان)، و با زاویه ها (فازها) 45، 135، 225 و 315 درجه قرار گرفته اند (فاز هر نقطه زاویه ایست که با جهت مثبت محور x می سازد). در شکل ۲-۲۵ (الف) چهار ترکیب معتبر وجود دارد، که بدین ترتیب می توان در هر سمبل 2 بیت ارسال کرد - این همان QPSK است.

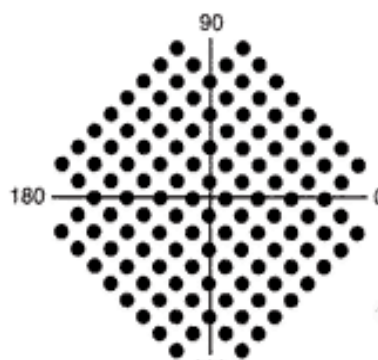
در شکل ۲-۲۵ (ب) مدولاسیون دیگری را می بینید، که در آن از چهار دامنه و چهار فاز مختلف استفاده شده و 16 ترکیب معتبر بدست می دهد. با این مدولاسیون می توان در هر سمبل چهار بیت ارسال کرد، و به آن QAM-16 (مدولاسیون دامنه چهارگانه - Quadrature Amplitude Modulation) گفته می شود (گاهی نیز 16-QAM خوانده می شود). با مدولاسیون QAM-16 می توان روی یک خط 2400-baud تا سرعت 9600 bps داده ارسال کرد.



شکل ۲-۲۵. (الف) مدولاسیون QPSK. (ب) مدولاسیون QAM-16. (ج) مدولاسیون QAM-64.



(ب)



(ج)

شکل ۲-۲۶. (الف) استاندارد V.32 برای 9600 bps. (ب) استاندارد V.32 bis برای 14,400 bps.

در شکل ۲-۲۵ (د) مدولاسیون دیگری را می بینید که شامل دامنه های بیشتری است. در این طرح ۶۴ ترکیب ممکنه وجود دارد، که بدین ترتیب می توان به ازای هر سمبل ۶ بیت را ارسال کرد. این مدولاسیون QAM-64 نام دارد (از مدولاسیون QAM با مراتب بالاتر نیز استفاده می شود).

دیاگرامهایی مانند شکل ۲-۲۵ که ترکیبات ممکنه دامنه و فاز را نشان می دهند، به دیاگرام فلکی (constellation diagram) معروفند. هر استاندارد مودمهای سرعت-بالا دارای دیاگرام فلکی خاص خود است، و فقط قادر به ارتباط با مودمهاییست که دارای دیاگرام مشابهی باشند (البته این مودمها می توانند تمام سرعتهای پائینتر را شبیه سازی کنند).

با بالا رفتن تعداد نقاط در دیاگرام فلکی حتی نویزهای کوچک نیز می توانند باعث بروز خطا در آشکارسازی دامنه یا فاز سیگنال شوند، که بدنبال آن بینهای زیادی از دست می رود. برای کاهش احتمال خطا، در مودمهای سرعت-بالا با اضافه کردن بیت های اضافی نوعی تصحیح خطا (error correction) انجام می شود. به این روش TCM (مدولاسیون کُدگذاری تاروپودی - Trellis Coded Modulation) می گویند. برای مثال، استاندارد V.32 از ۳۲ نقطه فلکی برای ارسال ۴ بیت داده و یک بیت برابری (parity) در هر سمبل استفاده کرده، و با نرخ 2400-baud به سرعت 9600 bps (با تصحیح خطا) دست می یابد. دیاگرام فلکی این استاندارد را در شکل ۲-۲۶ (الف) مشاهده می کنید. (چرخش ۴۵ درجه ای حول محور مختصات فقط بدلائیل مهندسی اتخاذ شده است، و هیچگونه تأثیری روی ظرفیت اطلاعات ندارد).

بعد از 9600 bps قدم بعدی 14,400 bps است، که استاندارد آن V.32 bis نامیده می شود. برای رسیدن به این سرعت باید (با نرخ 2400-baud) در هر سمبل ۶ بیت داده و ۱ بیت برابری ارسال شود. دیاگرام فلکی این مودم را، که (وقتی از QAM-128 استفاده شود) ۱۲۸ نقطه دارد، در شکل ۲-۲۶ (ب) می بینید. مودمهایی که قابلیت فکس دارند، برای ارسال فکس از این استاندارد استفاده می کنند. مدولاسیون QAM-256 در هیچ یک از مودمهای تلفنی استفاده نمی شود، ولی در شبکه های کابلی کاربرد دارد (بعداً آنرا خواهید دید).

استاندارد مودم تلفنی بعدی V.34 است، که با سرعت 28,800 bps (12 data bits/symbol در 2400-baud) کار می کند. آخرین مودم در این سری استاندارد V.34 bis نام دارد، که به سرعت 33,600 bps (14 data bits/symbol در 2400-baud) دست پیدا می کند.

برای رسیدن به سرعتهای بیشتر، بسیاری از مودمها از تکنیکهای فشرده سازی استفاده می کنند تا به سرعتهای

بالتر از 33,600 bps برسند. از سوی دیگر، تقریباً تمام مودمها قبل از شروع ارسال داده ها کیفیت خط را چک می کنند، و اگر نقصی در کیفیت خط وجود داشته باشد، سرعت خود را آنقدر پائین می آورند تا ارسال مطمئن داده ها امکان پذیر باشد. بهمین دلیل، سرعت مؤثر یک مودم می تواند کمتر، مساوی، و یا بیشتر از سرعت رسمی آن باشد.

تمام مودمهای جدید (با استفاده از فرکانسهای متفاوت برای ارسال و دریافت) اجازه می دهند تا ارسال و دریافت همزمان انجام شود. به چنین ارتباطی دو-طرفه همزمان (full duplex) گفته می شود. مانند یک اتوبان دو-سبانه. اگر در هر لحظه فقط از یک طرف ارتباط ممکن باشد، به آن دو-طرفه ناهمزمان (half duplex) گفته می شود (مانند راه آهن های یک-خطه). و اگر ارتباط فقط در یک جهت مجاز باشد، به آن یکطرفه (simplex) می گویند (مانند یک خیابان یکطرفه). یک رشته فیبر نوری که در یک سمت فقط دیود لیزری و در سمت دیگر فقط آشکارساز نوری دارد، نیز سیستمی یکطرفه است.

همانطور که قبلاً گفتیم، طبق قانون شانون سرعت انتقال روی خطوط تلفن از 35-kbps نمی تواند فراتر رود، بهمین مودمهای استاندارد از سرعت 33,600 bps بالاتر نمی روند. شاید بپرسید پس مودمهای 56-kbps چگونه کار می کنند؟ کمی صبر کنید، به آن هم خواهیم رسید.

اما این حد 35-kbps از کجا آمده است؟ این محدودیت به طول متوسط مدارهای پایانی و کیفیت آنها بستگی دارد. به شکل ۲-۲۳ نگاه کنید: ارتباطی که از کامپیوتر سمت چپ شروع شده و به 1 ISP ختم شود، از دو مدار پایانی آنالوگ (یکی در مبدأ و دیگری در مقصد) عبور می کند. هر یک از این حلقه ها میزان نویز سیگنال را بالا می برند. اگر بتوانیم یکی از این حلقه ها را حذف کنیم، می توانیم حداکثر سرعت را به دو برابر برسانیم.

این دقیقاً همان کاریست که 2 ISP انجام داده: این ISP ارتباط خود با نزدیکترین ایستگاه پایانی را بصورت کاملاً دیجیتال در آورده است. سیگنال دیجیتال ترانک مستقیماً به 2 ISP فرستاده شده، و مودمها و کودکهای آنالوگ بکلی حذف شده اند. بدین ترتیب، حداکثر سرعت ارتباط با این ISP می تواند به 70-kbps نیز برسد. بین دو نقطه که از مودم و خطوط آنالوگ استفاده می کنند، حداکثر سرعت همان 33.6 kbps است.

اما علت اینکه مودمهای 56-kbps می توانند در عمل کار کنند، به قضیه نایکوئیست برمی گردد. پهنای باند کانالهای تلفنی (بانهضام باند محافظ) 4000 Hz است، و طبق قضیه نایکوئیست حداکثر نرخ نمونه برداری در چنین کانالی 8000 خواهد بود. در ایالات متحده آمریکا تعداد بیت های هر نمونه ۸ است، که (با احتساب یک بیت کنترلی) حداکثر سرعت مجاز به 56,000 bits/sec می رسد. در اروپا از بیت کنترلی استفاده نمی شود، و می توان تمام بیتها را به داده ها اختصاص داد، بنابراین حداکثر سرعت در آنجا می تواند تا 64,000 bits/sec افزایش یابد، ولی طبق یک توافق بین المللی همان سرعت 56,000 انتخاب شده است.

این استاندارد V.90 نامیده می شود. در این استاندارد سرعت ارسال کاربر به ISP همان 33.6 kbps است، ولی سرعت دریافت از ISP به 56-kbps می رسد (چون معمولاً آنچه که کاربر از ISP می گیرد، بسیار بیشتر از آن چیزیست که به آن می فرستند). از لحاظ نظری امکان رساندن کانال ارسال به ISP تا 56-kbps نیز وجود داشت، ولی (بعلم نویزی بودن خطوط تلفنی) تصمیم گرفته شد تا این کانال به 33.6 kbps محدود شده، و پهنای باند آن به کانال دریافت از ISP داده شود تا احتمال رسیدن آن به سرعت 56-kbps افزایش یابد.

قدم بعدی در این بازی استانداردها، V.92 است. مودمهای این استاندارد می توانند تا 48-kbps روی کانال ارسال داده بفرستند (البته مشروط باینکه خط تلفن توانایی آنرا داشته باشد). زمان شناسایی کیفیت خط و تعیین سرعت مناسب در این استاندارد نیز بسیار کمتر از مودمهای دیگر است (تقریباً نصف ۳۰ ثانیه ای که مودمهای دیگر صرف این کار می کنند). و بالاخره، مودمهای استاندارد V.92 اجازه می دهند تا یک تماس تلفنی بتواند تماس اینترنتی را قطع کند، مشروط باینکه خط دارای سرویس انتظار مکالمه (call waiting) باشد.

خط مشترک دیجیتال (DSL)

وقتی شرکتهای تلفن بالاخره موفق شدند سرویس 56-kbps ارائه کنند، خیلی به خود غره شدند؛ اما در همان حال، شرکتهای تلویزیون کابلی ارتباطاتی با سرعت 10 Mbps (و ماهواره‌ها سرعت ارسال 50 Mbps) عرضه می‌کردند. با داغ شدن بازار اینترنت، شرکتهای تلفن (LEC) دریافته‌اند که برای رقابت به محصول جدیدی نیاز دارند. پاسخ آنها به این وضعیت ارائه سرویسهای دیجیتال روی مدارهای پایانی بود. این سرویسها پهنای باند بیشتری داشتند و به آنها سرویس باند-وسیع (broadband) گفته می‌شد (اگرچه این نامگذاری بیشتر جنبه تبلیغاتی داشت، تا فنی).

در ابتدا این سرویسها بسیار متنوع بودند، و تحت نام xDSL (خط مشترک دیجیتال: Digital Subscriber Line) - که در آن x نوع سرویس را مشخص می‌کند) دسته بندی می‌شدند. مهمترین این سرویسها - که عامل اصلی موفقیت آن هم بود - ADSL (DSL نامتقارن - Asymmetric DSL) نام دارد، که در زیر با آن آشنا خواهید شد. از آنجائیکه ADSL هنوز در حال توسعه و تکامل است و استانداردهای آن هنوز بطور کامل تدوین نشده، ممکنست در آینده تغییراتی در آن رخ دهد، ولی تصویر کلی همان است که خواهید دید. برای اطلاعات بیشتر درباره ADSL به (Summers, 1999; Vetter et al., 2000) نگاه کنید.

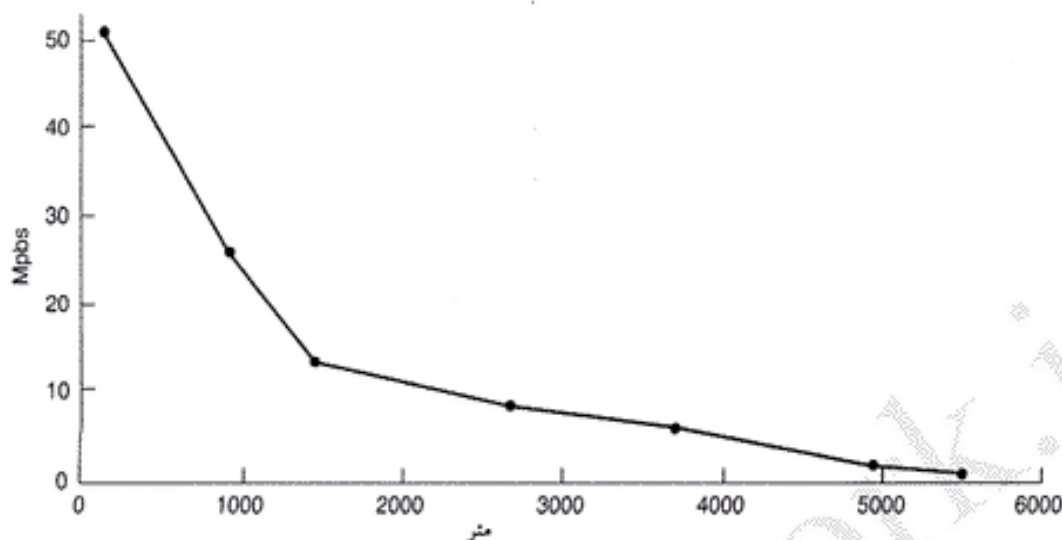
علت آن که مودمها اینقدر گد هستند اینست که، شبکه تلفن اساساً برای انتقال صدای انسان طراحی و توسعه داده شده، و سرویسهای داده فرزندخوانده آن محسوب می‌شود. در نقطه‌ای که ما، پایانی وارد ایستگاه پایانی تلفن می‌شود، فیلترهایی قرار داده شده که تمام فرکانسهای زیر 300 Hz و بالای 3400 Hz را تضعیف می‌کنند. البته قطع فرکانس بیکباره صورت نمی‌گیرد - از لحاظ فنی، 300 Hz و 3400 Hz نقاط 3 dB هستند - بهمین دلیل پهنای باند را معمولاً 4000 Hz فرض می‌کنند، اگرچه فاصله نقاط 3 dB فقط 3100 Hz است. داده نیز به همین باند باریک محدود است.

برای اجتناب از این وضعیت در سرویس xDSL، خط مشترک بدون عبور از فیلترهای مزبور مستقیماً به نوع خاصی از سونچها متصل می‌شود، تا بتواند از تمام ظرفیت مدار پایانی استفاده کند. در این حالت دیگر محدودیت پهنای باند فقط به خواص فیزیکی مدار پایانی بستگی دارد، نه به محدوده‌ای که فیلترها بطور مصنوعی برای آن بوجود آورده‌اند.

متأسفانه، ظرفیت مدار پایانی نامحدود نیست، و به عواملی از قبیل طول خط، ضخامت سیم، و کیفیت کلی آن بستگی دارد. در شکل ۲-۲۷ نمودار پهنای باند بر حسب مسافت را ملاحظه می‌کنید - در اینجا فرض شده که سایر عوامل بهینه هستند (سیمهای نو، کابلهای نه چندان قطور، و غیره).

نمودار فوق مشکل اصلی شرکتهای تلفن را بخوبی نشان می‌دهد: شعاع ارائه این سرویس به مشترکان بشدت محدود است. این بدان معناست که وقتی کاربری که خارج از این شعاع زندگی می‌کند، برای دریافت سرویس xDSL مراجعه کند، باید با کمال تأسف از او عذرخواهی کنیم که امکان ارائه سرویس به وی را نداریم. برای بیشتر کردن شعاع سرویس، باید سرعت آنرا پائین بیاوریم، ولی پائین آوردن سرعت همان و از دست دادن جذابیت همان. اینجاست که تکنولوژی مغلوب اقتصاد می‌شود. (راه حل دیگر آنست که ایستگاههای کوچک و پراکنده‌ای در نقاط نزدیک به هم تأسیس کنیم، که البته این هم اقتصادی نیست).

سرویس xDSL با اهداف مشخصی طراحی شده است: اول اینکه، این سرویسها باید بتوانند با خطوط زوج تاییده Cat 3 کار کنند؛ دوم اینکه، این سرویسها نباید هیچ اختلالی در دستگاههای تلفن و فکس معمولی بوجود آورند؛ سوم اینکه، باید از 56-kbps سریعتر باشند؛ و چهارم اینکه، این سرویسها باید دائماً برقرار باشند، و هزینه آنها هم ثابت (و مثلاً ماهیانه) باشد - نه مانند تلفنهای معمولی، بصورت دقیقه‌ای.

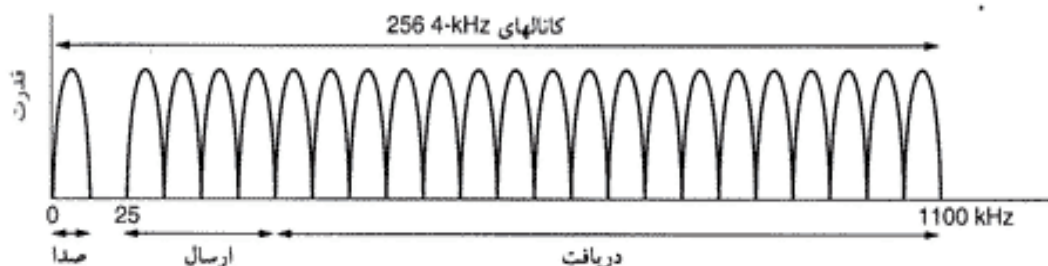


شکل ۲-۲۷. نمودار پهنای باند سرویس DSL بر حسب طول سیم در کابل Cat 3 UTP.

اولین سرویس ADSL توسط AT&T ارائه شد، که در آن پهنای باند موجود در مدار پایانی (که تقریباً 1.1 MHz است) به سه باند تقسیم شده بود: باند POTS (سرویس تلفن معمولی - Plain Old Telephone Service)، باند ارسال از کاربر (به ایستگاه پایانی)، و باند ارسال به کاربر (از ایستگاه پایانی). تکنیک تقسیم پهنای باند به فرکانسهای مختلف مالتی پلکس تقسیم فرکانس (frequency division multiplexing) نام دارد، که بعداً درباره آن مفصلاً صحبت خواهیم کرد. شرکتهای بعدی برای ارائه این سرویس از تکنیک متفاوتی استفاده کردند، و از آنجائیکه احتمالاً این روش غالب خواهد شد، ابتدا آنرا توضیح می دهیم.

این روش که DMT (تون چندگانه گسسته - Discrete MultiTone) نامیده می شود، در شکل ۲-۲۸ نشان داده شده است. در حقیقت، کاری که در اینجا انجام شده تقسیم پهنای باند موجود (1.1 MHz) به 256 کانال مستقل 4312.5 Hz است. از کانال 0 بعنوان POTS (سرویس تلفن معمولی) استفاده می شود. کانالهای 1-5 خالی رها شده اند، تا تداخلی بین صدا و داده پیش نیاید. از 250 کانال باقیمانده، یکی برای کنترل ارسال از کاربر و یکی برای کنترل ارسال به کاربر تخصیص یافته، و از بقیه کانالها می توان برای داده استفاده کرد.

در تئوری می توان از هر یک از این کانالها برای ارتباط دو-طرفه همزمان استفاده کرد، ولی هارمونیکها، همشنوایی و اثرات دیگر باعث می شود تا در عمل ظرفیت سیستم بسیار کمتر باشد. این ارائه دهنده سرویس است که تعیین می کند چند کانال برای ارسال از کاربر اختصاص یافته، و چند کانال برای ارسال به کاربر. از نظر تکنیکی می توان این نسبت را بصورت 50-50 تعریف کرد، ولی از آنجائیکه اغلب کاربران اطلاعات دریافتی خیلی بیشتری دارند، ۸۰ الی ۹۰ درصد پهنای باند به دریافت کاربر اختصاص داده می شود. از اینجا است که حرف "A" (بمعنای



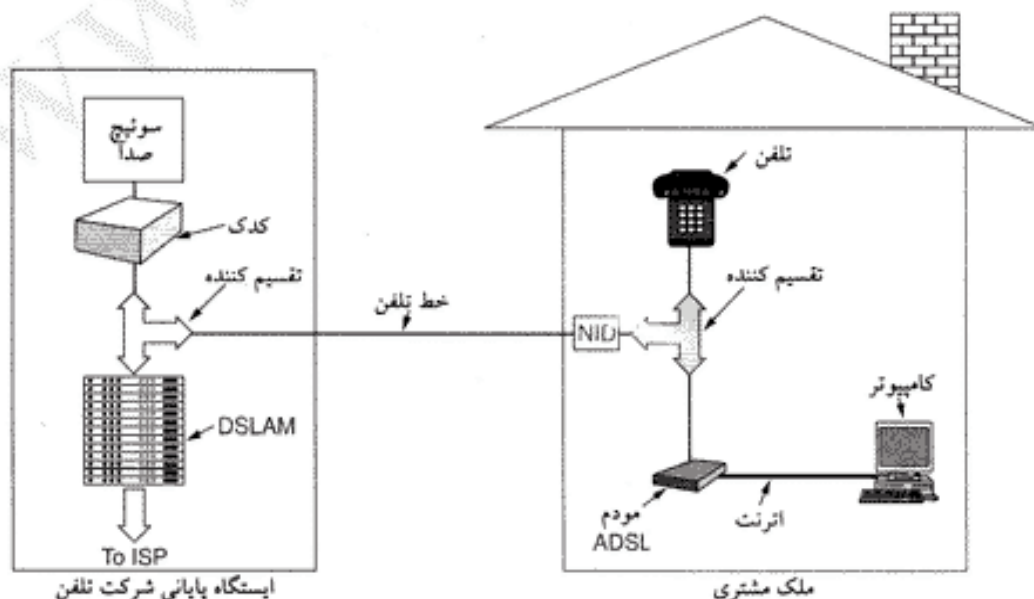
شکل ۲-۲۸. عملکرد ADSL با مدولاسیون تون چندگانه گسسته.

"نامتقارن" در ADSL ظاهر می‌شود. در اغلب موارد ۳۲ کانال به ارسال کاربر، و بقیه به دریافت آن اختصاص می‌یابد. به‌منظور افزایش پهنای باند، امکان اختصاص چند کانال فوقانی به ارتباط دوطرفه نیز وجود دارد، ولی این روش به تجهیزات اضافی برای خنثی کردن پژواک (echo) در خط نیاز دارد.

در استاندارد ADSL (ANSI T1.413 و ITU G.992.1) تا 8 Mbps برای دریافت کاربر و تا 1 Mbps برای ارسال کاربر مجاز است. اما، کمتر شرکتی پیدا می‌کنید که چنین سرویسهایی ارائه کند. در عمل، این سرویسهها معمولاً بصورت 512 kbps دریافت و 64 kbps ارسال (سرویس معمولی)، و 1 Mbps دریافت و 256 kbps ارسال (سرویس ویژه) ارائه می‌شوند.

در هر کانال از مدولاسیونی شبیه V.34 (با نرخ نمونه‌برداری 4000-baud بجای 2400-baud) استفاده می‌شود. کیفیت خط در هر کانال مستقلاً (و بصورت پیوسته) ارزیابی شده، و در صورت نیاز سرعت مجدداً تنظیم می‌شود، بنابراین سرعت کانالها می‌تواند کاملاً متفاوت باشد. برای ارسال داده‌ها از مدولاسیون QAM با نرخ 15 bits/ baud (شبهه شکل ۲-۲۵ ب) استفاده می‌شود. برای مثال، اگر 224 کانال دریافت با نرخ 15 bits/ baud در 4000-baud داشته باشیم، پهنای باند دریافتی معادل 13.44 Mbps خواهیم داشت. در عمل، نرخ سیگنال سه-نویز هرگز آنقدر خوب نیست که اجازه چنین سرعت‌های را بدهد، ولی در مسافت‌های کوتاه و با کابل‌کشی مناسب می‌توان به سرعت 8 Mbps رسید، که همین برای استقبال عامه کافیست.

در شکل ۲-۲۹ یک طرح ADSL را مشاهده می‌کنید. در این طرح، تکنسین شرکت تلفن بایستی یک NID (دستگاه واسط شبکه - Network Interface Device) در محل سکونت مشتری نصب کند. این جعبه پلاستیکی کوچک انتهای مایملک شرکت تلفن و شروع مایملک مشتری را مشخص می‌کند. کنار NID (یا حتی گاهی بصورت ترکیبی با آن) یک تقسیم‌کننده (splitter - فیلتر آنالوگی که باند POTS را از باند داده جدا می‌کند) نصب می‌شود. سیگنال POTS به دستگاه تلفن، و سیگنال داده به یک مودم ADSL داده می‌شود. مودم ADSL در واقع یک DSP (پردازشگر سیگنال دیجیتال - Digital Signal Processor) است، که بگونه‌ای تنظیم شده تا بعنوان مودم QAM موازی (با فرکانسهای مختلف) کار کند. از آنجائیکه اکثر مودمهای ADSL بصورت خارج از



شکل ۲-۲۹. یک طرح ساده ADSL.

کامپیوتر هستند، ارتباط آنها با کامپیوتر نیز باید از نوع پرسرعت باشد. برای این منظور از پورت های اترنت یا USB استفاده می شود. بدون شک در آینده شاهد به بازار آمدن مودم های داخلی ADSL نیز خواهیم بود. در انتهای دیگر خط (در ایستگاه پایانی)، یک تقسیم کننده متناظر قرار می گیرد. در اینجا، سیگنال صدا (0-4000 Hz) از داده جدا شده و به سونچ های معمولی تلفن فرستاده می شود. سیگنال بالای 26 kHz نیز جدا شده، و به دستگاهی بنام DSLAM (مالتی پلکسر دسترسی DSL - DSL Access Multiplexer) که بسیار شبیه مودم ADSL است، می رود. پس از تبدیل این سیگنال به جریان بیت های دیجیتال، از آنجا به ISP فرستاده می شود. جداسازی کامل سیستم صدا از ADSL، پیاده سازی آنرا برای شرکت های تلفن بسیار ساده کرده است: تمام کاری که باید انجام دهد اینست که در سمت خود یک تقسیم کننده و DSLAM، و در سمت مشتری یک تقسیم کننده ساده نصب کند. در سیستم های پرسرعت دیگر (مانند، ISDN) تجهیزات بسیار پیچیده تری باید نصب شود.

یکی از معایب طرح شکل ۲-۲۹ وجود تجهیزاتی است که باید در محل سکونت مشترک نصب شود (NID و تقسیم کننده). این کار مستلزم آنست که یک تکنسین به آنجا مراجعه کرده، و وسایل را نصب کند. به همین دلیل استاندارد جدیدی که به تقسیم کننده نیازی ندارد، توسعه داده شده است. نام غیررسمی این استاندارد G.lite است، ولی رسماً به آن ITU G.992.2 گفته می شود. این طرح شبیه شکل ۲-۲۹ است، که فقط تقسیم کننده حذف شده، و از خط تلفن بهمان صورت موجود استفاده می شود. تنها تفاوت اینست که باید میکروفیلترهایی را بین پریز تلفن و تجهیزات (دستگاه تلفن و مودم ADSL) قرار داد. میکروفیلتر تلفن یک فیلتر پائین گذر (low-pass) است که فرکانسهای بالای 3400 Hz را حذف می کند؛ میکروفیلتر مودم ADSL یک فیلتر بالاگذر (high-pass) است که فرکانسهای زیر 26 kHz را حذف می کند. با این حال کارایی G.lite مانند زمانی که تقسیم کننده بکار می رود، نیست و تنها می تواند به سرعت 1.5 MHz برسد (ولی هزاران مراجعه مستقیم به منازل مشترکان را حذف می کند). در استاندارد G.lite به تقسیم کننده ایستگاه پایانی همچنان نیاز هست.

باید بیاد داشت که، ADSL یک استاندارد لایه فیزیکی است، و آنچه که روی آن اجرا می شود به کاربرد تلفن بستگی دارد. یکی از این کاربردها ATM است، چون در ATM امکان کنترل کیفیت وجود دارد و بسیاری از شرکت های تلفن زیرساخت های آنرا در اختیار دارند.

مدار پایانی بیسیم

از سال ۱۹۹۶ در ایالات متحده آمریکا (و کمی بعد در کشورهای دیگر)، شرکت هایی که مایل بودند وارد رقابت با انحصارگر تلفن شهری (که به ILEC شهرت داشت) شوند، اجازه آنرا یافتند. شرکت های تلفن راه دور (IXC ها) اولین کسانی بودند، که وارد گود شدند. هر IXC که می خواست وارد بازار تلفن شهری شود، باید کارهای ذیل را انجام می داد. اول، زمین یا ساختمانی برای تأسیس ایستگاه پایانی (end office) بخرد یا اجاره کند. دوم، تجهیزات و سونچ های لازم را خریداری و در ایستگاه پایانی نصب کند. سوم، یک (یا چند) رشته فیبر نوری بین این ایستگاه پایانی و نزدیکترین ایستگاه بین شهری (tool office) بکشد، تا مشترکان آن بتوانند به شبکه تلفن کشوری دسترسی داشته باشند. چهارم، در صدد جلب مشتری برآید (با تبلیغ و ارائه سرویس های بهتر و قیمت کمتر). و قسمت سخت کار همین جاست.

فرض کنید یک مشتری پیدا شده و می خواهد از شرکت تلفن شهری جدید (که آنرا CLEC می نامیم) سرویس تلفن بخرد. این شرکت چگونه باید مشتری را به ایستگاه پایانی خود - که خیلی هم برای آن خرج کرده - وصل کند؟ خریدن زمین در تمام طول مسیر، کندن کانال، و کشیدن یک خط تلفن بسیار پرهزینه است. بسیاری از این CLEC ها راه ساده تر و کم هزینه تری پیدا کرده اند: WLL (مدار پایانی بیسیم - Wireless Local Loop).

از یک نظر تلفنهای ثابت با مدار پایانی بیسیم شبیه تلفنهای همراه هستند، ولی سه تفاوت فنی مهم و اساسی با آنها دارند: اول، مشترکان WLL دسترسی اینترنت پرسرعت (با سرعتی حداقل معادل ADSL) می‌خواهند. دوم، این قبیل مشترکان انتظار ندارند (و اغلب اجازه نمی‌دهند) یک آنتن بشقابی بزرگ بالای بام خانه‌هایشان نصب شود. سوم، جای این تلفن‌ها اساساً ثابت است، و مشکلات جابجا شدن سیگنال را (که در تلفنهای همراه وجود دارد، و بعداً در همین فصل با آنها آشنا خواهید دید) ندارند. بدین ترتیب یک صنعت کاملاً جدید متولد می‌شود: بیسیم ثابت (سرویس تلفن معمولی و اینترنت روی مدار پایانی بیسیم).

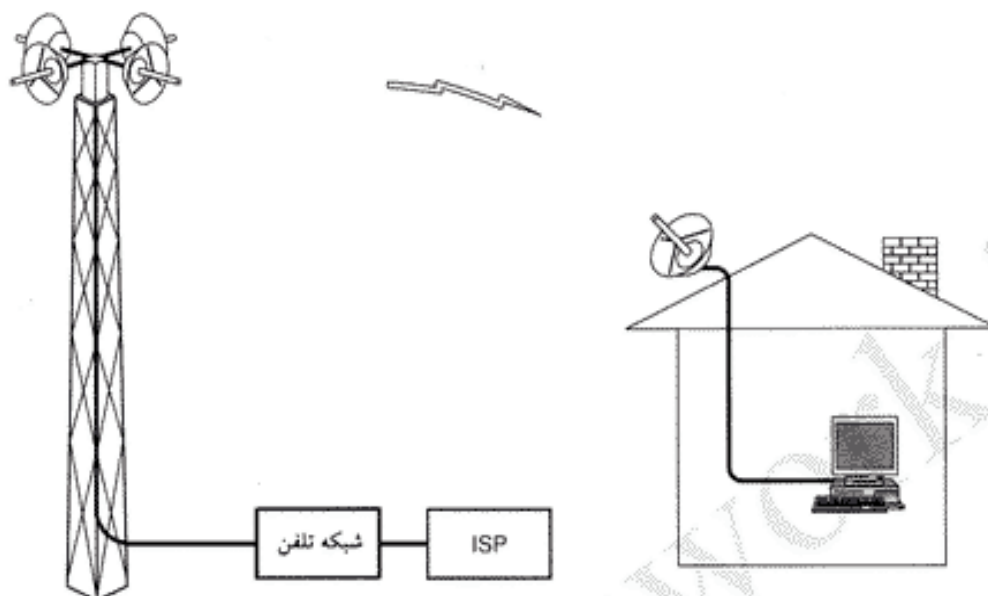
با اینکه آغاز بکار جدی WLL از سال ۱۹۹۸ است، ولی منشأ آن به سال ۱۹۶۹ برمی‌گردد. در این سال، FCC دو کانال تلویزیونی (هر یک با پهنای باند 6 MHz در فرکانس 2.1 GHz) به مصارف آموزشی اختصاص داد. در سالهای بعد تعداد این کانالها به ۳۳ (و پهنای باند مجموع آنها به 198 MHz در 2.5 GHz) افزایش یافت.

تلویزیون آموزشی هرگز با گرفت و در سال ۱۹۹۸، FCC این کانالها را پس گرفت و به ارتباطات رادیویی دوطرفه اختصاص داد - که بلافاصله در مدارهای پایانی بیسیم بکار گرفته شد. در این فرکانسها، طول موج امواج 10-12 cm است. بُرد این امواج حدود ۵۰ کیلومتر است، و از گیاهان و قطرات باران نسبتاً خوب عبور می‌کنند. سرویس جدیدی، که MMDS (سرویس توزیع چندکاناله چندنقطه‌ای - Multichannel Multipoint Distribution Service) نام گرفت، تمامی پهنای باند 198 MHz را مصرف کرد؛ MMDS را می‌توان (مانند پسر عمویش LMDS - که در همین قسمت آنرا توضیح می‌دهیم) یک شبکه شهری (MAN) در نظر گرفت. مزیت بزرگ این سرویس تکنولوژی جا افتاده و موجود بودن تجهیزات آن است؛ و بزرگترین عیب آن این است که پهنای باند متوسطی دارد، که آن هم باید بین افراد زیادی در یک محدوده جغرافیایی وسیع به اشتراک گذاشته شود.

پهنای باند پائین MMDS باعث جذب شرکتهای تلفن به امواج میلیمتری شد. در فرکانسهای 28-31 GHz در آمریکا (و 40 GHz در اروپا) تمام باندها آزاد هستند، چون تکنولوژی تجهیزات نیمه‌هادی که در این فرکانسها کار کنند، پیچیده و گرانقیمت است. اما این مشکل هم با اختراع مدارات مجتمع گالیوم-آرسناید (Ga-As IC) حل شده، و راه برای ارتباطات رادیویی در باندهای میلیمتری هموار شده است. در پاسخ به تقاضاهای موجود، FCC یک پهنای باند 1.3 GHz به سرویس بیسیم جدید بنام LMDS (سرویس توزیع چندنقطه‌ای محلی - Local Multipoint Distribution Service) اختصاص داد - که این بزرگترین پهنای باندی بود که FCC تا به آن روز به یک کاربرد خاص اختصاص می‌داد. پهنای باند مشابهی نیز در اروپا (در فرکانس 40 GHz) به این سرویس اختصاص داده شده است.

عملکرد LMDS را در شکل ۲-۳۰ ملاحظه می‌کنید. در این شکل یک برج مخابراتی می‌بینید، که تعدادی آنتن در جهت‌های مختلف روی آن نصب شده است. از آنجائیکه امواج میلیمتری شدت خطی هستند، هر آنتن زاویه محدودی (که به قطاع معروف است) را پوشش می‌دهد، که مستقل از سایر آنتنهاست. در این فرکانس، بُرد امواج 2-5 km است، و بهمین دلیل برای پوشش کامل یک شهر به تعداد زیادی از این برجهای مخابراتی نیاز هست.

در سرویس LMDS نیز، مانند ADSL، تخصیص پهنای باند نامتقارن (و به نفع کانال دریافت) است. با تکنولوژی موجود، هر قطاع می‌تواند تا 36 Mbps روی کانال دریافت کاربر و تا 1 Mbps روی کانال ارسال کاربر ظرفیت داشته باشد، که بین تمام کاربران موجود در آن قطاع به اشتراک گذاشته می‌شود. اگر هر کاربر در هر دقیقه سه صفحه 5-KB دریافت کند، بطور متوسط 2000 bps از پهنای باند را اشغال می‌کند، که بدین ترتیب همزمان حداکثر 18,000 کاربر می‌توانند در هر قطاع کار کنند. البته برای قابل قبول بودن سرویس، نباید در هر لحظه بیش از ۹۰۰۰ کاربر فعال داشته باشیم. با فرض داشتن چهار قطاع در یک برج (مانند شکل ۲-۳۰)، تعداد کاربرانی که در



شکل ۲-۳. معماری یک سیستم LMDS.

هر لحظه می توانند فعال باشند به ۳۶,۰۰۰ می رسد - و اگر فرض کنیم در ساعات اوج مصرف از هر سه کاربر یکی فعال است، تعداد کل مشترکانی که یک برج مخابراتی در شعاع ۵ کیلومتری می تواند پوشش دهد، به ۱۰۰,۰۰۰ بالغ می شود. (بر مبنای همین محاسبات بود که تعداد زیادی از CLEC ها به این نتیجه رسیدند که با یک سرمایه گذاری متوسط در برجهای مخابراتی امواج میلیمتری، می توانند در بازار تلفن شهری و اینترنت با شرکت های تلویزیون کابلی رقابت کنند.)

اما، LMDS هم مشکلات خاص خود را دارد. اول اینکه، امواج میلیمتری کاملاً خطی هستند، و هیچ مانعی در خط دید برج مخابراتی و آنتن گیرنده نبایستی وجود داشته باشد. دیگر اینکه، برگ درختان جاذب خوبی برای این امواج است، بنابراین برج فرستنده بایستی آنقدر مرتفع باشد که درختان مانع امواج آن نشوند (و باید توجه داشت مسیری که در زمستان صاف است، ممکنست در تابستان که درختان پر از برگ هستند، چنین نباشد). باران نیز امواج میلیمتری را جذب می کند. این مشکل را می توان تا حدی با گداهای تصحیح خطا، و افزایش توان تشعشعی فرستنده در روزهای بارانی جبران کرد. معهذ، سرویس LMDS در مناطق خشک شانس بیشتری برای موفقیت دارد، تا مناطق مرطوب و پُر باران.

مدار پایانی بیسیم بدون وجود استانداردهای معتبر شانس برای موفقیت ندارد، چون فقط در صورت وجود این استانداردهاست که سازندگان تجهیزات الکترونیکی رغبت ساخت وسایل مورد نیاز این صنعت را پیدا می کنند، و مشتریان نیز می توانند با خیال راحت (بدون نگرانی از اینکه با عوض کردن شرکت تلفن، وسایل را هم باید تعویض کنند) مشترک این سرویسها شوند. بمنظور تدوین استانداردهای LMDS، کمیته 802.16 توسط IEEE تأسیس شد، که این استاندارد را (که شبکه شهری بیسیم نامیده می شود) در آوریل ۲۰۰۲ عرضه کرد.

استاندارد IEEE 802.16 تلفن دیجیتال، دسترسی اینترنت، ارتباط بین شبکه های محلی، ایستگاههای پخش برنامه های رادیویی و تلویزیونی (و چندین کاربرد دیگر) را در بر می گیرد. در فصل ۴ درباره این استاندارد بیشتر صحبت خواهیم کرد.

۴-۵-۲ ترانک‌ها و مالتی‌پلکس کردن

ساخت و ساز نقش مهمی در اقتصاد سیستمهای تلفن بازی می‌کند. نصب یک خط اصلی با پهنای باند زیاد (high-bandwidth trunk) بین دو مرکز سوئیچینگ به همان اندازه یک خط اصلی با پهنای باند کم (low-bandwidth trunk) پول لازم دارد، چون هزینه اصلی در اینجا هزینه حفر کانالهاست نه هزینه کابل مسی یا فیبر نوری. بهمین دلیل، شرکت‌های تلفن از روشهای پیچیده‌ای برای ارسال همزمان چندین مکالمه روی یک خط فیزیکی استفاده می‌کنند، که به این روشها مالتی‌پلکس (multiplex) گفته می‌شود. تکنیکهای مالتی‌پلکس کردن به دسته بزرگ تقسیم می‌شوند: FDM (مالتی‌پلکس تقسیم فرکانس - Frequency Division Multiplexing) و DM (مالتی‌پلکس تقسیم زمان - Time Division Multiplexing). در FDM، طیف فرکانسی به باندهای مختلفی تقسیم می‌شود، و هر کاربر انحصاراً از یک باند استفاده می‌کند. در TDM، هر کاربر برای لحظه‌ای کوتاه (برش کوچکی از زمان) کل پهنای باند را در اختیار می‌گیرد.

در رادیوهای AM هر دو نوع مالتی‌پلکس دیده می‌شود. طیف تخصیص داده شده به هر کانال (ایستگاه) در حدود 1 MHz (تقریباً بین 500-1500 kHz) است، و هر ایستگاه در باند اختصاصی خود کار می‌کند. بین ایستگاهها نیز آنقدر فاصله وجود دارد، که با هم تداخل نکنند. این سیستم نمونه‌ای از FDM است. علاوه بر آن، در برخی از کشورها هر ایستگاه دارای دو زیرکانال منطقی است: کانال موزیک و کانال آگهی. فرستنده در برشهای کوتاه زمانی (و بصورت یک در میان) موزیک و آگهی پخش می‌کند، که در این حالت بصورت TDM کار می‌کند. در ادامه، ابتدا مالتی‌پلکس تقسیم فرکانس (و کاربرد آن در فیبرهای نوری، که به مالتی‌پلکس تقسیم طول موج معروفست) را خواهید دید. سپس با مالتی‌پلکس تقسیم زمان (و یکی از کاربردهای پیشرفته آن در فیبرهای نوری، بنام SONET) آشنا می‌شوید.

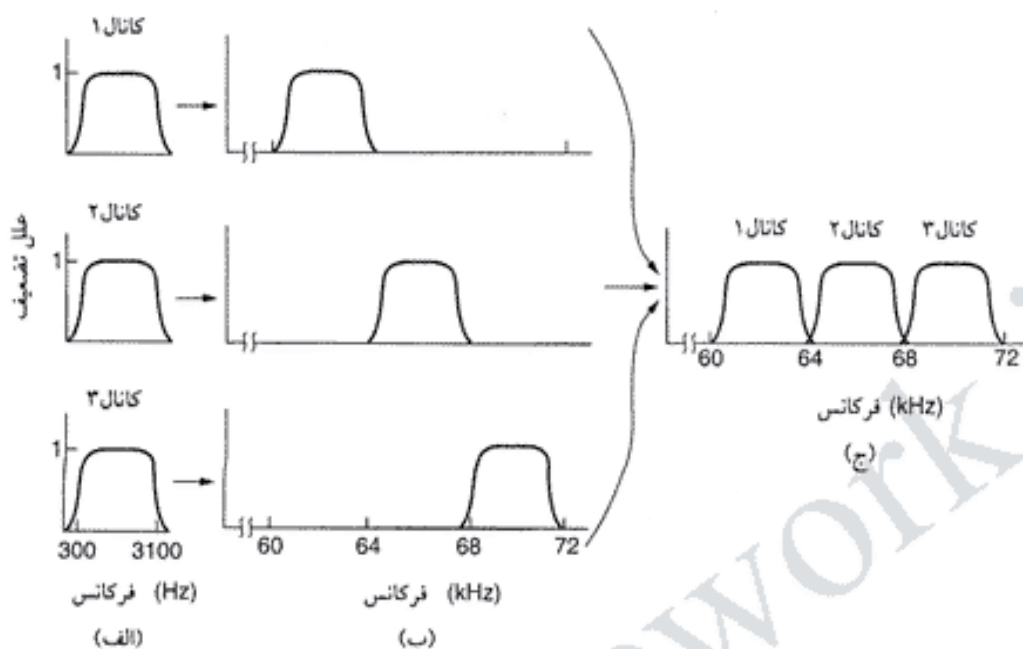
مالتی‌پلکس تقسیم فرکانس

در شکل ۲-۳۱ مالتی‌پلکس ۳ کانال صوتی با استفاده از FDM نشان داده شده است. پهنای باند هر کانال با استفاده از فیلترهای مخصوص به 3100 Hz محدود شده، ولی هنگام مالتی‌پلکس کردن پهنای هر کانال 4000 Hz در نظر گرفته می‌شود تا بین آنها تداخل پیش نیاید. در اولین قدم، فرکانس هر کانال به مقدار مشخصی (که با سایر کانالها متفاوت است) بالا برده می‌شود. سپس می‌توان این کانالها را با هم ترکیب کرد، چون دیگر هیچکدام از آنها فرکانس یکسانی ندارند و با هم مخلوط نمی‌شوند. دقت کنید که با وجود در نظر گرفتن یک حاشیه امنیتی برای هر کانال، آنها تا حدی روی هم می‌افتند، چون نقطه قطع فیلترها کاملاً تیز نیست. این روی هم افتادگی باعث بروز نوعی نویز غیرحرارتی در کانالهای مجاور می‌شود.

روشهای FDM که در سرتاسر دنیا بکار برده می‌شوند، تا حدی استاندارد شده‌اند. یکی از این استانداردها مالتی‌پلکس کردن دوازده کانال صوتی 4000-Hz روی باند 60-108 kHz است. این واحد گروه (group) نامیده می‌شود. گاهی اوقات از باند 12-60 kHz نیز بعنوان یک گروه استفاده می‌شود. بسیاری از شرکت‌های تلفن سرویسهای خطوط اجاره‌ای 48 kbps تا 56 kbps را بصورت همین گروه‌ها به مشتریان خود ارائه می‌کنند. از مالتی‌پلکس پنج گروه (یعنی ۶۰ کانال صوتی) یک فوق‌گروه (supergroup) بوجود می‌آید؛ و از مالتی‌پلکس پنج (در استاندارد CCITT) یا ده (در استاندارد شرکت پل) فوق‌گروه یک ابرگروه (mastergroup) شکل می‌گیرد. در برخی از استانداردها حتی تا ۲۳۰,۰۰۰ کانال صوتی در یک باند مالتی‌پلکس می‌شود.

مالتی‌پلکس تقسیم طول موج

برای کانالهای فیبر نوری از نوع دیگری از مالتی‌پلکس تقسیم فرکانس استفاده می‌شود، که مالتی‌پلکس تقسیم



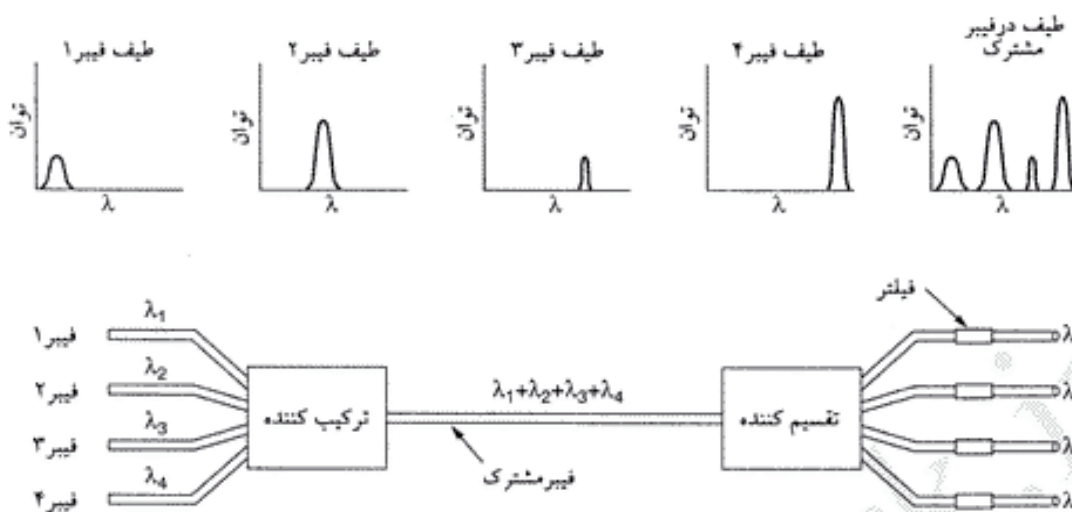
شکل ۲-۳۱. مالتی پلکس تقسیم فرکانس. (الف) کانالهای اولیه. (ب) بالا بردن فرکانس کانالها. (ج) کانالهای مالتی پلکس شده.

طول موج (Wavelength Division Multiplexing) نام دارد. اصول WDM در شکل ۲-۳۲ نشان داده شده است. در اینجا پرتوهای چهار فیبر نوری، که هر کدام طول موج متفاوتی دارند، وارد یک ترکیب کننده نوری شده، و برای ارسال آماده می شوند. در انتهای دیگر نیز عمل عکس انجام شده، و پرتوها مجدداً تفکیک می شوند. این کار توسط فیلترهای مخصوصی که فقط به یک طول موج اجازه عبور می دهند، انجام می شود.

در واقع هیچ چیز جدیدی در این تکنیک وجود ندارد، و این یک مالتی پلکس تقسیم فرکانس در فرکانسهای بسیار بالاست. مادامیکه هر کانال دارای فرکانس (یا، طول موج) خاص خود باشد، می توان آنها را با هم ترکیب (مالتی پلکس) کرد. تنها فرق این روش با FDM الکتریکی اینست که در سیستمهای نوری از فیلترهای انکساری منعزل استفاده می شود، که بسیار قابل اطمینان هستند.

سرعت پیشرفت تکنولوژی WDM آنقدر زیاد است که صنعت کامپیوتر به گرد آن هم نمی رسد. این تکنولوژی در سال ۱۹۹۰ اختراع شد، و اولین سیستم تجاری WDM دارای هشت کانال 2.5 Gbps بود. در سال ۱۹۹۸، سیستمهایی با ۴۰ کانال 2.5 Gbps به بازار عرضه شد، و در سال ۲۰۰۱ شاهد سیستمهایی با ۹۶ کانال 10 Gbps بودیم، که ظرفیت کل آنها به 960 Gbps می رسد (این پهنای باند برای ارسال ۳۰ فیلم کامل - با فرمت MPEG-2 - در هر ثانیه کافیت). در آزمایشگاهها سیستمهایی با ۲۰۰ کانال نیز تست شده اند. وقتی تعداد کانالها خیلی زیاد باشد، و طول موجها فاصله کمی (در حد 0.1 nm) با هم داشته باشند، به آن DWDM (WDM چگال - Dense WDM) گفته می گویند.

علت محبوبیت سیستمهای WDM اینست که انرژی هر فیبر نوری فقط چند گیگاهرتز پهنای دارد، چون در حال حاضر امکان تبدیل سریعتر سیگنالهای الکتریکی به پالسهای نوری (و بالعکس) وجود ندارد. با ترکیب چند کانال با طول موجهای مختلف، پهنای باند بصورت خطی افزایش پیدا می کند. از آنجائیکه پهنای باند یک فیبر نوری در حدود 25,000 GHz است (شکل ۲-۶ را ببینید)، ظرفیت آن می تواند به ۲۵۰۰ کانال 10-GHz برسد - و این



شکل ۲-۳۲. مالتی پلکس تقسیم طول موج.

در bit/Hz است، که البته نرخهای بالاتر هم امکان دارد.

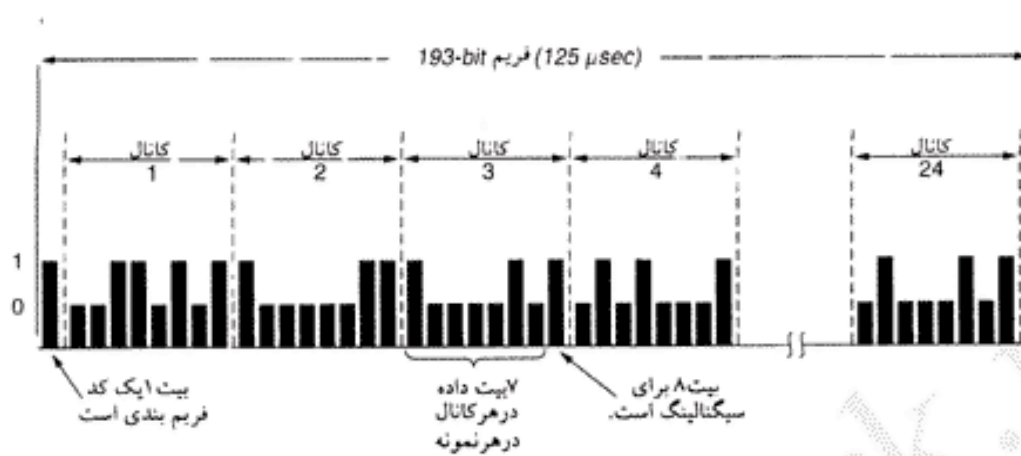
پیشرفت جدیدی که در این زمینه بدست آمده، تقویت کننده های تمام نوری است. در گذشته، لازم بود تا در فواصل ۱۰۰ کیلومتری پالسهای نوری به سیگنالهای الکتریکی تبدیل شده، و پس از تقویت دوباره به پالسهای نوری تبدیل و ارسال شوند. امروزه، تقویت کننده های تمام نوری می توانند (بدون نیاز به تبدیل کننده های نوری-الکتریکی) عمل تقویت را در هر ۱۰۰۰ کیلومتر انجام دهند.

مثال شکل ۲-۳۲ یک سیستم ثابت است: بیت های فیبر ۱ به خروجی ۳ می روند، بیت های فیبر ۲ به خروجی ۱ می روند، و الی آخر. ولی، ساخت سیستم های سونیچینگ WDM نیز امکان پذیر است. در این قبیل دستگاهها، می توان فیلترهای خروجی را با استفاده از تداخل سنجهای فابری-پروت یا ماخ-زندر تنظیم کرد. برای کسب اطلاعات بیشتر در زمینه WDM و کاربرد آن در سونیچینگ بسته اینترنت، به (Elmirghani and Moustafa, 2001; Hunter and Andonovic, 2000; Listani et al., 2001) مراجعه کنید.

مالتی پلکس تقسیم زمان

تکنولوژی WDM بسیار جالب و مهیج است، ولی هنوز هزاران هزار کیلومتر سیم مسی در شبکه تلفن موجود است که باید فکری هم بحال آنها کرد. با اینکه FDM هنوز کاربرد گسترده ای در کابلهای مسی و کانالهای مایکروویو دارد، اما این تکنولوژی اساساً آنالوگ است و نمی توان آنرا با کامپیوتر انجام داد. ولی TDM (مالتی پلکس تقسیم زمانی - Time Division Multiplexing) را می توان بطور کامل دیجیتالی کرد، و بهمین دلیل در سالهای اخیر کاربرد گسترده ای یافته است. متأسفانه، از TDM فقط برای داده های دیجیتال می توان استفاده کرد، و از آنجائیکه مدارهای پایانی سیگنالهای آنالوگ تولید می کنند، باید این سیگنالها را در ایستگاه پایانی به دیجیتال تبدیل کرده، و سپس روی ترانکها (که دیجیتال هستند) ارسال کرد.

در این قسمت نحوه دیجیتالی کردن سیگنالهای آنالوگ صدا، و ترکیب آنها برای ارسال روی ترانکهای دیجیتال را توضیح می دهیم - سیگنالهایی که کامپیوترها از طریق مودم ارسال می کنند نیز آنالوگ است، بنابراین توضیحات زیر در مورد آنها هم صادق است. سیگنالهای آنالوگ در ایستگاه پایانی توسط دستگاهی بنام کدک (codec) دیجیتالی شده، و یکسری اعداد ۸ بیتی تولید می شود. این کدک در هر ثانیه ۸۰۰۰ نمونه می گیرد (در هر



شکل ۲-۳۳. کاربرد T1 (1.544 Mbps).

۱۲۵ میکروثانیه یک نمونه)، چون طبق قضیه نایکوئیست این مقدار برای گرفتن تمام اطلاعات کانالی با پهنای ماند 4-kHz کافیست - با نرخ نمونه برداری پانسیتر اطلاعات از دست می‌رود، و با نرخ بالاتر اطلاعات بیشتری بدست نمی‌آید. این تکنیک که PCM (مدولاسیون کُد پالس - Pulse Code Modulation) نامیده می‌شود، قلب سیستمهای جدید تلفن است. در نتیجه، تمام فواصل زمانی در سیستمهای تلفن مضاربی از $125 \mu\text{sec}$ هستند. با ورود تکنولوژی مخابرات دیجیتال، CCITT نتوانست بر سر استاندارد بین‌المللی برای آن به توافق برسد، بهمین دلیل سیستمهای دیجیتال متعددی در سراسر دنیا مورد استفاده قرار گرفتند که عمدتاً با هم ناسازگار بودند. تکنیکی که در آمریکای شمالی و ژاپن از آن استفاده شد، کاربرد T1 است که آنرا در شکل ۲-۳۳ ملاحظه می‌کنید. (بخواهیم دقیقتر صحبت کرده باشیم، این فرمت DS1 و کاربرد آن T1 خوانده می‌شوند، ولی در اینجا قصد نداریم با آنچه در صنعت جا افتاده مخالفت کنیم) هر کاربرد T1 عبارتست از ۲۴ کانال صوتی که با یکدیگر مالتی پلکس شده‌اند. معمولاً، این سیگنالهای آنالوگ در فواصل زمانی منظم و متوالی به یک کُدک داده شده و دیجیتال می‌شوند (بجای آنکه از ۲۴ کُدک استفاده کرده، و خروجی آنها را ترکیب کنیم). هر یک از این ۲۴ کانال بنوبت ۸ بیت به خروجی می‌دهند. با احتساب ۷ بیت داده و یک بیت کنترل، در هر کانال $7 \times 8000 = 56,000 \text{ bps}$ داده و $1 \times 8000 = 8000 \text{ bps}$ اطلاعات کنترلی خواهیم داشت.

هر فریم دارای $24 \times 8 = 192$ بیت است، که (با احتساب یک بیت اضافی برای فریم‌بندی) ۱۹۳ بیت در هر $125 \mu\text{sec}$ خواهیم داشت، که بدین ترتیب نرخ داده ۱.۵۴۴ Mbps به می‌رسد. بیت ۱۹۳ ام برای سنکرون کردن فریم مورد استفاده قرار می‌گیرد. این بیت شکل $0101010101\dots$ بخود می‌گیرد، و گیرنده بکمک این بیت می‌تواند از سنکرون بودن اطلاعات اطمینان یابد. اگر گیرنده همزمانی خود را با فرستنده از دست دهد، می‌تواند با جستجوی این طرح بیت دوباره با فرستنده سنکرون شود. کاربران آنالوگ بکلی نمی‌توانند چنین طرحی از بیت‌ها تولید کنند، چون این طرح بیت معادل موج سینوسی 4000-Hz است که حذف خواهد شد. کاربران دیجیتال می‌توانند چنین طرحی را تولید کنند، ولی مشکل اینجاست که این طرح با لغزش فریم نیز می‌تواند ظاهر شود. برای بازیابی سریعتر سیستم در صورت بروز چنین لغزش‌هایی، هنگامی که از T1 فقط برای ارسال داده استفاده می‌شود، کانال ۲۴ ام به یک طرح خاص سنکرون کردن اختصاص می‌یابد (و فقط در ۲۳ کانال داده ارسال می‌شود).

وقتی بالاخره CCITT بر سر استانداردهای PCM بتوافق دست یافت، احساس کرد که ۸۰۰۰ بیت کنترلی

خیلی زیاد است، بنابراین استاندارد 1.544-Mbps بر اساس بسته‌های داده ۸ بیتی بنا شده نه ۷ بیتی؛ بعبارت دیگر، هر سیگنال آنالوگ بجای ۱۲۸ سطح به ۲۵۶ سطح مجزا دیجیتایز می‌شود. دو ویرایش (ناسازگار) از این استاندارد تهیه شد. در سیگنالینگ کانال مشترک (common-channel signaling) بیت اضافی (که در اینجا بجای ابتدا به انتهای فریم ۱۹۳ بیتی چسبانده می‌شود) در فریمهای فرد مقدار 10101010... بخود می‌گیرد، و در فریمهای زوج (تمام کانالها) حاوی اطلاعات سیگنالینگ است.

در ویرایش دیگر، سیگنالینگ وابسته به کانال (channel-associated signaling)، هر کانال دارای زیرکانال سیگنالینگ مخصوص بخود است. در هر زیرکانال یک بیت از هر هشت بیت داده در هر شش فریم به سیگنالینگ اختصاص داده می‌شود، بنابراین از هر شش فریم پنج تای آنها ۸ بیتی و یکی ۷ بیتی است. یکی دیگر از پیشنهادات CCITT، کاربرد PCM با ظرفیت 2.048 Mbps است که E1 نام دارد. این کاربرد دارای ۳۲ بسته ۸ بیتی در هر 125 μsec است، که ۳۰ تای آنها برای داده و دو تا برای سیگنالینگ بکار می‌روند. در هر گروه فریمی ۶۴ بیت سیگنالینگ وجود دارد، که نیمی از آن به سیگنالینگ وابسته به کانال اختصاص داده شده، و نیمی دیگر برای سنکرون کردن فریمها کنار گذاشته شده است (کشورهای مختلف می‌توانند از این نیمه بدلیخواه استفاده کنند). خارج از آمریکای شمالی و ژاپن بجای T1 از کاربرد E1 استفاده می‌شود.

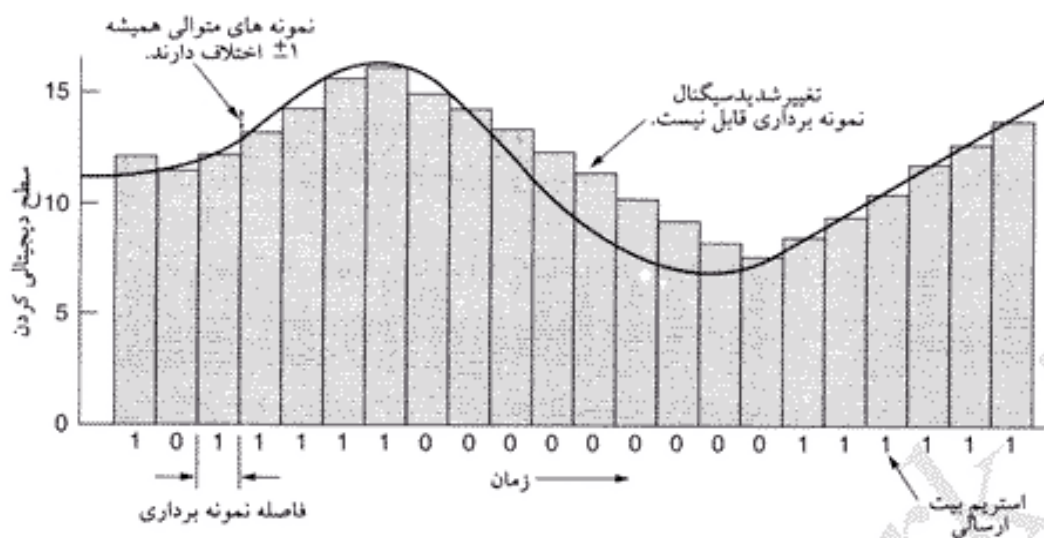
بعد از دیجیتایز کردن سیگنال صوتی، می‌توان با استفاده از تکنیکهای آماری تعداد بیتهای مورد نیاز هر کانال را کاهش داد (فشرده‌سازی). از این تکنیکها برای دیجیتایز کردن هر نوع سیگنال آنالوگ می‌توان استفاده کرد. تمام تکنیکهای فشرده‌سازی بر این اصل استوارند که تغییرات سیگنال نسبت به فرکانس نمونه برداری نسبتاً کمتر است، بنابراین قسمت اعظم سطوح دیجیتال ۷ یا ۸ بیتی تکراری و زائد هستند.

در یکی از این روشها، که مدولاسیون کُد پالس تفاضلی (differential pulse code modulation) نام دارد، نه خود دامنه دیجیتایز شده بلکه تفاوت آن با مقدار قبلی بکار برده می‌شود. از آنجائیکه در یک مقیاس ۱۲۸ تایی پرشهایی بیشتر از ± 16 ممکن نیست، بجای ۷ بیت به فقط ۵ بیت نیاز داریم. اگر سیگنال در مواردی خاص بیش از ۱۶ سطح اختلاف داشته باشد، برای نمونه برداری آن به بیش از یک دوره زمانی نیاز هست. خطایی که بدین ترتیب روی می‌دهد، در دیجیتایز کردن صدا قابل چشم‌پوشی است.

در نوع اصلاح شده این تکنیک فشرده‌سازی، هر نمونه باید با نمونه قبلی ± 1 اختلاف داشته باشد. در این شرایط بیتی که فرستاده می‌شود برای آنست که مشخص کند، نمونه جدید بیشتر از قبلی است یا کمتر از آن. این تکنیک را، که مدولاسیون دلتا (delta modulation) نامیده می‌شود، در شکل ۲-۳۴ ملاحظه می‌کنید. مانند سایر تکنیکهای فشرده‌سازی (که فرض را بر تغییرات اندک در سیگنال آنالوگ می‌گذارند) مدولاسیون دلتا هم در تغییرات شدید سیگنال دچار خطا می‌شود (که در چنین مواردی اطلاعات از دست می‌رود).

روش بهبود یافته PCM تفاضلی بر پیش‌بینی مقدار بعدی سیگنال با استفاده از برون‌یابی چند مقدار قبلی، و سپس محاسبه تفاضل این مقدار پیش‌بینی شده با مقدار واقعی سیگنال استوار است - به این روش کُد کردن پیشگویانه (predictive encoding) گفته می‌شود. البته در این روش فرستنده و گیرنده هر دو باید از روش پیش‌بینی یکسانی استفاده کنند. روش کُد کردن پیشگویانه باعث کم شدن اندازه اعدادی که باید کُد شوند، می‌شود و بهمین دلیل تعداد بیتهای ارسالی کاهش خواهد یافت.

تکنیک TDM اجازه می‌دهد تا چندین کاربرد T1 روی یک کاربرد مرتبه بالاتر مالتی‌پلکس شوند - در شکل ۲-۳۵ این تکنیک را ملاحظه می‌کنید. در اینجا چهار کانال T1 (سمت چپ) روی یک کانال T2 مالتی‌پلکس شده‌اند. در کانال T1 مالتی‌پلکس بصورت بایت به بایت انجام می‌شود، ولی از T2 به بعد مالتی‌پلکس بیت به بیت صورت می‌گیرد. چهار استریم T1 که هر کدام 1.544 Mbps هستند، بایستی خروجی 6.167 Mbps تولید کند،



شکل ۲-۳۴. مدولاسیون دلتا.

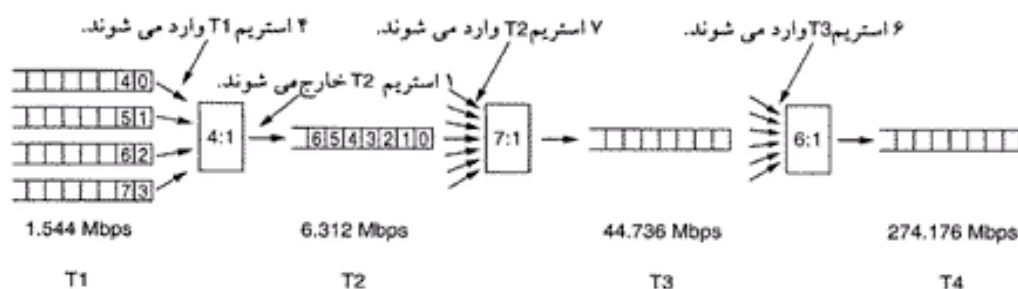
ولی ظرفیت T2 در واقع 6.312 Mbps است. پهنای اضافی برای فریم بندی و بازیابی کانال در صورت لغزش کاربر منظور شده اند. کاربرهای T1 و T3 در میان مشترکین شناخته شده تر هستند، چون T2 و T4 بیشتر برای مصارف داخلی سیستم تلفن بکار می رود.

در مرحله بعد، هفت استریم T2 بصورت بیت به بیت روی یک استریم T3 مالتی پلکس می شوند؛ از ترکیب شش استریم T3 هم یک T4 بوجود می آید. در هر مرحله مقدار کمی سرآیند نیز برای فریم بندی و تصحیح خطای لغزش کاربر اضافه می شود.

همانطور که بر سر کاربر اصلی توافق کمی بین ایالات متحده و سایر کشورهای جهان وجود دارد، نحوه مالتی پلکس آنها نیز استاندارد دقیقی ندارد. کمتر کسی را در دنیا پیدا می کنید که با ضرایب پیشنهادی آمریکا (۴، ۷ و ۶) موافق باشد، به همین دلیل CCITT از مضرب یکنواخت ۴ در تمام مراحل استفاده می کند. همچنین نحوه فریم بندی و بازیابی کانال هم در ایالات متحده و CCITT متفاوت است. در استاندارد پیشنهادی CCITT تعداد کانالها به ترتیب ۳۲، ۱۲۸، ۵۱۲، ۲۰۴۸ و ۸۱۹۲ است، که با سرعتهای 2.048 Mbps، 8.848 Mbps، 34.304 Mbps، 139.264 Mbps و 565.148 Mbps کار می کنند.

SONET/SDH

در روزهای اولیه فیبر نوری، هر شرکت تلفن برای خود یک سیستم TDM نوری اختصاصی داشت. بعد از تجزیه



شکل ۲-۳۵. مالتی پلکس استریم T1 روی کاربرهای مرتبه بالاتر.

AT&T در سال ۱۹۸۴، شرکتهای تلفن شهری مجبور شدند به کاربرهای بین شهری مختلف که هر کدام سیستم TDM نوری خاص خود را داشت متصل شوند، بهمین دلیل به فکر استاندارد کردن آن افتادند. در ۱۹۸۵، شرکت Bellcore (بازوی تحقیقاتی شرکتهای RBOC) کار روی استاندارد بنام SONET (شبکه نوری سنکرون - Synchronous Optical Network) را شروع کرد. بعدها CCITT نیز وارد این معرکه شد، که نتیجه آن استانداردهای موازی SONET، G.707 و G.709 بود. توصیه های CCITT که SDH (سلسله مراتب دیجیتال سنکرون - Synchronous Digital Hierarchy) خوانده می شوند، فقط چند تفاوت جزئی با SONET دارند. امروزه تقریباً تمامی ترافیک تلفن راه دور در ایالات متحده و اکثر نقاط دنیا، از SONET در لایه فیزیکی ترانک استفاده می کنند. برای کسب اطلاعات بیشتر درباره SONET به (Bellamy, 2000; Goralski, 2000; Shepard, 2001) مراجعه کنید.

در طراحی SONET چهار هدف اصلی مد نظر بوده است. اول و از همه مهمتر، SONET بایستی کاری می کرد که کاربرهای مختلف بتوانند با هم کار کنند. برای رسیدن به این هدف به یک استاندارد سیگنالینگ (در زمینه های طول موج، تایمینگ، ساختار فریم بندی و غیره) نیاز بود.

دوم، SONET باید راهی برای یکپارچه کردن سیستمهای دیجیتال آمریکایی، اروپایی و ژاپنی (که همگی از کانالهای PCM 64-kbps، ولی با روشهای مختلف و ناسازگار، استفاده می کردند) پیدا می کرد.

سوم، SONET باید راهی برای مالتی پلکس کردن کانالهای دیجیتال متعدد پیدا می کرد. در زمان تدوین SONET سریعترین کاربر دیجیتال، که کاربرد گسترده ای در ایالات متحده داشت، T3 با سرعت 44.736 Mbps بود. کاربر T4 تعریف شده، ولی هنوز بطور کامل عملیاتی نشده بود (برای بالاتر از آن هم هنوز کسی فکری نکرده بود). بخشی از مأموریت SONET ادامه راه تا gigabits/sec و بالاتر بود. همچنین باید راهی برای مالتی پلکس کردن کانالهای کندتر در یک کانال SONET پیش بینی می شد.

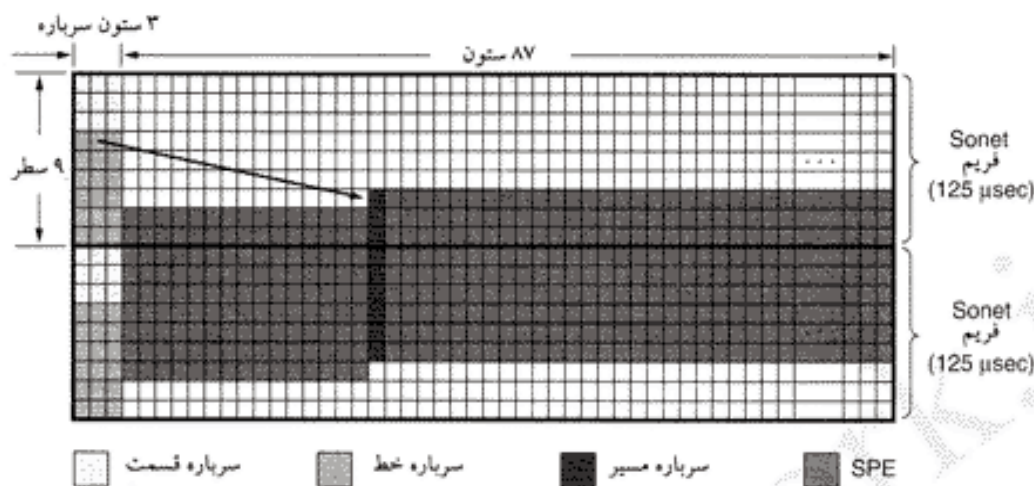
چهارم، SONET باید از یک سیستم جامع عملیات، مدیریت و نگهداری (OAM) پشتیبانی می کرد - وظیفه ای که در سیستمهای قبلی توجه چندانی به آن نشده بود.

از همان ابتدا تصمیم بر آن شد تا SONET یک سیستم ساده TDM باشد، بگونه ای که تمام پهنای باند فیبر نوری یک کانال واحد را تشکیل دهد، و هر زیرکانال برای کسری از ثانیه کانال را در اختیار بگیرد. بدین ترتیب، SONET یک سیستم سنکرون خواهد بود، که توسط یک ساعت اصلی (با دقت 1 در 10^9) کنترل می شود. در یک خط SONET بیت ها با فواصل زمانی فوق العاده دقیق (که توسط ساعت اصلی کنترل می شود) ارسال می شوند. وقتی بعدها سونیچینگ سلول بعنوان مبنای ATM انتخاب شد، برای تمایز آن با روش سنکرون SONET، از کلمه آسنکرون (Asynchronous Transfer Mode) استفاده شد. در SONET، فرستنده و گیرنده به یک ساعت مشترک وابسته اند، در حالیکه در ATM چنین نیست.

در SONET فریمها 810 بیتی هستند، که در فواصل زمانی 125 μ sec ارسال می شوند. از آنجائیکه SONET یک سیستم سنکرون است، خواه داده ای باشد یا نباشد، فریمها روی خط فرستاده می شوند. با این نرخ (8000 frames/sec)، SONET کاملاً با کانالهای PCM سازگار خواهد بود.

فریمهای ۸۱۰ بیتی SONET را می توان جدولی با ۹۰ ستون و ۹ سطر فرض کرد. بنابراین در هر ثانیه ۸۰۰۰ فریم $810 \times 8 = 6480$ بیتی ارسال می شود، که سرعت کل آنرا به 51.84 Mbps می رساند. این کانال پایه SONET است، که STS-1 (سیگنال انتقال سنکرون 1 - Synchronous Transport Signal-1) نامیده می شود. تمام ترانک های SONET مضاربی از STS-1 هستند.

همانطور که در شکل ۲-۳۶ می بینید، سه ستون اول هر فریم به اطلاعات مدیریت سیستم اختصاص یافته اند.



شکل ۲-۳۶. دو فریم متوالی SONET.

در سه سطر اول این بایتها محتوی سرآیند قسمت (section overhead) هستند، و در شش سطر بعدی محتوی سرآیند خط (line overhead). سرآیند قسمت در ابتدا و انتهای هر قسمت ایجاد و چک می‌شود، و سرآیند خط در ابتدا و انتهای خط.

فرستنده SONET فریمهای ۸۱۰ بایتی را بصورت متوالی (back-to-back) ارسال می‌کند، حتی اگر هیچ داده‌ای در کار نباشد - که در این حالت اطلاعات ساختگی ارسال می‌شود. از دید گیرنده این یک استریم پیوسته از بیتهاست، پس چگونه می‌تواند تشخیص دهد ابتدای هر فریم کجاست؟ پاسخ اینست که دو بایت ابتدایی هر فریم طرح خاصی دارند، که گیرنده می‌تواند آنها را تشخیص دهد. بمحض یافتن این طرح، گیرنده با فرستنده سنکرون می‌شود. شاید بگویند که احتمال دارد این طرح بیت‌ها در داده‌های کاربر نیز وجود داشته باشد، اما از آنجائیکه داده‌های چندین کاربر در یک فریم مالتی‌پلکس می‌شوند (و به دلایل دیگر) این اتفاق نخواهد افتاد.

ستونهای باقیمانده (۸۷ ستون) حاوی داده‌های کاربر هستند ($87 \times 9 \times 8 \times 8000 = 50.112 \text{ Mbps}$)، که به آنها SPE (بسته کاری سنکرون - Synchronous Payload Envelope) گفته می‌شود. اما داده‌های کاربر همیشه از سطر ۱ ستون ۴ شروع نمی‌شوند - SPE می‌تواند از هر کجای فریم شروع شود. اشاره گر نقطه شروع SPE در اولین سطر از سرآیند خط نوشته می‌شود. اولین ستون SPE سرآیند مسیر (path overhead) است - که سرآیند پست برای پروتکل‌های نقطه به نقطه.

با این تمهید (شروع SPE از هر نقطه فریم SONET، و امکان ادامه آن در فریمهای بعدی) انعطاف پذیری سیستم بسیار بالا می‌رود (شکل ۲-۳۶ را ببینید). برای مثال، اگر پس از شروع یک فریم SONET (و ارسال مقداری اطلاعات ساختگی) داده‌های واقعی از راه برسند، می‌توان بجای انتظار تا شروع فریم بعدی، آنها را از همان نقطه وارد فریم کرد.

سلسله مراتب مالتی‌پلکس SONET را در شکل ۲-۳۷ ملاحظه می‌کنید؛ نرخها از STS-1 تا STS-192 تعریف شده‌اند. هر STS-n یک کاربر نوری متناظر بنام OC-n دارد، که بیت به بیت با آن یکسان است (باستثنای یک جابجایی کوچک که برای سنکرون شدن لازم است). نامگذاری در استاندارد SDH متفاوت است و از OC-3 شروع می‌شود، چون در سیستمهای CCITT نرخی نزدیک به 51.84 Mbps وجود ندارد. علت وجود کاربر OC-9 هم وجود یک خط اصلی پرسرعت در ژاپن است، که چنین سرعتی دارد. از کاربرهای OC-9 و OC-18

SONET		SDH		سرعت داده (Mbps)	
الکترونیکی	نوری	نوری	ناخالص	SPE	کاربر
STS-1	OC-1		51.84	50.112	49.536
STS-3	OC-3	STM-1	155.52	150.336	148.608
STS-9	OC-9	STM-3	466.56	451.008	445.824
STS-12	OC-12	STM-4	622.08	601.344	594.432
STS-18	OC-18	STM-6	933.12	902.016	891.648
STS-24	OC-24	STM-8	1244.16	1202.688	1188.864
STS-36	OC-36	STM-12	1866.24	1804.032	1783.296
STS-48	OC-48	STM-16	2488.32	2405.376	2377.728
STS-192	OC-192	STM-64	9953.28	9621.504	9510.912

شکل ۲-۳۷. نرخهای مالتی پلکس SONET و SDH.

در ژاپن استفاده می شود. نرخ داده ناخالص شامل تمام سرآیندها می شود؛ در نرخ داده SPE سرآیندهای قسمت و خط حذف شده اند؛ و در نرخ داده کاربر تمام سرآیندها حذف، و فقط ۸۶ ستون محاسبه شده است. اگر یک کاربر مالتی پلکس نشده و فقط شامل داده های یک منبع باشد، یک حرف C (بمعنای پیوسته) به انتهای نام آن اضافه می شود. برای مثال، OC-3 نشاندهنده یک کاربر 155.52-Mbps است که از مالتی پلکس شدن سه خط OC-1 حاصل شده، ولی OC-3c استریمی است از یک منبع واحد با نرخ 155.52 Mbps. سه استریم OC-1 در یک استریم OC-3c بصورت ستون-در-میان چیده می شوند: ستون ۱ از استریم ۱، سپس ستون ۱ از استریم ۲، سپس ستون ۱ از استریم ۳، بدنبال آن ستون ۲ از استریم ۱، و الی آخر - که بدین ترتیب فریمی مرکب از ۲۷۰ ستون و ۹ سطر بوجود می آید.

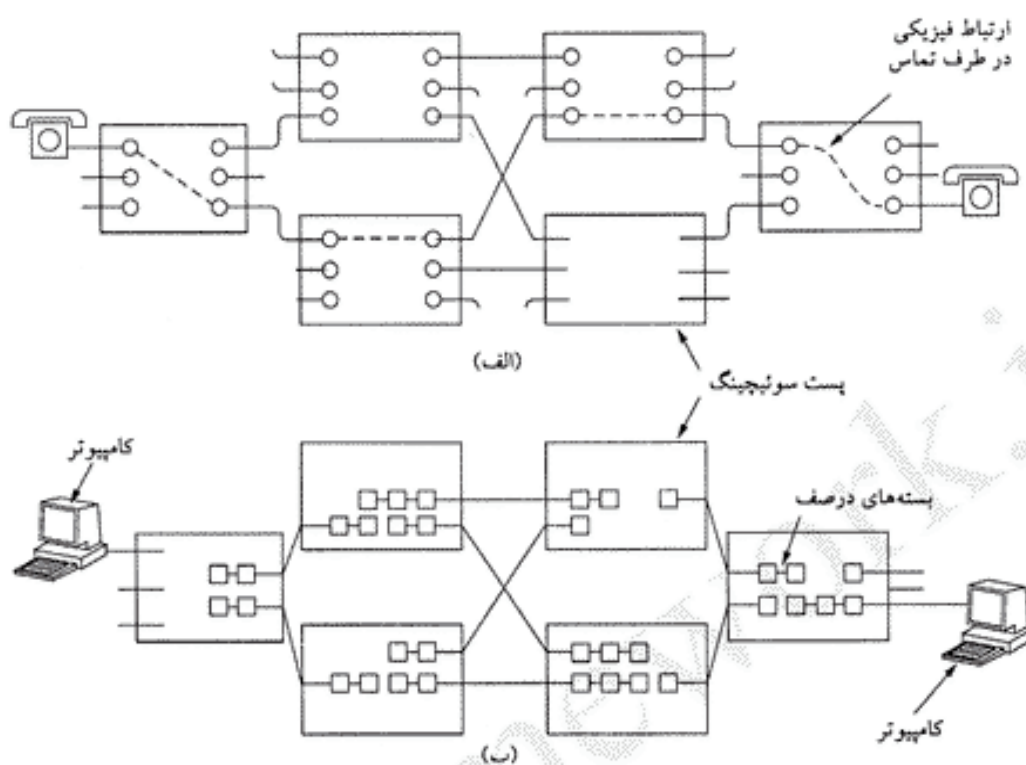
۵-۵-۲ سوئیچینگ

از دید اکثر مهندسان مخابرات، شبکه تلفن دو بخش عمده دارد: خارجی (مدارهای پایانی و ترانکها - تجهیزاتی که در فضای باز و بیرون مرکز تلفن نصب می شوند) و داخلی (سوئیچها - تجهیزاتی که در فضای بسته و داخل مرکز تلفن نصب می شوند). تا اینجا با بخش خارجی آشنا شدیم؛ اکنون وقت آنست که نگاهی به داخل مرکز تلفن بیندازیم.

امروزه از دو تکنیک متفاوت سوئیچینگ استفاده می شود: سوئیچینگ مدار (circuit switching) و سوئیچینگ بسته ای (packet switching). ابتدا هر دو تکنیک را مختصراً معرفی می کنیم، و سپس مفصلاً به سوئیچینگ مدار (که تکنیک اصلی شبکه تلفن است) می پردازیم. در فصلهای آینده درباره سوئیچینگ بسته ای بیشتر صحبت خواهیم کرد.

سوئیچینگ مدار

وقتی یک تماس تلفنی می گیرید، دستگاههای سوئیچینگ سیستم تلفن در صدد یافتن یک مسیر فیزیکی بین شما و تلفن مقصد برمی آیند. به این تکنیک که آنرا در شکل ۲-۳۸ (الف) ملاحظه می کنید، سوئیچینگ مدار گفته می شود. هر یک از مستطیلهایی که در این شکل می بینید، یک مرکز سوئیچینگ (شهری، بین شهری، و غیره) است. در این مثال، هر مرکز سوئیچینگ سه خط ورودی و سه خط خروجی دارد. وقتی تماس تلفنی از یکی از این مراکز سوئیچینگ می گذرد، یک ارتباط فیزیکی بین آن خط ورودی و یکی از خطوط خروجی برقرار می شود، که در این



شکل ۲-۳۸. (الف) سوییچینگ مداری. (ب) سوییچینگ بسته ای.

شکل با خط چین نشان داده شده است.

در روزهای اولیه تلفن، این ارتباط توسط اپراتور و یکمک یک سیم فنی که پریش ورودی را به خروجی متصل می کرد، انجام می شد. اختراع دستگاه سوییچینگ خودکار تلفن داستان جالبی دارد: این دستگاه در قرن نوزدهم در ایالت میسوری توسط فردی بنام آلمون ب. استراوگر، که شغل وی کفن و دفن بود، اختراع شد. در آن روزها وقتی کسی می مرد، یکی از بازماندگان وی با اپراتور تلفن شهر تماس می گرفت و می گفت، «لطفاً مرا به یک مؤسسه کفن و دفن وصل کنید.» در شهر آقای استراوگر دو مؤسسه کفن و دفن وجود داشت، و از شانس بد این آقای اپراتور تلفن همسر رقیب بود. آقای استراوگر خیلی زود دریافت که اگر می خواهد ورشکست نشود، باید یک دستگاه سوییچینگ خودکار تلفن اختراع کند - و این کار را کرد. همه آنهايي که در سراسر دنیا با دستگاههای سوییچینگ خودکار تلفن سروکار دارند، آنها را با نام «دستگاه استراوگر» می شناسند. (تاریخ نمی گوید آیا این خانم بعد از بیکار شدن توانست شغلی مانند اپراتور اطلاعات تلفن - که باید به سؤالاتی از قبیل «لطفاً شماره یک مؤسسه کفن و دفن را بدهید،» پاسخ دهد - بدست آورد یا خیر؟)

البته شکل ۲-۳۸ (الف) بسیار ساده شده است، چون مسیر فیزیکی بین دو تلفن می تواند از لیستکهای مایکروویو یا فیبر نوری (که هزاران تماس تلفنی روی آنها مالتی پلکس می شود) عبور کند. با این حال مفهوم کلی آن همچنان معتبر است: وقتی تماس تلفنی برقرار می شود، یک مسیر فیزیکی بین دو دستگاه تلفن بوجود می آید که تا آخر تماس باقی می ماند.

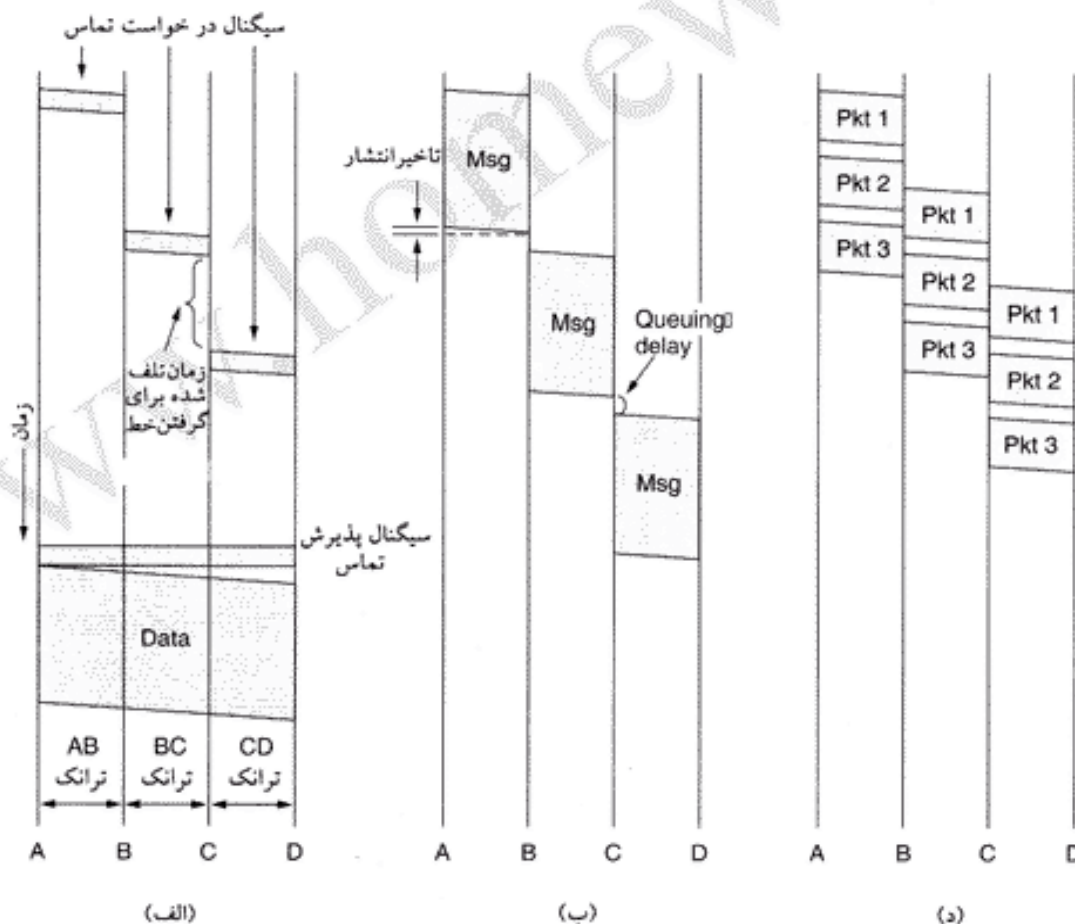
در شکل ۲-۳۸ (ب) روش جایگزین سوییچینگ مداری، که سوییچینگ بسته ای نام دارد، را می بینید. در این تکنولوژی هر بسته مستقلاً فرستاده می شود، بدون آنکه از قبل مسیری به آن اختصاص داده شده باشد؛ این بر عهده هر بسته است که راه خود را به سمت مقصد پیدا کند.

مهمترین ویژگی سوئیچینگ بسته‌ای اینست که قبل از ارسال هر گونه داده‌ای، بایستی یک مسیر نقطه-به-نقطه بین مبدأ و مقصد برقرار شده باشد. فاصله زمانی بین شماره‌گیری در مبدأ و زنگ خوردن تلفن مقصد براحتی می‌تواند به ۱۰ ثانیه برسد (که این زمان در تماسهای راه‌دور و بین‌المللی حتی بیشتر است). در این مدت سیستم تلفن بدنبال یافتن یک مسیر مناسب است (شکل ۲-۳۸ الف) را ببینید. توجه داشته باشید که قبل از شروع ارسال داده، سیگنال درخواست (request) باید تمام مسیر را تا مقصد طی کرده، و تصدیق (acknowledgement) آن بازگردد. این تأخیر در بسیاری از کاربردهای کامپیوتری (مانند بررسی اعتبار مشتری در خریدهای اعتباری) قابل قبول نیست.

اما بمحض آنکه مدار برقرار شد، دیگر تأخیر چندانی بین گیرنده و فرستنده وجود ندارد (فقط زمان تأخیر انتشار امواج الکترومغناطیس، که آن هم چیزی در حدود $5 \mu\text{sec}/\text{km}$ است). دیگر اینکه (بدلیل وجود مدار اختصاصی)، در سوئیچینگ مداری پدیده‌ای بنام ازدحام (congestion) وجود ندارد (البته بدلیل اینکه ظرفیت خطوط و مراکز سوئیچینگ نامحدود نیست، قبل از برقراری مدار همیشه احتمال شنیدن بوق اشغال وجود دارد).

سوئیچینگ پیام

یکی دیگر از اشکال سوئیچینگ، که آنرا در شکل ۲-۳۹ (ب) ملاحظه می‌کنید، سوئیچینگ پیام



شکل ۲-۳۹. همزمانی رویدادها در (الف) سوئیچینگ مداری، (ب) سوئیچینگ پیام، (ج) سوئیچینگ بسته‌ای.

(message switching) است. در این نوع از سوییچینگ نیز مسیر فیزیکی ثابتی بین فرستنده و گیرنده وجود ندارد. وقتی فرستنده یک بلوک از داده ها را ارسال می کند، این داده ها در اولین مرکز سوییچینگ (که همان روتر است) ذخیره شده (store) و سپس به مرکز بعدی هدایت (forward) می شود. هر بلوک ابتدا بطور کامل دریافت شده، از نظر خطا بررسی و سپس ارسال می شود. همانطور که در فصل قبل هم گفتیم، به شبکه ای که با این روش کار می کند شبکه ذخیره-هدایت (store-and-forward) گفته می شود.

اولین تجهیزات مخابرات الکترومغناطیسی تجهیزات سوییچینگ پیام بودند، که برای ارسال تلگرام بکار گرفته شدند. پیام ابتدا در اداره تلگراف روی نوارهای کاغذی سوراخ می شد، و بعد از خوانده شدن توسط دستگاههای خاص به مرکز بعدی فرستاده می شد، که در آنجا بصورت یک نوار کاغذی سوراخ شده از دستگاه بیرون می آمد. اپراتور مسئول دستگاه کاغذ را پاره کرده، و در یک دستگاه نوارخوان (tape reader) می گذاشت تا بتواند پیام را بخواند (هر خط مخابراتی یک دستگاه نوارخوان داشت). این هم در واقع نوعی سوییچینگ است، که به سوییچینگ نوار پاره (torn tape switching) معروف بود. نوارهای کاغذی و سوییچینگ پیام مدتهاست که از دور خارج شده اند، و ما هم بیش از این درباره آنها صحبت نخواهیم کرد.

سوئیچینگ بسته ای

اندازه پیام در شبکه سوییچینگ پیام هیچ محدودیتی ندارد، و این به آن معناست که دستگاههای مسیریاب برای نگهداری پیامها به وسایل ذخیره سازی (از قبیل، دیسک) نیاز دارند. پیامد دیگر این روش آنست که یک پیام واحد می تواند آنقدر بزرگ باشد که برای دقایقی خط مسیریاب-مسیریاب را اشغال کند، که بدین ترتیب کاربرد سوییچینگ پیام را در ارتباطات تعاملی (interactive) بشدت محدود می کند. برای حل این مشکل، سوییچینگ بسته ای (packet switching) اختراع شد. در شبکه های سوییچینگ بسته ای روی اندازه بسته ها محدودیت شدیدی اعمال می شود، و بهمین دلیل مسیریابها نیازی به دیسک برای ذخیره کردن بسته ها ندارند، و می توانند آنها در حافظه اصلی خود ذخیره کنند. با محدود کردن اندازه بسته ها، و اطمینان از اینکه یک کاربر نمی تواند خط انتقال را برای مدتی طولانی - که البته در اینجا منظور از طولانی بیش از چند میلی ثانیه است - به انحصار خود در آورد، شبکه های سوییچینگ بسته ای برای کاربردهای تعاملی بسیار مناسبند. با مقایسه شکل های ۲-۳۹ (ب) و (ج) یکی دیگر از مزایای روش سوییچینگ بسته ای بر سوییچینگ پیام را مشاهده می کنید: بسته اول یک پیام چندبسته ای می تواند حتی قبل از رسیدن بسته دوم به مسیریاب بعدی فرستاده شود، که این زمان تأخیر را کاهش داده و کارایی سیستم را بالا می برد. به دلایل فوق، شبکه های کامپیوتری اغلب از سوییچینگ بسته ای استفاده می کنند؛ سوییچینگ مداری نیز در موارد خاصی بکار برده می شود، ولی سوییچینگ پیام هیچ کاربردی در شبکه های کامپیوتری ندارد.

سوئیچینگ مداری و سوییچینگ بسته ای تفاوت های زیادی با یکدیگر دارند. برای مثال، در یک شبکه سوییچینگ مداری قبل از ارسال اطلاعات بایستی مدار فیزیکی بین فرستنده و گیرنده برقرار شده باشد، در حالیکه در شبکه های سوییچینگ بسته ای چنین الزامی وجود ندارد و ارسال بسته ها می تواند بلافاصله شروع شود. پیامد لزوم برقراری مدار ثابت در سوییچینگ مداری، اختصاص پهنای باند در تمام طول مسیر بین فرستنده و گیرنده است: تمام بسته ها باید از این مسیر عبور کنند. از طرف دیگر وقتی تمام بسته ها مجبور به عبور از یک مسیر باشند، نمی توانند خارج از ترتیبی که ارسال شده اند، به مقصد برسند. در سوییچینگ بسته ای هیچ مسیر ثابتی وجود ندارد، و بسته ها می توانند از هر مسیری که (در آن لحظه خاص) در شبکه موجود است عبور کنند، و حتی خارج از نظم و ترتیب اولیه به مقصد برسند.

ویژگی تحمل خطا در شبکه های سونیچینگ بسته ای بسیار بهتر از شبکه های سونیچینگ مداری است - و در واقع دلیل اختراع آن هم همین بوده است. وقتی در شبکه سونیچینگ بسته ای یک مسیریاب از کار می افتد، بسته ها می توانند از مسیرهای دیگری که وجود دارد، استفاده کرده و مسیریاب مرده را دور بزنند.

البته وجود یک پهنای باند اختصاصی در شبکه های سونیچینگ مداری این مزیت را دارد که بسته ها بمحض رسیدن به یک مسیریاب به مسیریاب بعدی فرستاده می شود، و زمان تأخیر ارسال پشددت کاهش می یابد؛ در حالیکه در شبکه های سونیچینگ بسته ای چنین پهنای باندی اختصاصی وجود ندارد، و بسته ها باید تا رسیدن نوبت ارسال در صف منتظر بمانند.

وجود مدار اختصاصی در شبکه های سونیچینگ مداری بدان معناست که (بعد از برقراری مدار) دیگر حالت ازدحام - انتظار برای باز شدن راه - بروز نخواهد کرد. البته همیشه این احتمال وجود دارد که در شروع ارتباط بدلیل شلوغی شبکه، امکان اختصاص مدار وجود نداشته باشد؛ این نوع دیگری از ازدحام - انتظار برای تخصیص مدار - است.

پهنای باندی که به یک کاربر تخصیص داده می شود، در تمام مدت در اختیار وی است، حتی اگر هیچ چیز برای ارسال نداشته باشد. امکان استفاده از این مدار برای کاربران دیگر وجود ندارد. در شبکه های سونیچینگ بسته ای اتلاف پهنای باند به شکل فوق وجود ندارد، و بهمین دلیل کارایی کلی آن بسیار بهتر است. درک این تفاوت بین سونیچینگ بسته ای و سونیچینگ مداری بسیار مهم است: تفاوت تضمین سرویس به قیمت اتلاف منابع، با استفاده بهینه از منابع به قیمت عدم تضمین سرویس.

سونیچینگ بسته ای از تکنیک ذخیره - هدایت استفاده می کند. در این روش هر بسته قبل از ارسال به مسیریاب بعدی باید بطور کامل دریافت و در حافظه مسیریاب ذخیره شود. این روش تأخیر نسبتاً قابل ملاحظه ای ایجاد می کند؛ در حالیکه در سونیچینگ مداری، بیت ها بطور پیوسته روی مدار منتقل می شوند و چنین تأخیری وجود ندارد.

تفاوت دیگر اینست که سونیچینگ مداری بطور کامل شفاف است: فرستنده و گیرنده می توانند از هر نرخ بیت، فرمت، و یا روش فریم بندی که می خواهند استفاده کنند؛ کاربر در این مورد هیچ چیز نمی داند، و به آن اهمیتی هم نمی دهد. اما در سونیچینگ بسته ای این کاربر است که پارامترهای اصلی را تعیین می کند. تفاوت این دو تقریباً مانند جاده و راه آهن است: در جاده این مسافر است که سرعت، اندازه و نوع وسیله نقلیه را انتخاب می کند، در حالیکه در راه آهن انتخاب این پارامترها بر عهده شرکت مسافربری (کاربر) است. همین شفافیت است که به سیستم تلفن اجازه می دهد انواع اطلاعات (صوت، فکس و داده) را منتقل کند.

آخرین تفاوت سونیچینگ مداری و سونیچینگ بسته ای روش محاسبه هزینه است. در سونیچینگ مداری (بدلیل تاریخی) هزینه بر اساس مسافت و مدت محاسبه می شود. در تلفنهای همراه، مسافت (البته باستثنای مکالمات بین المللی) نقشی ندارد، و مدت مکالمه نیز نقش ناچیزی دارد (برای مثال، تفاوت هزینه مکالمه در روز، شب یا ایام تعطیل). در سونیچینگ بسته ای، اساساً چیزی بنام مدت مکالمه وجود ندارد، و فقط گاهی حجم ترافیک نقشی در هزینه ها بازی می کند. در مصارف خانگی، معمولاً هزینه ها بصورت ماهیانه ثابت اخذ می شود، چون این روش برای ISP ها ساده تر است و کاربران نیز راحتتر با آن کنار می آیند، ولی کاربرهای اصلی شبکه هزینه ها را بر اساس حجم ترافیک دریافت می کنند. تفاوتهایی را که در این قسمت برشمردیم، بصورت خلاصه در شکل ۲-۴۰ ملاحظه می کنید.

سونیچینگ مداری و بسته آنقدر مهم هستند که بزودی دوباره به این مبحث برگشته، و تکنولوژیهای مختلف آنها را به تفصیل بررسی خواهیم کرد.

سوئیچینگ بسته ای	سوئیچینگ مداری	آیتم
لازم ندارد	لازم دارد	برقراری تماس
خیر	بلی	مسیر فیزیکی اختصاصی
خیر	بلی	تمام بسته ها از یک مسیر عبور می کنند
خیر	بلی	بسته ها به ترتیب دریافت می شوند
خیر	بلی	خرابی سوئیچ مرگ آور است
متغیر	ثابت	پهنای باند موجود
در هر بسته	در لحظه شروع	زمان از حاکم احتمالی
خیر	بلی	پهنای باند تلف شده
بلی	خیر	ذخیره - هدایت
خیر	بلی	شفافیت
به ازای هر بسته	در دقیقه	هزینه

شکل ۲-۴۰. مقایسه شبکه های سوئیچینگ مداری و سوئیچینگ بسته ای.

۶-۲ شبکه تلفن همراه

سیستم تلفن معمولی (حتی اگر سرعت آن به دهها گیگابایت برسد) دسته خاصی از کاربران را هرگز راضی نخواهد کرد: آنهایی که در یک جا بند نمی شوند. امروزه مردم می خواهند از هر جایی که هستند (داخل هواپیما، سوار بر اتومبیل، کنار دریا، نوک قله کوهها و یا از اعماق جنگل) تلفن بزنند - و حتماً تا چند سال دیگر انتظار دارند از این نقاط ایمیل خود را نیز چک کنند، و یا در وب گشت بزنند. پیامد این گرایش رشد چشمگیر تلفنهای غیر ثابت در سالهای اخیر است. در این قسمت قصد داریم مبحث تلفن همراه را بررسی کنیم.

تلفنهای غیر ثابت به دو دسته بزرگ تقسیم می شوند: گوشی بیسیم (cordless phone)، و تلفن همراه (mobile phone) - که گاهی به آن تلفن سلولی (cell phone) نیز گفته می شود. گوشی بیسیم عبارتست از یک دستگاه مرکزی و یک گوشی متحرک، که برای مصارف خانگی (مسافتهای کوتاه) طراحی شده است. این وسیله هرگز کاربردی در شبکه نداشته، و ما هم بیش از این به آن نخواهیم پرداخت. تلفن همراه، که امروزه کاربرد گسترده ای در ارتباطات صدا و داده دارد، موضوع اصلی بحث ماست.

تکامل تلفنهای همراه سه نسل را با تکنولوژیهای متفاوت پشت سر گذاشته است:

۱. صدای آنالوگ
۲. صدای دیجیتال
۳. صدای دیجیتال و داده (اینترنت، ایمیل، و غیره)

با اینکه بیشتر تکنولوژی این سیستمها مورد علاقه ماست، اما جالبست بدانید که سیاست و تصمیمات اقتصادی چه تأثیری بر تکامل این سیستمها دارد. اولین سیستم تلفن همراه در ایالات متحده آمریکا و توسط AT&T اختراع شد، که FCC هم بلافاصله آنرا در تمام کشور اجباری کرد. در نتیجه، تمام ایالات متحده صاحب یک سیستم واحد تلفن همراه (آنالوگ) شد، و تلفنی که در کالیفرنیا خریده شده بود، در نیویورک هم کار می کرد. اما در اروپا اوضاع کاملاً عکس این بود، و هر کشوری سیستم خاص خود را طراحی کرد، که نتیجه آن یک هرج و مرج کامل بود.

اروپا از اشتباه خود درس گرفت و وقتی نوبت تلفن همراه دیجیتال رسید، تمام شرکتهای مخابرات دولتی دور

هم جمع شده و بر سر یک استاندارد واحد (GSM) به توفیق رسیدند، که در نتیجه تلفنهای موبایل اروپایی در تمام نقاط این قاره کار می‌کند. در همان زمان ایالات متحده به این نتیجه رسیده بود که دولت نباید در موضوع استاندارد دخالت کند، و در نتیجه سرنوشت تلفن همراه دیجیتال به بازار سپرده شد. این تصمیم باعث شد تا انواع مختلفی از تلفن همراه دیجیتال تولید و وارد بازار شود - ایالات متحده اکنون دو سیستم بزرگ تلفن همراه دیجیتال (بعلاوه یک سیستم کوچکتر) دارد، که هیچکدام با دیگری سازگار نیست.

اروپا فقط در یک دوره زمانی کوتاه در زمینه تعداد کاربران تلفن همراه از آمریکا عقب بود، ولی اکنون از آن بسیار پیش افتاده است، که یکی از دلایل آن سیستم یکپارچه تلفن همراه در اروپاست، اما دلایل دیگری هم برای آن وجود دارد. تفاوت دیگر سیستم تلفن همراه آمریکا و اروپا (که باعث سرافکنندگی آمریکایی‌هاست) شماره‌های آنهاست. در ایالات متحده شماره‌های تلفن همراه با شماره تلفنهای معمولی (ثابت) مخلوط است. هیچ راهی برای تماس گیرنده وجود ندارد تا تشخیص دهد شماره‌ای که دارد می‌گیرد (مثلاً، 234-5678 (212))، یک شماره معمولی (کم‌هزینه) است یا یک شماره تلفن همراه (با هزینه زیاد). برای عصبی‌تر کردن مردم، شرکت‌های تلفن قرار گذاشته‌اند تا هزینه تماسها را به پای صاحب تلفن همراه بنویسند. به همین دلیل اغلب مردم در خرید تلفن همراه تردید دارند، و از صورتحسابهای نجومی که ممکنست (در نتیجه تماس دیگران) برای آنها بیاید، می‌ترسند. در اروپا، شماره‌های تلفن همراه کد ناحیه مشخصی دارند (مانند تلفنهای 800 یا 900)، و بسادگی از تلفنهای معمولی قابل تشخیص هستند. هزینه تماس هم مثل تلفنهای معمولی به پای تماس گیرنده نوشته می‌شود (البته با استثنای تماسهای بین‌المللی، که هزینه تماس بین دو طرف تقسیم می‌شود).

دلیل دیگری که باعث پذیرش گسترده تلفن همراه در اروپا شده، ابداع تلفنهای همراه از پیش پرداخت شده (pre-paid) است، که تا ۷۵٪ تلفنهای این قاره را در برخی نقاط شامل می‌شود. این تلفن‌ها را می‌توان در هر فروشگاهی (بدون هیچگونه تشریفات خاص) خرید - فقط پولش را بده و استفاده کن. این تلفن‌ها معمولاً با اعتبار ۲۰ یا ۵۰ یورو (واحد پول اروپا) عرضه می‌شوند، و بعد از صفر شدن اعتبار می‌توان آنها را (با استفاده از یک PIN code سری) دوباره شارژ کرد. امروزه هر نوجوان (و حتی بچه‌ای) در اروپا یک چنین تلفن همراهی دارد، و والدین وی می‌توانند بسادگی محلی او را پیدا کنند (بدون اینکه از صورتحسابهای نجومی آن بترسند). این تلفن‌ها (البته اگر از آنها برای تماس گرفتن استفاده نشود) ماهها دوام می‌آورند، چون نه هزینه ماهیانه دارند نه برای تلفنهایی که به آنها می‌شود، اعتباری کسر می‌شود.

۲-۱۶ تلفن‌های همراه نسل اول: صدای آنالوگ

صحبت از سیاست و روشهای بازاریابی دیگر بس است؛ اجازه دهید به کار اصلی خود یعنی تکنولوژی پردازیم. حتی از اوایل قرن بیستم نیز تلفنهای متحرک رادیویی-سیم در زمینه‌های نظامی و مسافرت‌های دریایی کارایی خود را به اثبات رسانده بودند. در سال ۱۹۴۶، اولین تلفنهای مخصوص اتومبیل در سنت لوئیس راه‌اندازی شدند. در این سیستم آنتنهای (فرستنده-گیرنده) بزرگی روی ساختمانهای بلند نصب شده بود، و ارسال و دریافت روی یک کانال واحد صورت می‌گرفت. صحبت کردن و شنیدن در آن واحد امکان نداشت: برای صحبت کردن باید یک دکمه را فشار می‌دادید، و با رها کردن آن دیگر نمی‌توانستید حرف بزنید. تا اوایل دهه ۱۹۵۰، این سیستم، که به فشار بده-حرف بزن (push-to-talk) معروف بود، در بسیاری از شهرهای بزرگ آمریکا نصب شد. این سیستم امروزه هم در اتومبیل‌های پلیس و ناخسی تلفنی مورد استفاده قرار می‌گیرد.

در دهه ۱۹۶۰، IMTS (سیستم تلفن همراه بهبودیافته - Improved Mobile Telephone System) نصب و راه‌اندازی شد. این سیستم نیز از فرستنده-گیرنده‌های قوی (۲۰۰ وات) در نقاط مرتفع استفاده می‌کرد، ولی برای ارسال و دریافت دو فرکانس متفاوت داشت، که بدین ترتیب نیازی به دکمه «فشار بده-حرف بزن» نبود. علاوه بر

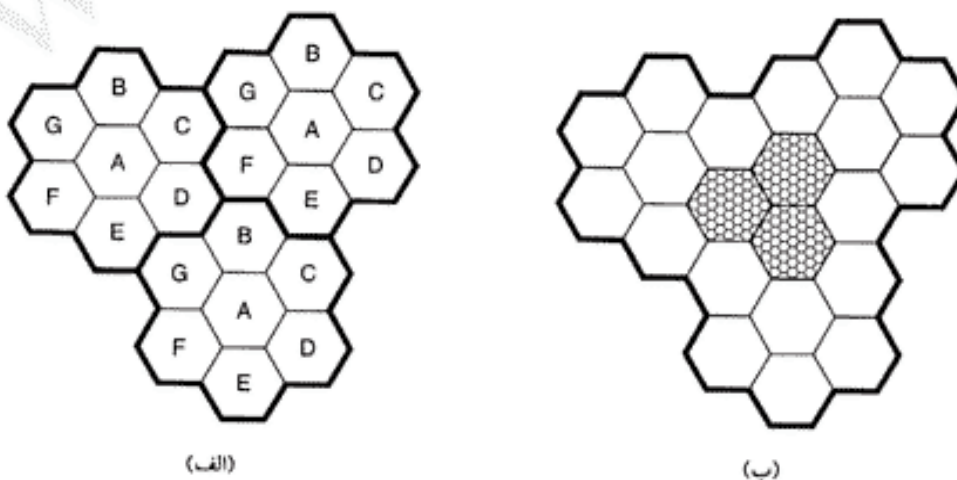
آن، برای ارسال نیز از فرکانسهای متعددی استفاده می‌شود، و خطر شنیده شدن صحبت دیگران (بر خلاف تلفنهای تاکسی تلفنی) نیز وجود نداشت.

سیستم IMTS از ۲۳ کانال (در طیف 150-450 Mhz) پشتیبانی می‌کرد. بدلیل کم بودن تعداد این کانالها، کاربران مجبور بودند برای شنیدن بسوق آزاد مدت زیادی صبر کنند. همچنین بدلیل قدرت زیاد فرستنده-گیرنده‌های این سیستم، تا شعاع چند صد کیلومتری در سیستمهای رادیویی اختلال ایجاد می‌کردند. همه این دلایل IMTS را به سیستمی غیر عملی تبدیل کرده بود.

سیستم تلفن همراه پیشرفته

در سال ۱۹۸۲، با اختراع AMPS (سیستم تلفن همراه پیشرفته - Advanced Mobile Phone System) توسط شرکت Bell Labs اوضاع تغییر کرد. این سیستم در انگلستان (با نام TACS) و ژاپن (با نام MCS-L1) نیز نصب و بکار گرفته شد. با آنکه این سیستم دیگر تکنولوژی برتر محسوب نمی‌شود، اما از آنجائیکه بمنظور سازگاری با گذشته بسیاری از ویژگیهای بنیادی آن در تلفن همراه دیجیتال (DAMPS) لحاظ شده است، قدری درباره آن صحبت خواهیم کرد.

در تمام سیستمهای تلفن همراه، یک منطقه جغرافیایی به تعدادی سلول (cell) تقسیم می‌شود - که بهمین دلیل به آن تلفن سلولی (cell phone) نیز می‌گویند. در AMPS هر سلول ۱۰ تا ۲۰ کیلومتر قطر دارد؛ در سیستمهای دیجیتالی این سلولها کوچکترند. هر سلول دارای تعدادی فرکانس است، که سلولهای همسایه از آنها استفاده نمی‌کنند. ایده کلیدی در سیستمهای سلولی که باعث شده تا ظرفیت آنها بسیار بیشتر از سیستمهای قبلی باشد، استفاده از سلولهای نسبتاً کوچک و بکارگیری فرکانسها در سلولهای نزدیک هم (البته، نه سلولهای مجاور) است. در حالیکه یک سیستم IMTS با قطر ۱۰۰ کیلومتر می‌تواند روی هر فرکانس یک تماس داشته باشد، یک سیستم AMPS قادر است (با جدا کردن فرکانس سلولهای مجاور) روی هر فرکانس ۱۰ تا ۱۵ تماس برقرار کند. هر چه اندازه سلولها کوچکتر (و تعداد آنها بیشتر) باشد، ظرفیت سیستم بالاتر خواهد رفت. با کوچکتر شدن سلولها می‌توان قدرت فرستنده-گیرنده‌ها را نیز کمتر کرد، و از تجهیزات ساده‌تر و ارزاتری استفاده کرد. طبق مقررات FCC، حداکثر قدرت تشعشعی دستگاههای تلفن همراه 0.6 وات، و فرستنده-گیرنده‌های مخصوص اتومبیل 3 وات تعیین شده است.



شکل ۲-۴۱. (الف) فرکانسهای سلولهای مجاور یکسان نیستند. (ب) برای افزایش تعداد کاربران، می‌توان از سلولهای کوچکتر استفاده کرد.

ایده تکرار فرکانسها را در شکل ۲-۴۱ (الف) ملاحظه می کنید. سلولها تقریباً به شکل دایره هستند، ولی برای سادگی کار بهتر است آنها را شش ضلعی فرض کنیم. در این شکل تمام سلولها هم اندازه اند، و در گروههای هفت تایی دسته بندی شده اند. هر حرف نشان دهنده مجموعه ای از فرکانسها است. توجه کنید که هر مجموعه فرکانسی با نزدیکترین سلول مشابه حداقل و سلول فاصله دارد، که این باعث به حداقل رسیدن تداخل فرکانسها خواهد شد.

یکی از بزرگترین معضلات سیستمهای تلفن همراه، یافتن نقاط مناسب برای نصب آنتنهای مرکزی سلولهاست. این مشکل باعث شده تا شرکتهای تلفن برای استفاده از مناره های رفیع کلیساها دست بدامن آنها شوند. وقتی در یک منطقه تعداد کاربران افزایش می یابد و سیستم دیگر جوابگوی بار مکالمات نیست، (با کم کردن قدرت آنتنها) هر سلول به سلولهای کوچکتر (microcell) تقسیم می شود، تا بتوان از فرکانسها بدفعات بیشتر استفاده کرد - شکل ۲-۴۱ (ب) را ببینید. در مواقعی که تعداد زیادی تلفن همراه برای مدتی کوتاه در یک نقطه گرد می آیند (مانند مسابقات ورزشی و کنسرتها موسیقی)، شرکتهای تلفن با استقرار آنتنهای متحرک (که ارتباط ماهواره ای دارند) موقتاً سلولهای کوچکتری ایجاد می کنند. اینکه اندازه یک سلول چقدر باید باشد، مسئله پیچیده ایست که در (Hac, 1995) درباره آن بحث شده است.

در مرکز هر سلول یک ایستگاه قرار دارد که تمام تلفنهای داخل سلول امواج خود را به آن می فرستند. در سیستمهای کوچک، تمام این ایستگاهها به یک دستگاه مرکزی بنام MTSO (مرکز سوئیچینگ تلفن همراه - Mobile Telephone Switching Office) یا MSC (مرکز سوئیچینگ همراه - Mobile Switching Center) متصل می شوند. در سیستمهای بزرگتر چندین MTSO وجود دارند، که بنوبه خود به یک MTSO بالاتر متصلند. MTSO ها، در واقع، همان ایستگاههای پایانی سیستم تلفن همراه هستند، که به حداقل یک ایستگاه پایانی سیستم تلفن ثابت نیز ارتباط دارند. ارتباط MTSO ها با تلفنهای همراه، با یکدیگر و با مراکز PSTN از طریق یک شبکه سوئیچینگ بسته ای صورت می گیرد.

هر تلفن همراه در هر لحظه در یک سلول (و تحت کنترل ایستگاه مرکزی آن) قرار دارد. وقتی این تلفن سلول را ترک می کند، ایستگاه مرکزی متوجه ضعیف شدن سیگنال آن شده و از تمام ایستگاههای مجاور میزان قدرت دریافتی از این تلفن را می پرسد. سپس، این ایستگاه کنترل تلفن مزبور را به ایستگاهی که بیشترین قدرت را گزارش کرده (و در واقع، تلفن وارد آن شده)، تحویل می دهد. رئیس جدید نیز به تلفن همراه اطلاع می دهد که (اگر مایل به ادامه مکالمه است) کانال خود را عوض کند (چون فرکانس قبلی آن در هیچیک از سلولهای همسایه وجود ندارد). این فرآیند (که به آن پاس کاری - handoff - گفته می شود) تقریباً 300 msec طول می کشد. تخصیص کانال توسط MTSO صورت می گیرد (چون ایستگاههای مرکزی چیزی جز رله های رادیویی ساده نیستند).

پاس کاری می تواند به دو طریق صورت گیرد. در پاس کاری نرم (soft handoff) تلفن همراه قبل از آن که ایستگاه قدیم آنرا قطع کند، به ایستگاه جدید متصل می شود. در این روش کاربر هیچ نوع قطعی احساس نمی کند. اما مشکل این روش آنست که دستگاه تلفن همراه باید بتواند در آن واحد خود را روی دو فرکانس (ایستگاه قبلی و ایستگاه جدید) تنظیم کند. تلفنهای همراه نسل اول و دوم هیچکدام قادر به انجام چنین کاری نیستند. در پاس کاری سخت (hard handoff) ایستگاه قدیمی قبل از اتصال تلفن همراه به ایستگاه جدید، آنرا قطع می کند. اگر ایستگاه جدید به هر دلیلی (مثلاً، نداشتن باند خالی) نتواند تلفن را تحویل بگیرد، ارتباط کاربر یکباره قطع خواهد شد. این یکی از مشکلات اجتناب ناپذیر تلفنهای همراه نسل اول و دوم است.

کانال ها

سیستم AMPS دارای ۸۳۲ کانال دو-طرفه همزمان است، که هر کانال خود از دو کانال یکطرفه ساده تشکیل می شود (۸۳۲ کانال دریافت و ۸۳۲ کانال ارسال). کانالهای یکطرفه دریافت روی فرکانسهای 824-849 MHz و کانالهای یکطرفه ارسال روی فرکانسهای 869-894 MHz کار می کنند (پهنای باند هر کانال 30 kHz است). همانطور که می بینید، AMPS از FDM برای تقسیم پهنای باند استفاده می کند.

امواج در فرکانس 800 MHz طول موجی در حدود 40 cm دارند، و به خط مستقیم حرکت می کنند. درختان و گیاهان این امواج را جذب می کنند، و در برخورد با زمین نیز منعکس می شوند. موجی که به ایستگاه مرکزی سلول می رسد، می تواند مستقیماً از گوشی تلفن همراه آمده باشد، و یا انعکاس آن از سطح زمین یا ساختمانها باشد، که این می تواند باعث پژواک (echo) یا محوشدگی چندمسیره (multipath fading) شود. گاهی حتی امکان شنیدن مکالماتی که در مسافتی طولانی چندین بار به سطح زمین خورده و منعکس شده، نیز وجود دارد.

۸۳۲ کانال AMPS به چهار دسته تقسیم می شوند:

۱. کانالهای کنترل (ایستگاه به تلفن همراه) برای مدیریت سیستم
۲. کانالهای فراخوانی (ایستگاه به تلفن همراه) برای اعلام اینکه مکالمه ای در انتظار کاربر است
۳. کانالهای دسترسی (دو طرفه) برای برقراری مکالمات و تخصیص کانال
۴. کانالهای داده (دو طرفه) برای صدا، فکس، یا داده

تعداد کانالهای کنترل ۲۱ عدد است، که بطور ثابت در PROM تلفنهای همراه نوشته شده است. از آنجائیکه استفاده از فرکانسهای مشابه در سلولهای مجاور مجاز نیست، تعداد کانالهای قابل استفاده در هر سلول بسیار کمتر از ۸۳۲ (و در واقع، چیزی نزدیک به ۴۵ کانال) است.

مدیریت مکالمه

هر تلفن همراه در سیستم AMPS دارای یک شماره سریال ۳۲ بیتی و یک شماره تلفن ۱۰ رقمی است، که در PROM آن نوشته شده است. شماره تلفن همراه دارای یک کد ناحیه ۳ رقمی (که بصورت ۱۰ بیتی کُد شده) و یک شماره مشترک ۷ رقمی (که بصورت ۲۴ بیتی کُد شده) - مجموعاً ۳۴ بیت - است.

وقتی تلفن روشن می شود، ۲۱ کانال کنترل از پیش برنامه ریزی شده را اسکن می کند، تا قویترین سیگنال را پیدا کند. سپس، شماره سریال ۳۲ بیتی و شماره تلفن ۳۴ بیتی خود را منتشر می کند. تمام اطلاعات کنترلی در AMPS دیجیتال هستند (برخلاف کانالهای صدا، که آنالوگ می باشند)، و این اطلاعات به دفعات و همراه با کدهای تصحیح خطا منتشر می شوند.

وقتی ایستگاه مرکزی سلول این اعلام را دریافت کرد، آنرا به MTSO می فرستد، که آن هم (پس از ثبت کاربر جدید) محل وی را به نزدیکترین MTSO اعلام می کند. در حالت عادی، یک تلفن همراه هر ۱۵ دقیقه خود را به ایستگاه مرکزی معرفی می کند.

برای برقراری یک تماس، کاربر (بعد از روشن کردن تلفن) شماره مورد نظر را وارد کرده، و دکمه CALL را فشار می دهد. تلفن این شماره را به همراه کد شناسایی خود روی یکی از کانالهای دسترسی (access channel) به ایستگاه مرکزی می فرستد. (اگر تداخلی پیش آید، تلفن این عملیات را بعداً تکرار می کند). وقتی ایستگاه مرکزی سلول این درخواست را دریافت کرد، به MTSO اطلاع می دهد. اگر تماس گیرنده یکی از مشترکین آن MTSO باشد، MTSO بدنبال یک کانال خالی می گردد تا به وی تخصیص دهد. اگر کانال خالی موجود بود، شماره آن روی کانال کنترل به تلفن برگشت داده می شود. با گرفتن کانال دسترسی از MTSO، تلفن همراه بطور خودکار به آن

کانال سوئیچ کرده و منتظر می ماند تا طرف مقابل گوشی را بردارد.

تماسهای ورودی به طریق دیگری عمل می کنند. تلفنهایی که بیکار هستند، دائماً به کانال فراخوانی (paging channel) گوش می کنند، تا پیامهایی را که برای آنان می رسد دریافت کنند. وقتی یک شماره تلفن همراه گرفته می شود (خواه از یک تلفن ثابت یا یک تلفن همراه دیگر)، یک بسته به MTSO ی آن تلفن ارسال می شود تا محل وی را پیدا کند. MTSO نیز که محل تمام مشترکین فعال خود را می داند، یک بسته به ایستگاه مرکزی آن تلفن می فرستد، که آن هم بنوبه خود یک پیام (با مضمون: «تلفن 14، تو آنجایی؟») روی کانال فراخوانی منتشر می کند. تلفن مقصد هم با ارسال پاسخ "Yes" روی کانال دسترسی جواب می دهد. وقتی ایستگاه مرکزی پاسخ "Yes" را دریافت کرد، با پیامی مانند این عکس العمل نشان می دهد: «تلفن 14، روی کانال 3 به تماس جواب بده.» در اینجا، تلفن همراه به کانال 3 سوئیچ کرده، و شروع به زنگ زدن (با نواختن یکی از ملودیهای عجیب و غریبی که این روزها همه جا شنیده می شود) می کند.

۲-۶-۲ تلفن های همراه نسل دوم: صدای دیجیتال

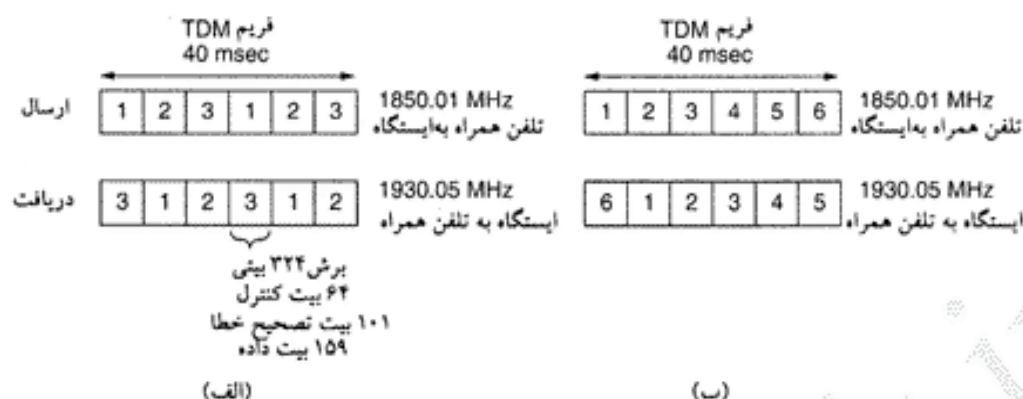
تلفن همراه نسل اول آنالوگ بود، و تلفن همراه نسل دوم دیجیتال. همانطور که نسل اول استاندارد جهانی نداشت، نسل دوم هم چنین استاندارد جهانی ندارد. امروزه چهار سیستم تلفن همراه نسل دوم مشغول بکار هستند: GSM، CDMA، D-AMPS و PDC. در این قسمت سه سیستم اول را بررسی خواهیم کرد؛ PDC فقط در ژاپن راه اندازی شده، و در واقع همان D-AMPS است که برای سازگاری با سیستمهای نسل اول آنالوگ ژاپن تغییراتی در آن صورت گرفته است. اصطلاح PCS (سرویسهای مخابرات شخصی - Personal Communications Services) از لحاظ فنی به معنای تلفن همراه دیجیتال با فرکانس 1900 MHz است، اما امروزه این کلمه عمدتاً به سیستمهای نسل دوم (یعنی، دیجیتال) تعبیر می شود.

D-AMPS : سیستم تلفن همراه پیشرفته دیجیتال

نسل دوم سیستمهای AMPS (که کاملاً دیجیتال است) D-AMPS (سیستم تلفن همراه پیشرفته دیجیتال - Digital Advanced Mobile Phone System) نام دارد. این سیستم تحت استاندارد بین المللی IS-54 (و خلف آن تحت استاندارد IS-136) تعریف شده است. سیستم D-AMPS بگونه ای طراحی شده که با AMPS سازگاری کامل داشته باشد، به همین دلیل تلفنهای نسل اول و دوم می توانند همزمان در یک سلول کار کنند. D-AMPS از همان کانالهای 30 kHz نسل اول (و دقیقاً با همان فرکانسها) استفاده می کند، بنابراین در یک سلول یک کانال می تواند آنالوگ کار کند، در حالیکه کانال مجاور در حال کار بصورت دیجیتال است. این MTSO ی مسئول سلول است که (بر اساس نسبت تلفنهای نسل اول و دوم در هر سلول) تعیین می کند کدام کانالها آنالوگ و کدام کانالها دیجیتال کار کنند - و (با تغییر این نسبت) می تواند ترکیب کانالها را بصورت دینامیک تغییر دهد.

بعد از عملیاتی شدن سرویسهای D-AMPS و بمنظور پاسخگویی به تقاضای رو به فزونی بازار، باند فرکانسی جدیدی به آن تخصیص داده شد. کانالهای دریافت از کاربر (upstream) در محدوده 1850-1910 MHz، و کانالهای ارسال به کاربر (downstream) در محدوده 1930-1990 MHz هستند. طول موج در این باند فقط 16 cm است، بنابراین آنتن استاندارد (آنتن یک چهارم طول موج) تلفنهایی که در این باند کار می کنند، فقط 4 cm خواهد بود، و به همین دلیل این تلفنها می توانند بسیار کوچکتر ساخته شوند. با این حال، اغلب تلفنهای D-AMPS هر دو باند فرکانسی (850 MHz و 1900 MHz) را پشتیبانی می کنند.

در یک تلفن همراه D-AMPS، سیگنال خروجی میکروفون بعد از دیجیتالی شدن با استفاده از مدلهایی که بسیار بهینه تر از مدولاسیون دلتا (delta modulation) و کد کردن پیشگویانه (predictive encoding) هستند،



شکل ۲-۴۲. (الف) یک کانال D-AMPS با سه کاربر. (ب) یک کانال D-AMPS با شش کاربر.

فشرده می شود. این مدل های فشرده سازی با استفاده از خصوصیات دستگاه صوتی انسان، پهنای باند لازم را از 56-kbps به 8 kbps یا کمتر تقلیل می دهند. این کار که توسط دستگاهی بنام vocoder انجام می شود (Bellamy, 2000) را ببینید)، در دستگاه تلفن صورت می گیرد نه در ایستگاه مرکزی، که بدین ترتیب تعداد بیت های منتقل شده روی لینک هوایی کاهش خواهد یافت. در تلفن های ثابت این تکنیک هیچ مزیتی بدنبال ندارد، چون کاستن از ترافیک مدار پایانی هیچ تأثیری روی ظرفیت کل سیستم نخواهد داشت.

فشرده کردن صدای دیجیتال تلفن های همراه (همان کاری که در D-AMPS انجام می شود) مزیت فوق العاده ای به همراه دارد، چون می توان با استفاده از حالتی پلکس تقسیم زمانی (TDM) یک زوج فرکانسی (ارسال و دریافت) را بین سه کاربر به اشتراک گذاشت. هر زوج فرکانسی از 25 frame/sec (هر فریم معادل 40 msec) پشتیبانی می کند. هر فریم نیز بنوبه خود به شش بُرش زمانی 6.67 msec تقسیم می شود (شکل ۲-۴۲ (الف) را ببینید).

هر فریم به سه کاربر سرویس (ارسال و دریافت) می دهد، که به نوبت از آن استفاده می کنند. برای مثال در بُرش زمانی 1 در شکل ۲-۴۲ (الف)، کاربر 1 می تواند در حال ارسال به ایستگاه مرکزی باشد، در حالیکه در همان زمان کاربر 3 در حال دریافت است. هر بُرش زمانی طولی معادل 324 بیت دارد، که 64 بیت آن برای مصارف کنترلی اختصاص یافته، و بقیه 260 بیت در اختیار کاربر است. از این 260 بیت، 101 بیت برای گدهای تصحیح خطا (که در لینک های هوایی پرنویز بسیار هم لازمند) بکار می رود، که بدین ترتیب فقط 159 بیت برای انتقال صدا باقی می ماند. با احتساب 50 بُرش زمانی در هر ثانیه، پهنای باند موجود برای صدای فشرده فقط 8 kbps (پهنای باند PCM) است.

با استفاده از الگوریتم های فشرده سازی بهتر، می توان پهنای باند لازم را حتی به 4 kbps کاهش داد، تا شش کاربر بتوانند در آن واحد از یک فریم استفاده کنند (شکل ۲-۴۲ ب). از دید شرکتهای تلفن، توانایی فشرده کردن سه یا شش کاربر D-AMPS در کانالی که فقط یک کاربر AMPS می تواند از آن استفاده کند، یک بُرد واقعی است و دلیل محبوبیت PCM نزد شرکتهای تلفن همراه نیز همین است. البته کیفیت صدای 4 kbps هرگز به پای صدای 56 kbps نمی رسد، اما شرکتهای تلفن همراه هم کمتر روی کیفیت صدای سرویس های خود تبلیغ می کنند. در مورد سرویس های داده، کیفیت یک خط 8 kbps حتی با مودم های 9600-bps نیز قابل مقایسه نیست.

ساختار کنترلی D-AMPS نسبتاً پیچیده است: هر ۱۶ فریم تشکیل یک ابرفریم (superframe) می دهند، که اطلاعات کنترلی بدفعات محدود در این ابرفریم گنجانده می شوند. در کل ۶ کانال کنترلی مورد استفاده قرار

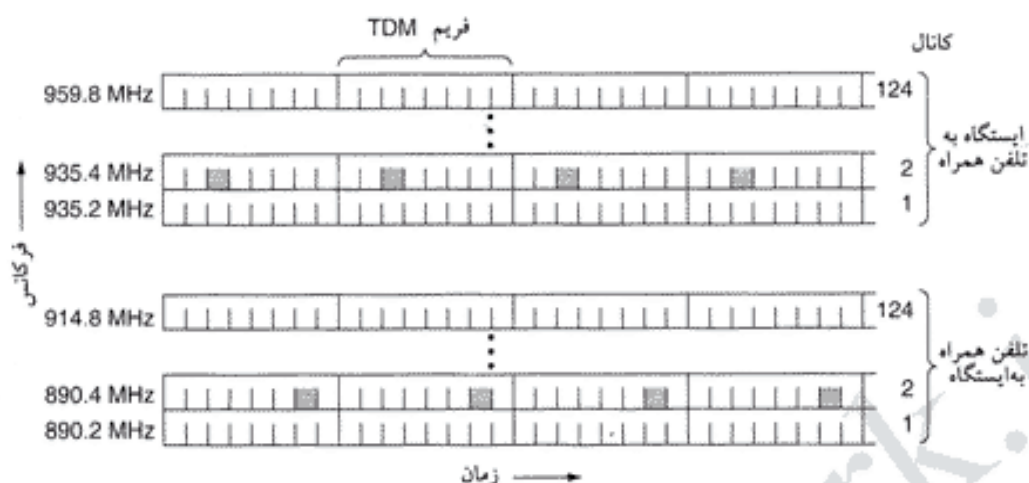
می‌گیرد: پیکربندی سیستم، کنترل بی‌درنگ (real-time) و غیر بی‌درنگ (nonreal-time)، فراخوانی (paging)، پاسخ دسترسی (access response) و پیام کوتاه (short message). طرز کار D-AMPS از نظر مفهومی شبیه AMPS است. وقتی گوشی تلفن همراه روشن می‌شود، با ایستگاه مرکزی تماس گرفته و بعد از معرفی خود، به کانالهای کنترلی گوش می‌کند. در این سیستم هم تعیین محل (سلول) تلفن همراه بر عهده MTSO است. یکی از تفاوت‌های D-AMPS با AMPS در نحوه پاس‌کاری است. در AMPS این کار بطور کامل بر عهده MTSO است و دستگاه تلفن هیچ دخالتی در آن ندارد. اما همانطور که در شکل ۲-۴۲ دیده می‌شود، در D-AMPS در $\frac{1}{3}$ از کل زمان تماس دستگاه تلفن هیچ کاری انجام نمی‌دهد، و می‌تواند از این زمان خالی برای اندازه‌گیری کیفیت خط استفاده کند. وقتی تلفن متوجه می‌شود که قدرت سیگنال در حال کاهش است، به MTSO اطلاع داده و MTSO هم ارتباط را قطع می‌کند، که در این لحظه تلفن می‌تواند به یک ایستگاه (فرکانس) قویتر سوییچ کند - به این تکنیک MAHO (پاس‌کاری یکمک تلفن همراه - Mobile Assisted HandOff) می‌گویند. در D-AMPS نیز (مانند AMPS) این فرآیند حدوداً 300 msec طول می‌کشد.

GSM : سیستم سراسری مخابرات همراه

سیستم D-AMPS بصورت گسترده در ایالات متحده (و بصورت اصلاح شده در ژاپن) در حال استفاده است. اما تقریباً در تمام نقاط دیگر دنیا از سیستمی بنام GSM (سیستم سراسری مخابرات همراه - Global System for Mobile Communications) استفاده می‌کنند، و حتی در آمریکا نیز این سیستم در نقاط محدودی راه‌اندازی شده است. GSM و D-AMPS از چند نظر شبیه یکدیگرند: هر دوی آنها سیستمهای سلولی هستند؛ هر دوی آنها از FDM و فرکانسهای متفاوت برای ارسال و دریافت (که فرکانس دریافت تلفن بیشتر از فرکانس ارسال آن است - 80 MHz بالاتر در D-AMPS و 55 MHz بالاتر در GSM) استفاده می‌کنند؛ و در هر دو سیستم چندین تلفن با استفاده از TDM مشترکاً از یک زوج فرکانس استفاده می‌کنند. اما در GSM پهنای کانالها بسیار بیشتر از D-AMPS (200 kHz در مقابل 30 kHz) و تعداد کاربران هر کانال نیز نسبتاً کمتر است (8 کاربر در مقابل 3 کاربر)، که در نتیجه نرخ داده آن بسیار بهتر از D-AMPS خواهد بود. در زیر توضیح مختصری درباره ویژگیهای اصلی GSM خواهیم داد. اما توجه داشته باشید که استاندارد چاپی GSM متجاوز از ۵۰۰۰ برگ است، و در آن جنبه‌های فنی و مهندسی این سیستم به تفصیل تشریح شده، که ما در اینجا بدانها نخواهیم پرداخت.

همانطور که قبلاً هم گفته شد (و در شکل ۲-۴۳ می‌بینید)، هر باند فرکانسی 200 kHz پهنای دارد. هر سیستم GSM دارای ۱۲۴ زوج کانال یکطرفه ساده (simplex) است، که هر کانال 200 kHz پهنای دارد و از ۸ ارتباط همزمان (بصورت مالتی پلکس تقسیم زمانی) پشتیبانی می‌کند. به هر ایستگاه فعال در هر بُرش زمانی یک زوج کانال اختصاص داده می‌شود، که بدین ترتیب تعداد کانالهای موجود به 992 می‌رسد. البته (بدلیل جلوگیری از تداخل فرکانسی ایستگاههای مجاور) این تعداد کانال برای تمام ایستگاهها قابل استفاده نیست. در شکل ۲-۴۳، هشت بُرش زمانی خاکستری می‌بینید که همگی متعلق به یک تماس هستند (۴ تا برای ارسال، و ۴ تا برای دریافت). ارسال و دریافت همزمان انجام نمی‌شود، چون دستگاههای رادیویی GSM برای سوییچ کردن فرکانس به زمان نیاز دارند، و نمی‌توانند همزمان هر دو کار را انجام دهند. اگر به یک ایستگاه فرکانس 890.4/935.4 MHz اختصاص داده شده باشد و بُرش زمانی 2 بخواد چیزی به این ایستگاه بفرستد، باید از بُرشهای خاکستری رنگ قسمت پائین تصویر (به هر تعداد که نیاز دارد) برای کار خود استفاده کند.

بُرشهای TDM نشان داده شده در شکل ۲-۴۳ بخشی از یک سلسله مراتب پیچیده فریم‌بندی هستند. هر بُرش TDM دارای ساختار خاصیتیست که ترکیب آنها نیز به روشی خاص تشکیل یک فریم چندگانه می‌دهد. در شکل ۲-۴۴ شکل ساده شده‌ای از این ساختار سلسله مراتبی را ملاحظه می‌کنید. همانطور که در این شکل



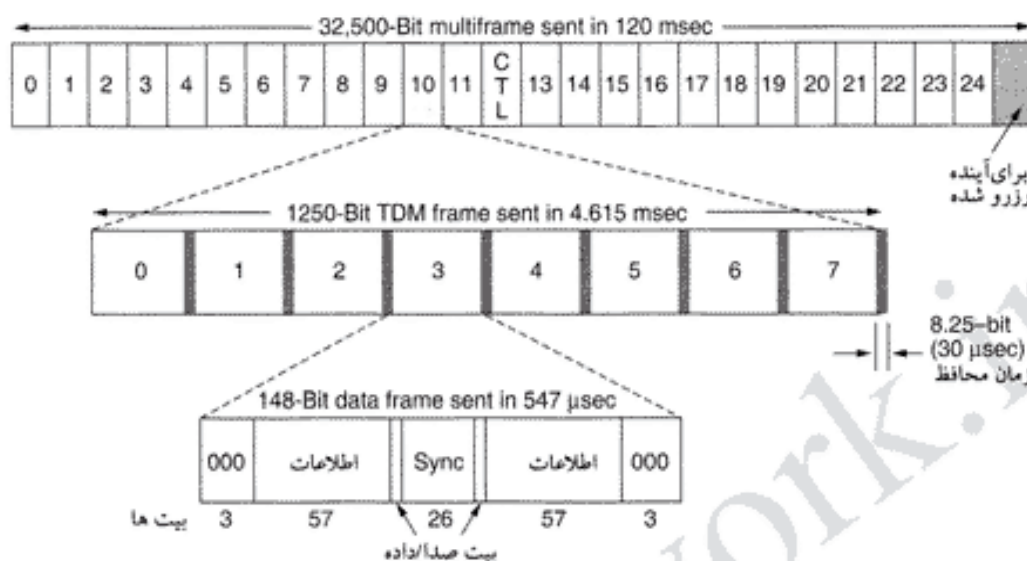
شکل ۲-۲۳. GSM از ۱۲۴ کانال فرکانسی استفاده می کند، که هر یک از آنها به ۸ بُرش زمانی تقسیم می شود.

می بینید، هر بُرش TDM از یک فریم داده ۱۴۸ بیتی تشکیل می شود، که کانال را برای مدت $577 \mu\text{sec}$ (شامل $30\text{-}\mu\text{sec}$ زمان حفاظت در هر بُرش) به اشغال خود در می آورد. بمنظور همزمان کردن فریمها، هر فریم داده با ۳ بیت 0 شروع و ختم می شود. هر فریم دو فیلد اطلاعاتی (Information) ۵۷ بیتی دارد، که هر کدام دارای ۱ بیت کنترلی هستند که مشخص می کند این فیلد اطلاعاتی حاوی داده است یا صدا. بین این فیلدهای اطلاعاتی یک فیلد نظم دهنده (Sync) ۲۶ بیتی وجود دارد که گیرنده از آن برای همزمان شدن با فرستنده استفاده می کند. ارسال هر فریم داده $547 \mu\text{sec}$ طول می کشد، ولی از آنجائیکه هر کانال بین ۸ ایستگاه به اشتراک گذاشته شده، فرستنده فقط در هر 4.615 msec می تواند یک فریم بفرستد. نرخ داده (ناخالص) هر کانال $270,833 \text{ bps}$ است که ۸ کاربر مشترکاً از آن استفاده می کنند. با یک تقسیم ساده مشخص می شود که نرخ داده ناخالص هر کاربر 33.854 kbps (بیش از دو برابر سیستمهای D-AMPS، یعنی 16.2 kbps) است. اما در اینجا هم مانند AMPS سرآیندها بخش زیادی از پهنای باند را می بلعند، و در نهایت 24.7 kbps برای داده واقعی کاربر (قبل از تصحیح خطا) باقی می ماند. آنچه پس از تصحیح خطا برای صدا باقی می ماند، 13 kbps است که باز هم کیفیت صدای بسیار بهتری نسبت به D-AMPS بدست می دهد (البته به قیمت مصرف پهنای باند بسیار بیشتر).

همانطور که در شکل ۲-۲۴ می بینید، هر ۸ فریم تشکیل یک فریم TDM و هر ۲۶ فریم TDM تشکیل یک فریم چندگانۀ 120-msec را می دهند. از ۲۶ فریم TDM، فریم ۱۲ برای کارهای کنترلی مورد استفاده قرار می گیرد، فریم ۲۵ نیز برای مصارف آتی کنار گذاشته شده است، و فقط ۲۴ فریم برای ارسال و دریافت باقی می ماند.

علاوه بر ساختار ۲۶ فریمی شکل ۲-۲۴، در GSM از ساختار دیگری با ۵۱ فریم نیز استفاده می شود. در این ساختار نیز برخی از فریمها کارکرد کنترلی (مدیریت سیستم) دارند. کانال کنترل پخش (broadcast control channel) یکی از این کانالهاست، که استریم خروجی پیوسته ایست که از ایستگاه مرکزی منتشر می شود و شامل هویت ایستگاه و اطلاعاتی درباره وضعیت کانالهای آن می شود. تمام دستگاههای تلفن همراه دائماً قدرت این سیگنال را چک می کنند تا بتوانند موقعیت خود را در میان سلولها تشخیص دهند.

کانال کنترل اختصاصی (dedicated control channel) یکی دیگر از کانالهای کنترلی است که برای به روز در آوردن موقعیت تلفنهای همراه، ثبت آنها و برقراری تماس بکار می رود. بویژه، هر ایستگاه مرکزی دارای پایگاه



شکل ۲-۴۴. بخشی از ساختار فریم بندی GSM.

داده‌ای از تمام تلفنهای موجود در محدوده قدرت خود است، که با استفاده از اطلاعاتی که روی این کانال فرستاده می‌شود، به روز در می‌آید.

دیگر کانال کنترلی، کانال کنترل مشترک (common control channel) است که به سه زیرکانال تقسیم می‌شود. اولین زیرکانال، کانال فراخوانی (paging channel) است که ایستگاه مرکزی از آن برای اعلام تماس ورودی استفاده می‌کند (و هر تلفن دائماً این کانال را چک می‌کند تا ببیند آیا تماسی دارد یا خیر). زیرکانال دوم، کانال دسترسی تصادفی (random access channel) است، که اجازه می‌دهد تا کاربران درخواستی برای یک بُرش از کانال کنترلی اختصاصی را روی آن ارسال کنند؛ کاربران از این بُرش برای برقراری تماس با دیگران استفاده می‌کنند. اگر دو کاربر همزمان وارد این کانال شوند، تداخل پیش می‌آید و باید (پس از کمی تأخیر) عملیات را از نو تکرار کنند. بُرش اختصاص داده شده به کاربر روی زیرکانال سوم، کانال اختصاص دسترسی (access grant channel)، به وی اعلام می‌شود.

CDMA: دسترسی چندگانه با تقسیم کد

سیستمهای GSM و D-AMPS هر دو سیستمهایی نسبتاً سستی و معمولی هستند، که از FDM و TDM برای تقسیم طیف فرکانسی به کانالها و تقسیم کانالها به بُرشهای زمانی استفاده می‌کنند. اما بازیگر سومی بنام CDMA (دسترسی چندگانه با تقسیم کد - Code Division Multiple Access) نیز در این صحنه حاضر است، که به روشی کاملاً متفاوت بازی می‌کند. وقتی CDMA برای اولین بار در صنعت مطرح شد، همان واکنشی را برانگیخت که پیشنهاد کریستف کلمب برای رسیدن به هندوستان از راه سفر به غرب، اما در نتیجه پایداری و مقاومت یک شرکت بنام Qualcomm، اکنون CDMA به جایی رسیده که نه تنها بعنوان یک سیستم قابل قبول مطرح است، بلکه به آن به چشم تنها مبنای مطمئن برای سیستمهای تلفن همراه نسل سوم نگاه می‌کنند. در ایالات متحده، CDMA حتی اکنون (در سیستمهای نسل دوم) نیز بعنوان رقیبی جدی برای D-AMPS مطرح است - برای مثال، شرکت Sprint PCS از CDMA استفاده می‌کند، در حالیکه AT&T Wireless از D-AMPS. استاندارد CDMA در سند IS-95 تدوین شده، و گاهی به این نام هم شناخته می‌شود؛ نام cdmaOne نیز یکی از نامهای تجاری آن است.

CDMA بکلی از AMPS، L-AMPS و GSM متفاوت است: بجای تقسیم طیف فرکانسی به کانالهای باریک، CDMA اجازه می دهد تا تمام تلفنهای و ایستگاهها از تمام طیف فرکانسی برای ارسال و دریافت استفاده کنند. و برای تفکیک آنها از یکدیگر از تئوری رمزگذاری (coding theory) استفاده می کند. در سیستمهای CDMA این فرض که فریمهای تداخل کرده غیر قابل استفاده اند را نیز بکلی کنار می گذارد. و بجای آن فرض می کند که این سیگنالها بصورت خطی با هم جمع می شوند.

برای درک بهتر CDMA از یک مثال آشنا استفاده می کنیم: سالتی پر از جمعیت که دو به دو مشغول صحبت هستند. TDM مانند آن است که این افراد را وسط سالن جمع کنید، ولی فقط به نوبت به آنها اجازه صحبت کردن بدهید. FDM مانند آن است که این افراد را با فاصله زیاد از یکدیگر بچینید. و به آنها اجازه دهید دو به دو همزمان (و البته مستقل) با هم صحبت کنند. CDMA نیز مانند این است که همه افراد را وسط سالن جمع کنید، ولی آنها (همزمان) به زبانهای مختلف با هم صحبت کنند. برای مثال، دو نفری که به فرانسه با هم حرف می زنند، هر چیزی غیر از کلمات فرانسوی را بعنوان پارازیت (نویز) نشنیده می گیرند. نکته کلیدی در CDMA همین استخراج سیگنال مورد نظر (و دور ریختن هر چیز دیگر) است. در زیر روش کار یک سیستم ساده شده CDMA را توضیح می دهیم.

در CDMA، هر بیت به m فاصله زمانی کوتاه، موسوم به چیب (chip)، تقسیم می شود معمولاً هر بیت دارای ۶۴ یا ۱۲۸ چیب است، ولی در اینجا برای سادگی هر بیت را به فقط ۸ چیب تقسیم کرده ایم. به هر تلفن همراه (یا ایستگاه) یک کد m بیتی منحصر بفرد، که به آن توالی چیب (chip sequence) می گویند، اختصاص داده می شود برای ارسال بیت 1، ایستگاه توالی چیب خود را می فرستد، و برای ارسال بیت 0 مکمل یک (one's complement) توالی چیب خود را - یک ایستگاه مجاز نیست هیچ چیز دیگری بفرستد. برای مثال، با فرض $m = 8$ ، اگر ایستگاهی بنام A دارای توالی چیب 00011011 باشد، برای ارسال بیت 1 توالی 00011011 و برای ارسال 0 توالی 11100100 را می فرستد.

افزایش مقدار اطلاعات ارسالی از b bits/sec به mb chips/sec فقط وقتی امکانپذیر است که پهنای باند موجود m برابر شود، که بدین ترتیب CDMA به سیستمی با طیف گسترده تبدیل می شود. اگر پهنای باند موجود برای ۱۰۰ ایستگاه 1 MHz باشد، با تکنیکهای FDM هر ایستگاه فقط 10 kHz در اختیار خواهد داشت و می تواند حداکثر 10 kbps (با فرض 1 bit/Hz) ارسال کند. با CDMA هر ایستگاه می تواند از تمامی پهنای باند 1-MHz استفاده کند، و به سرعت 1 megachips/sec برسد. اگر تعداد چیب بر بیت کمتر از ۱۰۰ باشد، پهنای باند مؤثر CDMA باز هم بیشتر از FDM است (و مشکلات تخصیص کانال نیز دیگر وجود ندارد).

در این قسمت برای درک بهتر مطلب روش دو-علامتی را بکار می بریم، یعنی بجای 0 باینری از -1 و بجای 1 باینری از +1 استفاده خواهیم کرد. توالی چیبها را نیز در پرانتز نمایش می دهیم، که بدین ترتیب توالی چیب ایستگاه A به $(+1 +1 -1 -1 +1 -1 -1 +1)$ تبدیل می شود. در شکل ۲-۴۵ (الف) توالی چیب چهار ایستگاه بنامهای A، B، C و D را می بینید؛ در شکل ۲-۴۵ (ب) نیز همین توالیها را به روش دو-علامتی نشان داده ایم. هر ایستگاه توالی چیب خاص خود را دارد. فرض می کنیم S بردار m -چیپی ایستگاه S، و \bar{S} بردار منفی (مکمل یک) آن است. تمام بردارهای توالی چیب متعامد (orthogonal) هستند، بعبارت دیگر ضرب داخلی نرمال شده هر دو بردار S و T (S • T) صفر است. با استفاده از روشی بنام کدهای والش (Walsh codes) می توان چنین بردارهایی تولید کرد. متعامد بودن دو بردار را به زبان ریاضی می توان چنین نوشت:

$$S \cdot T \equiv \frac{1}{m} \sum_{i=1}^m S_i T_i = 0$$

همانطور که خواهید دید، متعامد بودن بردارها یکی از ویژگیهای کلیدی این سیستم است. توجه داشته باشید که اگر $S \cdot T = 0$. آنگاه $S \cdot \bar{T} = 0$. ضرب داخلی نرمال شده هر بردار در خودش نیز 1 است:

$$S \cdot S = \frac{1}{m} \sum_{i=1}^m S_i S_i = \frac{1}{m} \sum_{i=1}^m S_i^2 = \frac{1}{m} \sum_{i=1}^m (\pm 1)^2 = 1$$

از آنجائیکه هر یک از جملات این دنباله 1 است، مجموع m جمله آن معادل m می شود، که بعد از ضرب در $\frac{1}{m}$ به عدد 1 خواهیم رسید. همچنین توجه داشته باشید که، $S \cdot \bar{S} = -1$.

همانطور که گفتیم، هر ایستگاه می تواند برای ارسال بیت 1 توالی چیب خود، و برای ارسال بیت 0 مکمل یک توالی چیب خود را بفرستد، و یا اصلاً سکوت کند و چیزی نفرستد. فعلاً فرض را بر این می گذاریم که تمام ایستگاهها سکوت هستند، یعنی ارسال توالی های خود را در یک زمان شروع می کنند.

وقتی دو یا چند ایستگاه بطور همزمان شروع به ارسال می کنند، سیگنال دو-علامتی آنها بصورت خطی با هم جمع می شود. برای مثال، اگر (در دوره زمانی یک چیب) سه ایستگاه سیگنال +1 و یک ایستگاه سیگنال -1 بفرستند، مجموع آنها +2 خواهد شد. می توان این وضعیت را مانند جمع ولتاژها تصور کرد: مجموع سه ولتاژ +1 و یک ولتاژ -1 معادل +2 ولت می شود. در شکل ۲-۴۵ (ج) شش نمونه از ارسال همزمان یک یا چند ایستگاه را ملاحظه می کنید. در مثال اول، ایستگاه C مبادرت به ارسال یک بیت 1 می کند، عبارت دیگر توالی چیب خود (01011100) را می فرستد. در مثال دوم، هر دو ایستگاه B و C یک بیت 1 می فرستند، که بدین ترتیب مجموع توالی های دو-علامتی آنها چنین خواهد شد:

$$(-1 -1 +1 -1 +1 +1 +1 -1) + (-1 +1 -1 +1 +1 +1 -1 -1) = (-2 0 0 0 +2 +2 0 -2)$$

A: 00011011

B: 00101110

C: 01011100

D: 01000010

(الف)

A: (-1 -1 -1 +1 +1 -1 +1 +1)

B: (-1 -1 +1 -1 +1 +1 +1 -1)

C: (-1 +1 -1 +1 +1 +1 -1 -1)

D: (-1 +1 -1 -1 -1 -1 +1 -1)

(ب)

شش مثال

--1- C

-11- B+C

10-- A+B

101- A+B+C

1111 A+B+C+D

1101 A+B+C+D

 $S_1 = (-1 +1 -1 +1 +1 +1 -1 -1)$ $S_2 = (-2 0 0 0 +2 +2 0 -2)$ $S_3 = (0 0 -2 +2 0 -2 0 +2)$ $S_4 = (-1 +1 -3 +3 +1 -1 -1 +1)$ $S_5 = (-4 0 -2 0 +2 0 +2 -2)$ $S_6 = (-2 -2 0 -2 0 -2 +4 0)$

(ج)

 $S_1 \cdot C = (1 +1 +1 +1 +1 +1 +1)/8 = 1$ $S_2 \cdot C = (2 +0 +0 +0 +2 +2 +0 +2)/8 = 1$ $S_3 \cdot C = (0 +0 +2 +2 +0 -2 +0 -2)/8 = 0$ $S_4 \cdot C = (1 +1 +3 +3 +1 -1 +1 -1)/8 = 1$ $S_5 \cdot C = (4 +0 +2 +0 +2 +0 -2 +2)/8 = 1$ $S_6 \cdot C = (2 -2 +0 -2 +0 -2 -4 +0)/8 = -1$

(د)

شکل ۲-۴۵. (الف) توالی چیب باینری چهار ایستگاه. (ب) توالی چیب دو-علامتی همان

ایستگاهها. (ج) شش نمونه از ارسال همزمان ایستگاهها. (د) استخراج سیگنال ایستگاه C.

در مثال سوم، ایستگاه A یک بیت 1 و ایستگاه B یک بیت 0 می‌فرستد، و بقیه ایستگاهها ساکت هستند. در مثال چهارم، ایستگاههای A و C یک بیت 1 و ایستگاه B یک بیت 0 می‌فرستند. در مثال پنجم، هر چهار ایستگاه یک بیت 1 می‌فرستند؛ و بالاخره در مثال آخر، ایستگاههای A ، B و D یک بیت 1 می‌فرستند و ایستگاه C یک بیت 0. توجه داشته باشید که در تمام مثالهای فوق، توالی‌های S_1 تا S_6 فقط یک بیت را نشان می‌دهند.

برای تشخیص و استخراج استریم بیت‌های یک ایستگاه باید توالی چیپ آن ایستگاه را از قبل بدانیم. این کار را می‌توان با محاسبه ضرب داخلی نرمال شده توالی چیپ دریافت شده (جمع خطی سیگنال تمام ایستگاهها در آن لحظه) با توالی چیپ ایستگاه موردنظر انجام داد. اگر توالی دریافت شده را S بنامیم، و بخواهیم توالی چیپ (سیگنال ارسالی) ایستگاه C را از آن بیرون بکشیم، کافیست ضرب داخلی نرمال شده $S \cdot C$ را محاسبه کنیم.

برای اینکه ببینید این روش چگونه کار می‌کند، مثال چهارم شکل ۲-۴۵ (ج) را در نظر بگیرید. چیزی که گیرنده دریافت می‌کند، $S = A + \bar{B} + C$ است، و برای استخراج سیگنال C بایستی عبارت $S \cdot C$ را محاسبه کند:

$$S \cdot C = (A + \bar{B} + C) \cdot C = A \cdot C + \bar{B} \cdot C + C \cdot C = 0 + 0 + 1 = 1$$

همانطور که می‌بینید، دو جمله اول (بدلیل متعامد بودن بردارها) صفر شده‌اند. از اینجا می‌توانید علت انتخاب بردارهای متعامد را دریابید.

این وضعیت را می‌توان بصورت سه سیگنال مستقل نیز در نظر گرفت: سیگنال‌ها بصورت مستقل و جدا از هم دریافت شده، و پس از محاسبه ضرب داخلی، با هم جمع می‌شوند. بعلت متعامد بودن بردارها، تمام ضربهای داخلی (بجز $C \cdot C$) صفر خواهند شد. بعبارت دیگر، تقدم ضرب داخلی یا جمع تأثیری بر نتیجه نهایی ندارد.

برای درک بهتر روش از رمز خارج کردن سیگنالها به شکل ۲-۴۵ (د) نگاه کنید. فرض کنید گیرنده می‌خواهد از شش سیگنال S_1 تا S_6 بیت‌های ارسال شده از ایستگاه C را استخراج کند. برای این کار، گیرنده تک تک سیگنالهای S را در بردار C (شکل ۲-۴۵ ب) ضرب کرده، و سپس $1/8$ آنرا محاسبه می‌کند (چون، $m = 8$). همانطور که می‌بینید، گیرنده توانسته است بطور صحیح بیتهای ارسالی از C را استخراج کند.

در یک سیستم ایده‌آل (بدون نویز) CDMA تعداد ایستگاهها را می‌توان به تعداد دلخواه زیاد کرد (همانگونه که در یک کانال بدون نویز نایکونیست می‌توان نرخ نمونه برداری را بدلتخواه افزایش داد). اما در عمل، محدودیتهای فیزیکی ظرفیت سیستم را بشدت پائین می‌آورند. اول اینکه، فرض کردیم تمام چیپها از نظر زمانی سنکرون هستند، در حالیکه این وضعیت عملاً غیرممکن است. بهترین کاری که می‌توان کرد اینست که فرستنده با ارسال یک توالی چیپ از پیش تعریف شده (که باندازه کافی طولانی است) به گیرنده امکان دهد تا خود را با فرستنده سنکرون کند. در این حالت تمام سیگنالهای غیرسنکرون بعنوان نویز تلقی خواهند شد. اگر تعداد این سیگنالهای غیرسنکرون چندان زیاد نباشد، الگوریتم فوق همچنان بخوبی کار خواهد کرد. در زمینه بر هم نهی (superposition) توالی‌های چیپ با سطح نویز تحقیقات تئوریک مفصلی انجام شده است (Pickholtz et al., 1982). همانطور که می‌توان انتظار داشت، هر چه توالی چیپ طولانی تر باشد، احتمال استخراج آن در محیطهای پرنویز بیشتر خواهد بود. حتی می‌توان در توالی بیت‌ها از گداهای تصحیح خطا نیز استفاده کرد - البته در توالی چیپ هرگز از گداهای تصحیح خطا استفاده نمی‌شود.

یکی دیگر از مفروضات ضمنی بحث فوق یکسان بودن قدرت سیگنالهاست که به گیرنده می‌رسند. در سیستمهای تلفن همراه (و از جمله CDMA) فاصله تلفنهای همراه از ایستگاه مرکزی (و در نتیجه قدرت تشعشعی آنها) متغیر است، و قدرت سیگنالهایی که از تلفنها به ایستگاه مرکزی می‌رسد یکسان نیست. یکی از روشهای ابتکاری برای حل این مشکل آنست که تلفنهای همراه با کاهش سطح سیگنال دریافتی از ایستگاه مرکزی، توان تشعشعی خود را بالا ببرند؛ بعبارت دیگر، هر چه از ایستگاه مرکزی دور می‌شوند، سیگنال قویتری بفرستند.

ایستگاه مرکزی نیز می‌تواند صریحاً به تلفنهای همراه فرمان دهد تا (بسته به فاصله‌شان از ایستگاه) قدرت سیگنالهای خود را افزایش یا کاهش دهند.

همچنین فرض کردیم که گیرنده از هویت فرستنده آگاه است. از نظر تئوری این امکان هست که (با فرض وجود قدرت محاسباتی)، یک گیرنده به تمام سیگنالها گوش دهد و الگوریتمهای از رمز خارج کردن را روی تمام آنها اجرا کند. اما مثل معروفی هست که می‌گوید، «حرف زدن همیشه از عمل کردن ساده‌تر است.» CDMA پیچیدگیهای دیگری نیز دارد، که در این بحث مختصر آنها را نادیده گرفتیم. با این حال، CDMA یکی از هوشمندانه‌ترین سیستمهای مخابرات بیسیم محسوب می‌شود، که سرعت در حال گسترش است. این سیستم معمولاً از یک باند 1.25 MHz (برخلاف 30 kHz در D-AMPS و 200 kHz در GSM) استفاده می‌کند، و تعداد کاربرانی که پشتیبانی می‌کند، از هر دو سیستم قبلی بیشتر است. در عمل، پهنای باند موجود برای هر کاربر در سیستمهای CDMA حتی از GSM هم بهتر است.

یکی از بهترین منابع موجود در این زمینه (Lee and Miller, 1998) است. روشهای دیگری نیز در (Crespo et al., 1995) و (Sari et al., 2000) تشریح شده‌اند، که البته فهم آنها به دانش زیادی در زمینه مهندسی مخابرات نیاز دارد.

۳-۶-۲ تلفن‌های همراه نسل سوم: صدای دیجیتال و داده

آینده تلفن همراه چیست؟ اجازه دهید نگاه سریعی به آن بیندازیم. صنعت تلفن همراه عوامل محرک و پیشبرنده متعددی دارد. اول، ترافیک داده مدتهاست از ترافیک صدا پیشی گرفته و همچنان به رشد تصاعدی خود ادامه می‌دهد، ولی ترافیک صدا سالهاست دیگر رشد چندانی را تجربه نمی‌کند. بسیاری از متخصصان پیش‌بینی می‌کنند که بزودی در تلفنهای همراه نیز شاهد پیشی گرفتن ترافیک داده از ترافیک صدا خواهیم بود. دوم، صنایع تلفن، تفریحات و کامپیوتر همگی دیجیتالی شده‌اند و روز بروز بیشتر به یکدیگر نزدیک می‌شوند. مردم مدتهاست با شور و هیجان از دستگاه کوچک، سبک و قابل حملی صحبت می‌کنند که بتواند بعنوان تلفن، پخش CD، پخش DVD، ترینال ایمیل و وب، وسیله بازی و سرگرمی، واژه‌پرداز (و غیره و غیره) کار کند، و قادر باشد در هر نقطه‌ای بصورت بیسیم و با پهنای باند بالا به اینترنت متصل شود. این دستگاه رؤیایی همان تلفن همراه نسل سوم است؛ برای اطلاعات بیشتر (Huber et al., 2000; 2000) and Sarikaya, (Huber et al., 2000; 2000) را ببینید.

در سال ۱۹۹۲، ITU کوشید این رؤیا را کمی بیشتر به واقعیت نزدیک کند، و به همین منظور طرح اولیه‌ای بنام IMT-2000 (که IMT مخفف سیستم مخابرات تلفن همراه بین‌المللی - International Mobile Telecommunication است) منتشر کرد. عدد 2000 سه چیز را نشان می‌داد: (۱) سالی که این سیستم بایستی عملیاتی شود، (۲) فرکانسی (بر حسب MHz) که سیستم تحت آن کار می‌کند، و (۳) پهنای باند این سرویس (بر حسب kHz).

البته IMT-2000 به هیچیک از این اهداف نرسید. در سال ۲۰۰۰ هیچ سیستمی نصب و راه‌اندازی نشد. ITU پیشنهاد کرده بود که دولتها طیف 2 GHz را برای این منظور کنار بگذارند تا این سرویس بتواند بین کشورهای مختلف بدون اشکال کار کند؛ تنها کشوری که به این توصیه عمل کرد، چین بود. و بالاخره، مشخص شد که در حال حاضر پهنای باند 2 Mbps برای کاربرانی که بیش از حد متحرک هستند، عملی نیست (علت آن هم دشواری پاس‌کاری این قبیل کاربران است). پهنای باند عملی‌تر عبارتست از 2 Mbps برای کاربران ثابت خانگی (که می‌تواند با ADSL رقابت کند)، 384 kbps برای کاربرانی که قدم می‌زنند، و 144 kbps برای آنهایی که سوار اتومبیل هستند. با این حال، حوزه فعالیت 3G (نسل سوم) بسیار پُر جنب و جوش است. نسل سوم شاید کمی دیرتر از راه برسد و کمی کمتر از آنچه که انتظار می‌رود باشد، ولی حتماً می‌آید.

سرویس‌هایی که قرار است شبکهٔ IMT-2000 در اختیار کاربران خود بگذارد، عبارتند از:

۱. انتقال صدا با کیفیت عالی.
۲. پیام‌رسانی (سرویسی که جایگزین ایمیل، فکس، SMS، chat و غیره خواهد شد).
۳. مالتی‌مدیا (پخش موسیقی، تماشای ویدئو، فیلم، تلویزیون و غیره).
۴. دسترسی اینترنت (گشت و گذار در وب، از جمله صفحاتی که صدا، تصویر و فیلم دارند)

سرویس‌های دیگری از قبیل کنفرانس ویدئویی (video conferencing)، حضور از راه دور (telepresence)، بازیهای گروهی و تجارت-همراه (m-commerce)؛ پرداخت بهای اجناس خریداری شده در فروشگاه با نزدیک کردن تلفن همراه به صندوق) را نیز می‌توان از این شبکه انتظار داشت. علاوه بر آن، تمام این سرویسها بین‌المللی خواهند بود (یعنی در تمام نقاط دنیا می‌توان به آنها دسترسی داشت: در جاهایی که خطوط ارتباطی زمینی وجود ندارد، تلفن همراه بطور خودکار از لینکهای ماهواره‌ای استفاده خواهد کرد)، و همیشه با کیفیت تضمین شده در دسترس هستند.

ITU برای IMT-2000 تکنولوژی واحدی را در نظر گرفته است، بگونه‌ایکه تلفنهای همراه بتوانند در هر نقطه‌ای از دنیا کار کنند (مانند دستگاههای ضبط صوت و پخش CD، نه تلویزیونها و تلفنهای همراه امروزی). یکسان شدن تکنولوژی، علاوه بر تسهیل کار شرکتهای مخابرات، افراد بیشتری را به استفاده از این سرویسها تشویق خواهد کرد. جنگ تکنولوژیها (مانند آنچه بین ویدئوهای Betamax و VHS رخ داد) هیچگاه به نفع صنعت و تجارت نبوده است.

پیشنادهای متعددی ارائه شد، که بعد از غربال شدن آنها، سرانجام دو طرح باقی ماند. طرح اول، بنام W-CDMA (CDMA پهن‌باند - Wideband CDMA)، از طرف شرکت سوئدی اریکسون ارائه شد. این طرح از توالی مستقیم با طیف گسترده (مانند آنچه در بالا گفتیم) استفاده می‌کند. این سیستم در یک باند 5-MHz کار می‌کند، و بگونه‌ای طراحی شده که بتواند با شبکه‌های GSM (البته نه GSM قدیمی) کار کند. تلفنهای این سیستم می‌توانند بدون اختلال در ارتباط یک سلول W-CDMA را ترک کرده و وارد یک سلول GSM شوند. اتحادیه اروپا بشدت از این سیستم پشتیبانی می‌کند، و به آن نام UMTS (سیستم مخابرات تلفن همراه جهانی - Universal Mobile Telecommunications System) را داده است.

طرح دیگر که از طرف شرکت Qualcomm (مخترع CDMA) پیشنهاد شده، CDMA2000 نام دارد. این سیستم نیز اساساً همان توالی مستقیم با طیف گسترده (شکل اصلاح شدهٔ استاندارد IS-95) است، که با آن سازگاری دارد. CDMA2000 نیز از یک باند 5-MHz استفاده می‌کند، ولی نمی‌تواند با سلولهای GSM (و طبیعتاً، با سلولهای D-AMPS) پاس‌کاری انجام دهد. CDMA2000 تفاوت‌های دیگری، از قبیل نرخ چیب، زمان فریم، طیف فرکانسی، و روش سنکرون شدن، نیز با W-CDMA دارد.

اگر مهندسان اریکسون و Qualcomm را در اتافی حبس کنند و از آنها بخواهند یک سیستم مشترک طراحی کنند، به احتمال زیاد می‌توانند. بهر حال، هر دو سیستم از تکنیکهای CDMA و یک کانال 5-MHz استفاده می‌کنند، و چیزهای دیگر هم مسلماً ارزش خودکشی ندارند. در اینجا هم (مثل همیشه) مشکل نه فنی و مهندسی بلکه سیاسی است. اروپا سیستمی می‌خواهد که با GSM سازگاری داشته باشد، و ایالات متحده آمریکا دنبال سیستمی است که با IS-95 سازگار باشد. هر دو نیز از شرکتهای محلی خود دفاع می‌کنند: اروپا از اریکسون (که شرکتهای سوئدی است) و آمریکا از Qualcomm (که در کالیفرنیا است). اریکسون و Qualcomm دعواهای حقوقی بیشماری نیز بر سر حق اختراع CDMA داشته‌اند.

بالاخره در مارس ۱۹۹۹ آریکسون Qualcomm را خرید، و دعوای خاتمه یافت. آنها بر سر یک استاندارد 3G نیز به توافق رسیدند، ولی این استاندارد ناسازگاری‌های زیادی داشت. اما این منازعات بالاخره به پایان می‌رسد، و بزودی شاهد تلفن‌ها و سرویس‌های 3G خواهیم بود.

در باره سیستم‌های 3G مطالب زیادی نوشته شده، و برخی آنرا بعنوان یکی از بزرگترین اختراعات بشری ستوده‌اند. برای نمونه به: (Collins and Smith, 2000; De Vriendt et al., 2002; Harte et al., 2002; 2002; Lu, and Sarikaya, 2000) اشاره می‌کنیم. اما 3G متقدانی نیز دارد که معتقدند اساساً راه را اشتباه می‌رود، و می‌توان از میان آنها از (Garber, 2002; and Goodman, 2000) نام برد.

عده‌ای هم که از جنگ 3G خسته شده‌اند، به فکر افتادند تا قدم کوچکی به سمت آن بردارند، و نام آنرا هم 2.5G گذاشتند (اگر چه شاید نام 2.1G برای آن مناسب‌تر است). یکی از این سیستم‌ها EDGE (نسخه داده بهبود یافته برای تکامل GSM - Enhanced Data rates for GSM Evolution) نام دارد، که اساساً چیزی نیست جز همان GSM با bits/ baud بیشتر. مشکل اینجا است که bits/ baud بیشتر یعنی خطای بیشتر، و به همین دلیل EDGE در سرعت‌های متفاوت از 9 روش مختلف برای مدولاسیون و تصحیح خطا استفاده می‌کند.

یکی دیگر از طرح‌های 2.5G، GPRS (سرویس عمومی بسته رادیویی - General Packet Radio Service) نام دارد، که عبارتست از یک شبکه سوئیچینگ بسته‌ای که روی GSM یا D-AMPS کار می‌کند. این سیستم اجازه می‌دهد تا تلفن‌های همراه در سلول‌های صوتی بسته‌های IP رد و بدل کنند. GPRS برای این منظور از بُرش‌های زمانی و فرکانس‌های اختصاصی استفاده می‌کند، که تعداد و محل آنها (به نسبت ترافیک صوت و داده در سلول) بطور دینامیک توسط ایستگاه مرکزی تعیین می‌شود.

بُرش‌های زمانی موجود به چندین کانال منطقی تقسیم شده، و به مصارف مختلف می‌رسند. تخصیص بُرش‌های زمانی به هر کانال توسط ایستگاه مرکزی صورت می‌گیرد. یکی از این کانال‌های منطقی برای ارسال بسته‌ها از ایستگاه مرکزی به تلفن‌های همراه است، و هر بسته می‌داند که مقصد آن کجاست. برای ارسال یک بسته IP، تلفن همراه درخواستی برای تخصیص یک یا چند بُرش زمانی به ایستگاه مرکزی می‌فرستد. اگر این درخواست بدون مشکل به ایستگاه مرکزی برسد، ایستگاه مرکزی فرکانس و بُرش‌های زمانی تخصیص یافته را به تلفن همراه اعلام می‌کند. همین که بسته‌های IP ارسالی از تلفن همراه به ایستگاه مرکزی رسید، ایستگاه مرکزی (از طریق ارتباطاتی که دارد) آنها را به اینترنت منتقل می‌کند. از آنجائیکه GPRS فقط لایه‌ایست روی سیستم‌های صوتی موجود، می‌توان آنرا در بهترین حالت یک قدم به سمت 3G بشمار آورد.

با اینکه شبکه‌های 3G هنوز بطور کامل راه‌اندازی نشده‌اند، برخی از محققان آنرا موضوعی تحقق یافته (که دیگر ارزش توجه ندارد) تلقی کرده و سراغ سیستم‌های نسل چهارم (4G) رفته‌اند. (Berezdivin et al., 2002; Guo and Chaskar, 2002; Huang and Zhuang, 2002; Kellerer et al., 2002; and Misra et al., 2002). برخی از مشخصات پیشنهادی سیستم‌های 4G عبارتند از: پهنای باند زیاد، دسترسی در همه جا، یکپارچگی کامل با شبکه‌های کابلی (و بویژه شبکه‌های IP)، مدیریت تطبیقی منابع و طیف فرکانسی، رادیویی نرم‌افزاری، و سرویس‌های مالتی‌مدیا با کیفیت عالی.

از سوی دیگر، با توجه به تعداد بسیار زیاد شبکه‌های محلی بیسیم 802.11 که امروزه نصب شده، برخی معتقدند 3G به دنیا نیامده، مرده است. اینها می‌گویند، «هر جا که بروید بالاخره تحت پوشش یک شبکه 802.11 هستید، و این همان چیز است که 3G می‌خواهد با آن همه منت به شما بدهد.» پنج سال دیگر معلوم می‌شود حق با چه کسی بوده است.

۷-۲ تلویزیون کابلی

تا اینجا سیستم‌های تلفن ثابت و بیسیم را بررسی کردیم. هر دوی این سیستمها مسلماً نقش مهمی در آینده شبکه بازی خواهند کرد. اما بازیگر دیگری نیز در زمینه شبکه‌های ثابت وارد صحنه شده، و هر روز اهمیت بیشتری می‌یابد: **تلویزیون کابلی (Cable TV)**. امروزه افراد بسیاری سرویسهای تلویزیون و اینترنت خود را از طریق کابل دریافت می‌کنند، و شرکت‌های تلویزیون کابلی نیز با جدیت بدنبال سهم بیشتری از بازار هستند. در این قسمت تلویزیون کابلی را از دیدگاه شبکه مورد بررسی دقیقتر قرار خواهیم داد. برای کسب اطلاعات بیشتر نیز می‌توانید به (Laubach et al., 2001; Louis, 2002; Ovadia, 2001; and Smith, 2002) مراجعه کنید.

۱-۷-۲ تلویزیون با آنتن مرکزی

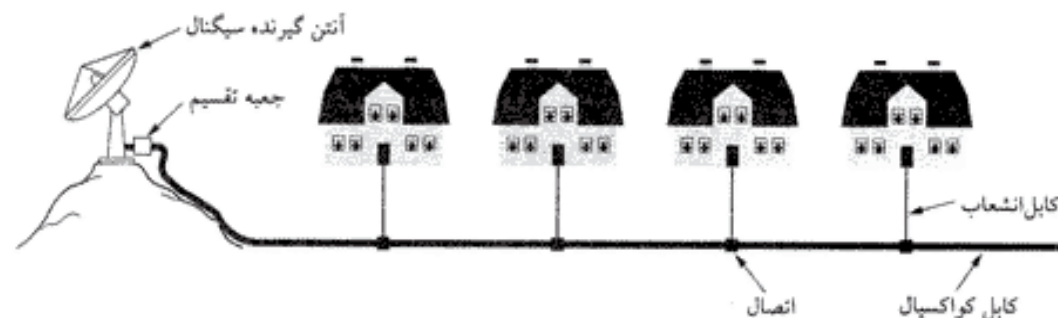
تلویزیون کابلی در اواخر دهه ۱۹۴۰ بعنوان راهی برای ارائه سرویسهای بهتر به مناطق روستایی و کوهستانی ابداع شد. این سیستم عبارت بود از یک آنتن بزرگ بر فراز نقطه‌ای مرتفع (برای دریافت امواج تلویزیونی)، یک تقویت‌کننده (موسوم به **جعبه تقسیم - head end**) برای تقویت سیگنال، و یک رشته کابل کواکسیال برای انتقال سیگنال تلویزیون به منازل؛ شکل ۲-۴۶ را ببینید.

در آن سالها به این سیستم **تلویزیون با آنتن مرکزی (Community Antenna Television)** گفته می‌شد. این سیستم بقدری ساده بود که هر کسی با کمترین تجربه سیمکشی نیز می‌توانست آنرا راه بیندازد، و هزینه آن هم معمولاً بین مشترکین سرشکن می‌شد. با اضافه شدن مصرف‌کنندگان فقط کافی بود انشعابهای بیشتری از کابل اصلی گرفته (و احياناً تقویت‌کننده‌های بیشتری نصب) شود. این سیستم اساساً یکطرفه (از جعبه تقسیم سیگنال به مشترکین) بود، و تا دهه ۱۹۷۰ هزاران نمونه از آن نصب شده بود.

در سال ۱۹۷۴ شرکت رسانه‌ای تایم کانال جدیدی بنام **Home Box Office** راه‌اندازی کرد که فیلمهای سینمایی پخش می‌کرد، و فقط از طریق کابل قابل دسترسی بود. کانالهای کابلی دیگر نیز بسرعت راه افتادند که اخبار، مسابقات ورزشی، آشپزی (و بسیاری موضوعات دیگر) پخش می‌کردند. این تحول باعث دو تغییر اساسی در صنعت تلویزیون شد. اول اینکه، شرکت‌های بزرگ شروع به خریدن شبکه‌های کابلی موجود در شهرها کردند، و در بسیاری جاها نیز خود رأساً شروع به کابل‌کشی و جذب مشترکین جدید کردند. دوم اینکه، برای ارائه سرویس در سراسر کشور لازم بود بین شهرها نیز کابل‌کشی شود، پس شرکت‌های مزبور شروع به متصل کردن شهرها به یکدیگر کردند. این وضعیت درست مانند شبکه تلفن در ۸۰ سال قبل بود، که شرکت‌های تلفن برای ارائه سرویسهای راه دور بین شهرها کابل تلفن کشیدند.

۲-۷-۲ اینترنت کابلی

در طول سالها سیستم‌های کابلی رشد کرد، و بین شهرهای مختلف فیبرهای نوری با پهنای باند زیاد کشیده شد. این

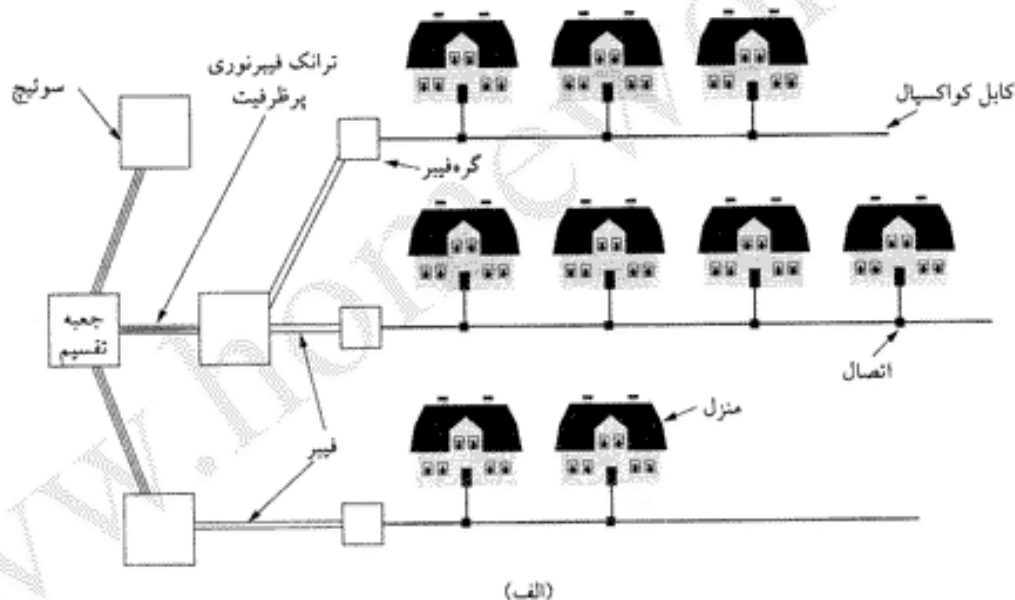


شکل ۲-۴۶. یک سیستم تلویزیون کابلی اولیه.

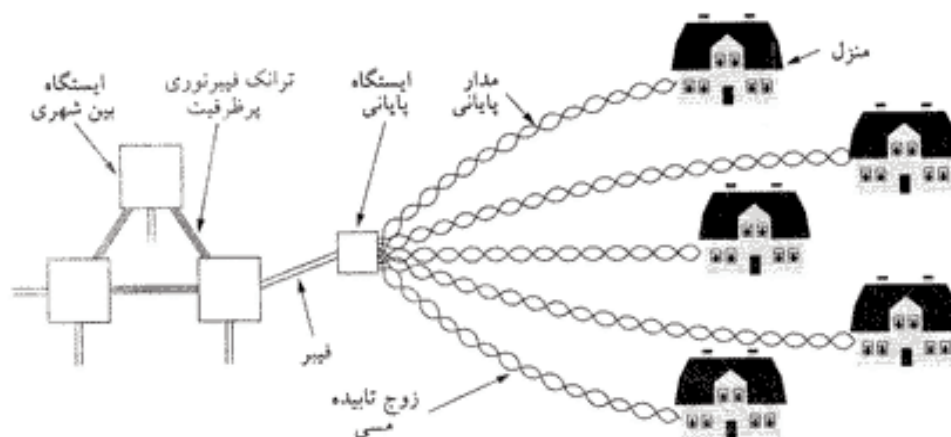
سیستم که ترکیبی بود از فیبر نوری (بین شهرها) و کابلهای کواکسیال (از مراکز توزیع تا منازل) HFC (آمیخته فیبر-کواکس - Hybrid Fiber Coax) نام دارد. در نقاطی که گره فیبر (fiber node) خوانده می شوند، سیگنالهای نوری به الکترونیکی (و بالعکس) تبدیل می شود. از آنجائیکه پهنای باند فیبر نوری بسیار بیشتر از کواکس است، یک گره فیبر می تواند تعداد زیادی کابل کواکسیال را تغذیه کند. در شکل ۲-۴۷ (الف) یک سیستم HFC مدرن را ملاحظه می کنید.

در سالهای اخیر، بسیاری از شرکتهای کابلی تصمیم گرفته اند وارد تجارت دسترسی اینترنت (و همچنین تلفن) شوند. با این حال تفاوت های فنی شبکه های کابلی و تلفن چنان است که تاکنون مانع از تحقق کامل این خواسته شده است. یکی از مشکلاتی که می توان بعنوان نمونه به آن اشاره کرد، تقویت کننده های یکطرفه ای است که باید با تقویت کننده های دوطرفه جایگزین شوند.

اما تفاوت مهمتری بین سیستم HFC شکل ۲-۴۷ (الف) و سیستم تلفن (شکل ۲-۴۷ ب) وجود دارد، که برطرف کردن آن بسیار مشکلتر است. در سیستم HFC یک رشته کابل کواکسیال بین خانه های متعددی مشترک



(الف)



(ب)

شکل ۲-۴۷. (الف) تلویزیون کابلی. (ب) سیستم تلفن ثابت.

است، در حالیکه در سیستم تلفن شهری هر خانه دارای یک اتصال (مدار پایانی) خاص خود است. در پخش برنامه های تلویزیونی این مشترک بودن کابل اهمیت چندانی ندارد: یک برنامه پخش می شود، و اهمیتی ندارد که ۱۰ نفر آنرا می بینند یا ۱۰۰,۰۰۰ نفر. اما وقتی پای دسترسی اینترنت به میان می آید، ۱۰ نفر با ۱۰۰,۰۰۰ نفر فرق بسیاری با هم دارند. اگر یکی از کاربران بخواهد فایل بزرگی را از اینترنت بگیرد، پهنای باند لازم برای این کار از کاربران دیگر گرفته می شود. هر چه تعداد کاربران بیشتر شود، رقابت بر سر پهنای باند شدیدتر خواهد بود. در سیستم تلفن چنین مشکلی وجود ندارد: اگر شما مشغول گرفتن یک فایل بزرگ روی خط ADSL خود باشید، تأثیری روی کار همسایه تان نخواهد گذاشت. از طرف دیگر، پهنای باند یک کابل کوآکس بسیار بیشتر از زوج-تاییده است.

شرکتهای کابلی برای غلبه بر این مشکل، کابلهای بلند را تکه تکه کرده و هر کدام را مستقیماً به یک گره فیبر متصل می کنند. اگر تعداد کاربران هر کابل کوآکس خیلی زیاد نباشد، پهنای باند فیبر نوری عملاً نامحدود خواهد بود و می توان ترافیک را بخوبی مدیریت کرد. امروزه هر کابل کوآکس به ۵۰۰ تا ۲۰۰۰ نفر سرویس می دهد، ولی با زیاد شدن تعداد مشترکین این سیستم شاید لازم باشد از کابلها و گره های فیبر بیشتری استفاده کرد.

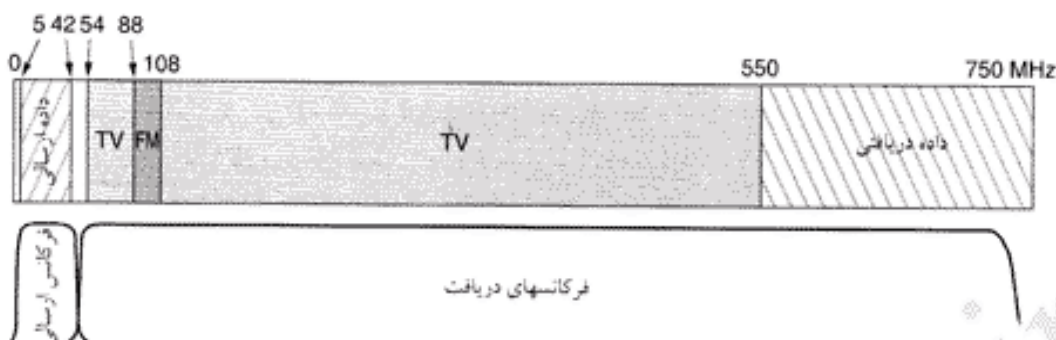
۳-۷-۲ تخصیص طیف فرکانسی

حذف تمام کانالهای تلویزیونی و بکارگیری زیرساخت های کابلی برای دسترسی اینترنت باعث نارضایتی تعداد زیادی از مشترکین خواهد شد، و بهمین دلیل شرکتهای کابلی در این کار تردید دارند. علاوه بر آن، در اغلب شهرها قوانین سختگیرانه ای برای کنترل آنچه روی کابلها فرستاده می شود، وجود دارد، و حتی اگر شرکتهای کابلی واقعاً هم بخواهند اجازه چنین کاری را ندارند. در نتیجه، آنها باید راهی برای همزیستی تلویزیون کابلی و اینترنت روی یک کابل می یافتند.

کانالهای تلویزیون کابلی در منطقه آمریکای شمالی در ناحیه 54-550 MHz طیف قرار دارند (البته با استثنای محدوده 88-108 MHz که به رادیوی FM اختصاص دارد). هر کانال (با احتساب باندهای محافظ) 6 MHz پهنای دارد. در اروپا، حد پائین طیف معمولاً 65 MHz است و کانالها نیز 6-8 MHz پهنای دارند، چون سیستمهای تلویزیونی PAL و SECAM کیفیت بالاتری دارند. از قسمت پائین این باند استفاده نمی شود. کابلهای جدید می توانند بالاتر از 550 MHz (اغلب تا 750 MHz) نیز کار کنند. راه حل انتخاب شده این بود که کانالهای ارسال به اینترنت (upstream) در باند 5-42 MHz (در اروپا، کمی بالاتر) و کانالهای دریافت از اینترنت (downstream) در فرکانسهای بالاتر تعریف شوند. در شکل ۲-۴۸ طیف فرکانسی کابل کوآکسیال را ملاحظه می کنید.

از آنجائیکه سیگنالهای تلویزیونی فقط در یک جهت (مرکز توزیع به منازل) منتشر می شوند، تقویت کننده های ارسال فقط در باند 5-42 MHz، و تقویت کننده های دریافت فقط در باند بالاتر از 54 MHz کار می کنند. همانطور که می بینید، بین باندهای ارسال و دریافت یک عدم تقارن بوجود آمده، و باند دریافت از اینترنت بسیار وسیعتر است (البته این وضعیت چندان هم بد نیست، چون اغلب ترافیک اینترنت در همین جهت است). قبلاً هم دیدید که شرکتهای تلفن عمداً (و بدون اینکه اجبار تکنیکی وجود داشته باشد) سرویس DSL را بصورت نامتقارن درمی آورند.

توانایی کابلهای بلند کوآکسیال در انتقال سیگنالهای دیجیتالی چندان بهتر از سیمهای زوج-تاییده نیست، و به همین دلیل آنها هم به نوعی مدولاسیون آنالوگ احتیاج دارند. برای مدولاسیون هر کانال 6 MHz (یا 8 MHz) از روش QAM-64 (و اگر کیفیت کابل فوق العاده خوب باشد، از QAM-256) استفاده می شود. با یک کانال 6-MHz و مدولاسیون QAM-64، نرخ انتقال داده ای معادل 36 Mbps بدست می آید، که اگر سرآیند از آن کسر شود، تقریباً 27 Mbps باقی می ماند. با مدولاسیون QAM-256 نرخ انتقال داده (بعد از کسر سرآیند) 39 Mbps



شکل ۲-۴. تخصیص فرکانس در یک سیستم تلویزیون کابلی برای دسترسی اینترنت.

خواهد بود. (ظرفیت سیستمهای اروپایی 1/3 بیشتر است.) در کانالهای ارسال به اینترنت بدلیل وجود نویز بالا (از منابع مایکروویو زمینی، و رادیوهای محلی) حتی QAM-64 نیز بخوبی کار نمی کند، و باید از روشهای محافظه کارانه تری (مانند QPSK) استفاده کرد. روش QPSK ظرفیت بسیار کمتری نسبت به مدولاسیون QAM دارد (شکل ۲-۲۵ را ببینید) - 2 bits/ baud بجای 6-8 bits/ baud. این اوصاف، عدم تقارن کانالهای ارسال و دریافت حتی از آنچه در شکل ۲-۴۸ دیده می شود، نیز بیشتر است. شرکت های کابلی مجبور شدند تقویت کننده های ساده خود را نیز با سیستم های دیجیتالی هوشمند جایگزین کنند. نام این تجهیزات نیز از جعبه تقسیم (headend) به CTMS (سیستم پایانه ای مودم کابلی - Cable Modem Termination System) تغییر کرده است، ولی ما در قسمتهای آتی همچنان از اصطلاح «جعبه تقسیم» برای اشاره به این تقویت کننده ها استفاده خواهیم کرد.

۴-۷-۲ مودمهای کابلی

مودم کابلی (cable modem) دستگاهیست با دو سر: یک سر به کامپیوتر وصل می شود، و سر دیگر به شبکه کابلی. در سالهای اولیه اینترنت کابلی، هر شرکت مودم کابلی خاصی داشت، که توسط تکنسینهای آن نصب می شد. اما بزودی معلوم شد که وجود سیستمی با استاندارد باز موجب گسترش رقابت در بازار، کاهش قیمتها، و در نتیجه اقبال عمومی به این سرویسها خواهد شد. علاوه بر آن، خریدن یک مودم از فروشگاه و نصب آن توسط خود مشتری بسیار ساده تر از مراجعه یک تکنسین برای نصب هر مودم است (که هزینه بسیار بالایی نیز دارد). در نتیجه، شرکت های کابلی بزرگ برای تولید یک مودم کابلی استاندارد (و تست سازگاری آنها) با شرکتی بنام CableLabs متحد شدند. این استاندارد که DOCSIS (Data Over Cable Service Interface Specification) نام داشت، فقط نقطه شروعی بود برای جایگزینی مودمهای متنوع و ناسازگار - ویرایش اروپایی این استاندارد نیز EuroDOCSIS نام گرفت. اما ایده استاندارد کردن مودمهای کابلی برای همه شرکت های کابلی جالب نبود، چون آنها پول خوبی بابت اجازه دادن مودم به علاقمندان این سرویسها به جیب می زدند. یک استاندارد باز پای دهها تولیدکننده جدید را به این قلمرو باز می کرد، و به این تجارت پُر منفعت پایان می داد. ارتباط مودم به کامپیوتر ساده و سراسر است: در حال حاضر از اینترنت 10-Mbps (و گاهی هم USB) برای این منظور استفاده می شود. بطور مسلم در آینده شاهد کارتهای مودم کابلی (شبه کارتهای V.9x فعلی) نیز خواهیم بود.

ارتباط در سر دیگر پیچیده تر است، و بخش عمده ای از استاندارد آن به مهندسی رادیو مربوط می شود (که از

حاصله این کتاب خارج است). تنها نکته‌ای که باید بیاد داشته باشید این است که، یک مودم کابلی (مانند مودمهای ADSL) همیشه روشن و متصل است. ارتباط این مودمها به محض روشن شدن برقرار شده و تا زمان خاموش شدن در همین حالت باقی می‌ماند (هزینه آنها نیز بر حسب زمان محاسبه نمی‌شود).

برای درک بهتر طرز کار مودمهای کابلی، اجازه دهید ببینیم با روشن کردن آنها چه اتفاقی می‌افتد. مودم بمحض روشن شدن تمام کانالهای دریافت را برای پیدا کردن بسته خاصی که در فواصل منظم از طرف منبع ارسال می‌شود (و حاوی پارامترهای سیستم برای مودمهایی که تازه روشن شده‌اند، می‌باشد) جستجو می‌کند. پس از پیدا کردن این بسته، مودم جدید حضور خود را از طریق یکی از کانالهای ارسال به منبع اعلام می‌کند. منبع نیز با تخصیص کانالهای ارسال و دریافت پاسخ مودم را می‌دهد. البته این کانالها هر زمان که منبع لازم ببیند (مثلاً، برای متعادل کردن بار)، می‌توانند عوض شوند.

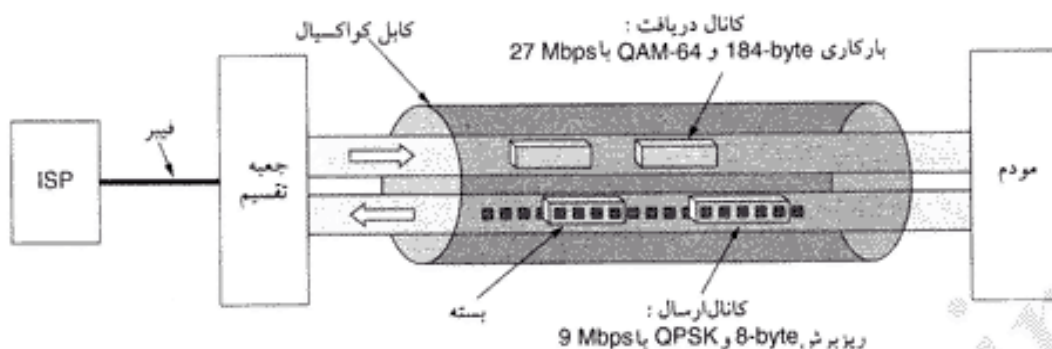
سپس مودم با ارسال یک بسته خاص به منبع و محاسبه زمان رفت و برگشت آن، فاصله خود با منبع را تعیین می‌کند؛ به این کار تعیین فاصله (ranging) گفته می‌شود. این کار برای کارکرد صحیح کانالهای ارسال و همزمانی لازم است. این فاصله زمانی به ریزبُرشها (minislots) تقسیم می‌شود؛ هر بسته باید دقیقاً در یک یا چند ریزبُرش متوالی جا شود. منبع در فواصل منظم دور جدیدی از ریزبُرشها را اعلام می‌کند، ولی (بدلیل یکسان نبودن فاصله مودمها از منبع) این شلیک شروع مسابقه همزمان در تمام مودمها شنیده نخواهد شد. هر مودم، با دانستن فاصله خود از منبع، می‌تواند بفهمد که اولین ریزبُرش واقعاً در چه زمانی شروع شده است. طول ریزبُرشها به شبکه بستگی دارد (و ظرفیت آن معمولاً معادل ۸ بایت است).

در حین آماده‌سازی اولیه، منبع به هر مودم یک ریزبُرش تخصیص می‌دهد که می‌تواند از آن برای درخواست بهنای پاند ارسال استفاده کند. فاعداً تعداد مودمها از ریزبُرشها بیشتر است و یک ریزبُرش به چند مودم اختصاص داده می‌شود، که این می‌تواند موجب کشمکش بین آنها شود. وقتی کامپیوتر اطلاعاتی را برای ارسال به مودم می‌فرستد، مودم تعداد ریزبُرشهای لازم را محاسبه کرده و درخواست تخصیص آنها را از منبع می‌کند. اگر این درخواست پذیرفته شود، منبع ریزبُرشهای تخصیص داده شده را (روی یک از کانالهای دریافت) به مودم اعلام می‌کند. مودم نیز در ریزبُرشهای اختصاص یافته اطلاعات خود را ارسال می‌کند (و اگر به ریزبُرشهای بیشتری نیاز داشت، می‌تواند آنها را از طریق یکی از فیلدهای سرآیند درخواست کند).

اگر جوابی از منبع نرسید (که درخواست همزمان چند مودم می‌تواند یکی از علت‌های آن باشد)، مودم مدتی صبر کرده و دوباره اقدام خواهد کرد. اگر اقدام دوباره هم با شکست مواجه شد، مودم زمان انتظار را بیشتر می‌کند. ایباد دارید که این همان روش تخصیص بُرش زمانی در شبکه ALOHA است؛ از تکنیکهای اینترنت نمی‌توان استفاده کرد، چون مودم راهی برای گوش کردن به رسانه و تشخیص تصادم ندارد. در فصل ۴ این مبحث را مصلتر بررسی خواهیم کرد.

مدیریت کانالهای دریافت از اینترنت بکلی متفاوت است. اول اینکه، در اینجا فقط یک فرستنده وجود دارد (منبع)، و بالطبع کشمکشی هم در میان نیست و منبع می‌تواند از تکنیکهای مالتی پلکس تقسیم زمانی آماری استفاده کند. دیگر اینکه، ترافیک دریافت از اینترنت معمولاً بسیار بیشتر از ترافیک ارسال به اینترنت است، و به همین دلیل از بسته‌هایی با طول ثابت (۲۰۴ بایت) استفاده می‌شود. قسمتی از این بسته سرآیندهاست (کُد تصحیح خطای Reed-Solomon و مانند آن)، و در نهایت ۱۸۴ بایت برای داده‌های کاربر باقی می‌ماند. این عدد برای سازگاری با تلویزیون دیجیتال MPEG-2 انتخاب شده، بنابراین فرمت کانالهای دریافت از اینترنت و تلویزیون یکسان است (شکل ۲-۲۹ را ببینید).

به فرآیند آماده‌سازی مودم برگردیم؛ همین که مودم فاصله خود را با منبع مشخص کرد، و کانالهای ارسال و



شکل ۲-۴۹. کانالهای ارسال و دریافت در منطقه آمریکای شمالی.

دریافت و ریزبُرشها را گرفت، آماده است تا اطلاعات خود را بفرستد. اولین بسته‌ای که مودم می‌فرستد، بسته خاصیت حاوی درخواست یک آدرس IP از ISP با استفاده از پروتکل DHCP (که در فصل ۵ درباره آن توضیح خواهیم داد). همچنین وقت دقیق نیز از منبع پرسیده شد، و منبع به آن جواب می‌دهد.

در قدم بعد ایمنی اطلاعات بایستی تضمین شود. از آنجائیکه کابل مودم بین تعداد زیادی از افراد مشترک است، هر کسی می‌تواند اطلاعات تمام کاربران دیگر را بخواند. برای جلوگیری از سرک کشیدن دیگران، ترافیک در هر دو جهت (ارسال و دریافت) رمز می‌شود. بخشی از فرآیند آماده‌سازی مودم شامل ایجاد کلیدهای رمز است. شاید در نگاه اول این کار غیرممکن بنظر برسد؛ چگونه می‌توان در روز روشن و زیر نگاه هزاران غریبه کلیدهای رمز را رد و بدل کرد؟ این کار با استفاده از الگوریتمی بنام دیفی-هلمان (Diffie-Hellman) ممکن است، اما اجازه دهید توضیحات بیشتر را به فصل ۸ موکول کنیم.

در پایان، مودم هویت منحصر بفرد (شامل نام کاربر و کلمه عبور) خود را روی کانالی که اکنون از امنیت کافی برخوردار است، به ISP اعلام می‌کند. در اینجا فرآیند آماده‌سازی به پایان می‌رسد، و کاربر می‌تواند به کار عادی خود ادامه دهد.

به همین توضیح مختصر درباره مودمهای کابلی بسنده می‌کنیم، ولی می‌توانید اطلاعات تکمیلی را در (Adams and Dulchinos, 2001; Donaldson and Jones, 2001; and Dutta-Roy, 2001) بیابید.

۲-۵-۷ مودم کابلی یا ADSL ؟

کدامیک بهتر است: ADSL یا مودم کابلی؟ این مانند آنست که پرسید کدام سیستم عامل، یا کدام زبان بهتر است جواب به کسی که این سؤال را از او می‌پرسید، بستگی دارد. اما اجازه دهید ADSL و مودم کابلی را از چند نقطه نظر با هم مقایسه کنیم. هر دوی این سیستمها در بخش ستون فقرات از فیبر نوری استفاده می‌کنند، ولی بخش انتهایی آنها متفاوت است. مودم کابلی در بخش انتهایی (مدار پایانی) از کابل کواکسیال استفاده می‌کند، و ADSL از زوج-تابیده. از نظر تئوری، ظرفیت کواکس صدها برابر زوج-تابیده است. اما، تمام این ظرفیت در اختیار کاربر اینترنت نیست، چون پهنای باند کابل صرف چیزهای بدر دنجوری (مثل برنامه‌های تلویزیونی) هم می‌شود.

در عمل، مقایسه ظرفیت این سیستمها بسیار مشکل است. در ADSL، شرکت سرویس دهنده ظرفیت خط را به صراحت مشخص می‌کند (مثلاً، 1 Mbps دریافت از اینترنت و 256 kbps ارسال به اینترنت)، و در اغلب موارد ۸۰٪ این ظرفیت نیز محقق می‌شود. اما شرکت‌های کابلی هرگز درباره ظرفیت سرویس خود صحبت نمی‌کنند، چون این ظرفیت به تعداد کاربرانی که در هر لحظه روی یک کابل هستند، بستگی دارد. گاهی کابل بهتر از ADSL است، و گاهی بدتر. اما چیزی را نمی‌توان از قبل پیش‌بینی کرد (و این از همه آزاردهنده‌تر است). در یک لحظه

کیفیت کابل خوب است، و لحظه بعد (وقتی یکی از آن خوره‌های اینترنت کامپیوتر خود را روشن می‌کند) غیر قابل تحمل.

در ADSL افزایش تعداد کاربران تأثیری روی کیفیت خط کاربران موجود نخواهد گذاشت، چون هر کاربر دارای یک خط مستقل است. در حالیکه در کابل این افزایش موجب افت کیفیت می‌شود. تنها راه مقابله با این مشکل هم انشعاب کابلهای بیشتر از گره فیبر، و کم کردن تعداد کاربران هر خط است. اما این یعنی هزینه بیشتر، چیزی که صاحبان شرکتهای کابلی در برابر آن مقاومت می‌کنند.

سیستمهای تلفن همراه هم وضعیتی شبیه مودم کابلی دارد: گروهی از کاربران بطور مشترک از یک بخش خاص از پهنای باند استفاده می‌کنند. از آنجائیکه ترافیک صدا نسبتاً یکنواخت است، تقسیم پهنای باند بین کاربران (که با تکنیکهای FDM و TDM انجام می‌شود) نیز یکنواخت و ثابت است. اما در ترافیک داده، تقسیم ثابت بهیچوجه کارایی ندارد، چون توزیع زمانی استفاده هر کاربر از کانال اختصاصی (روشی که در مودم کابلی بکار می‌رود) یکنواخت نیست، و موجب هدر رفتن منابع خواهد شد. با این همه، مودم کابلی حتی از این نظر نیز بیشتر شبیه تلفن همراه است تا تلفن ثابت.

سهولت دستیابی یکی دیگر از تفاوتهای ADSL و مودم کابلی است. هر کاربری یک خط تلفن دارد، اما همه آنها آنقدر به ایستگاه تلفن نزدیک نیستند که بتوانند ADSL بگیرند. از طرف دیگر، کابل نیز چیزی نیست که همه داشته باشند، اما اگر از قبل کابل دارید و شرکت طرف قرارداد شما دسترسی اینترنت هم عرضه می‌کند، باید گفت شانس آورده‌اید. در اینجا دیگر فاصله تا گره فیبر یا منبع اهمیتی ندارد. باید خاطر نشان کرد که، از آنجائیکه کابل اساساً رسانه‌ای تلویزیونی است (و از ابتدا برای این منظور نصب شده)، در اغلب شرکتهای و دفاتر اداری وجود ندارد.

خطوط ADSL (بدلیل ماهیت نقطه-به-نقطه آنها) ذاتاً از ایمنی بالاتری نسبت به کابل برخوردارند. هر کاربری که به کابل دسترسی داشته باشد، می‌تواند تمام بسته‌های منتشر شده روی آنرا بخواند. البته همانطور که قبلاً گفتیم، تمام شرکتهای معتبر سرویسهای کابلی ترافیک را در هر دو جهت رمز می‌کنند. اما همین که کسی بتواند اطلاعات رمز شده شما را هم بگیرد، ضریب ایمنی را پائین می‌آورد.

سیستم تلفن عموماً قابل اعتمادتر از کابل است. برای مثال، شرکتهای تلفن دارای سیستمهای برق اضطراری هستند، و سرویسهای آنها حتی در صورت بروز خاموشی کامل هم قطع نمی‌شود. اما در سیستمهای کابلی، اگر برق هر یک از تقویت‌کننده‌های میانی در یکی از زنجیره‌ها قطع شود، ترافیک در یافت آن خط بکلی متوقف خواهد شد.

و بالاخره اینکه، اغلب شرکتهایی که سرویس ADSL ارائه می‌کنند، انتخاب ISP را بر عهده خود شما می‌گذارند (در برخی نقاط حتی اجبار قانونی برای این کار وجود دارد). در مورد شرکتهای کابلی، اغلب خود آنها ارائه‌دهنده سرویس اینترنت هم هستند، و دست شما برای انتخاب ISP باز نیست.

نتیجه نهایی این بحث آنست که، ADSL و کابل بیشتر به هم شبیه‌اند تا متفاوت. آنها سرویسهای مشابهی ارائه می‌کنند، و با افزایش رقابت قیمت آنها هم احتمالاً روز بروز به یکدیگر نزدیکتر خواهد شد.

۸-۲ خلاصه

لایه فیزیکی اساس تمام شبکه‌هاست. طبیعت دو محدودیت بنیادی به تمام کانالهای ارتباطی تحمیل کرده، و همین محدودیت‌هاست که پهنای باند آنها را مشخص می‌کند. این دو محدودیت عبارتند از: حد نایکوئیست (Nyquist limit) که با کانالهای بدون نویز سروکار دارد، و حد شاننون (Shannon limit) که به کانالهای نویزدار مربوط می‌شود.

رسانه انتقال می تواند هدایت پذیر (guided) باشد یا هدایت ناپذیر (unguided). مهمترین رسانه های هدایت پذیر عبارتند از: زوج-تابیده، کابل کواکسیال، و فیبر نوری. رسانه های هدایت ناپذیر نیز عبارتند از: امواج رادیویی، مایکروویو، مادون قرمز، و لیزر. یکی از رسانه های انتقال رو به رشد ماهواره های مخابراتی (بویژه سیستم های مدار پائین - LEO) هستند.

سیستم تلفن یکی از کلیدی ترین اجزای شبکه های گسترده (WAN) است، که مهمترین عناصر آن عبارتند از: مدارهای پایانی (local loop)، ترانکها (trunk)، و سوئیچها (switch). مدارهای پایانی مدارهای آنالوگ زوج-تابیده هستند، که برای انتقال داده های دیجیتال روی آنها باید از مودم (modem) استفاده کرد. ADSL با استفاده از تکنیکهای مدولاسیون و تقسیم مدار پایانی به کانالهای مجازی می تواند به سرعت 50 Mbps برسد. مدارهای پایانی بیسیم (WLL) یکی از تکنولوژیهای جدیدیست که در آینده از آن (بویژه از LMDS) بیشتر خواهید شنید.

ترانک های شبکه تلفن دیجیتالی هستند، و از تکنیکهای مالتی پلکس (شامل FDM، TDM و WDM) در آنها استفاده می شود. تکنیکهای سوئیچینگ نیز بر دو نوع سوئیچینگ مداری (circuit switching) و سوئیچینگ بسته ای (packet switching) است، که هر دو اهمیت زیادی دارند.

برای کاربردهایی که تحرک زیادی دارند، سیستم تلفن ثابت چندان مناسب نیست. تلفنهای همراه امروزه بطور گسترده ای برای ارتباطات صدا مورد استفاده قرار می گیرند، و در آینده نزدیک ترافیک داده نیز در آنها بحد قابل ملاحظه ای خواهد رسید. نسل اول تلفنهای همراه آنالوگ بود، که عمدتاً به سیستمهای AMPS متکی بود. نسل دوم تلفن همراه دیجیتال است، که در آن از سیستمهای GSM، D-AMPS و CDMA استفاده می شود. نسل سوم تلفنهای همراه نیز دیجیتال خواهد بود، و در آن CDMA پهن باند استفاده خواهد شد. یکی از سیستمهایی که می توان از آن برای کاربردهای شبکه نیز بهره گرفت، تلویزیون کابلی (که از تلویزیون با آنتن مرکزی به سیستمهای آمیخته فیبر-کواکس تکامل یافته) است. این سیستم پهنای باند بالقوه زیادی دارد، ولی پهنای باند واقعی آن به تعداد کاربران فعال (و اینکه مشغول چه کاری هستند) بستگی دارد.

مسائل

۱. ضرایب فوریه تابع $f(t) = t$ را محاسبه کنید ($0 \leq t \leq 1$).
۲. هر 1 msec از یک کانال بدون نویز 4-kHz نمونه برداری می شود. حداکثر نرخ داده این کانال چقدر است؟
۳. کانالهای تلویزیونی 6 MHz پهنای دارند. اگر از یک سیگنال دیجیتال چهارسطحی استفاده کنیم، چند bit/sec می توان در این کانال مخابره کرد؟ فرض کنید کانال بدون نویز است.
۴. اگر یک سیگنال باینری در کانالی 3-kHz که نسبت سیگنال به نویز آن 20 dB است، مخابره شود، حداکثر نرخ داده قابل دستیابی چقدر است؟
۵. برای آن که بتوان کاربر T1 را روی یک خط 50-kHz قرار داد، نسبت سیگنال به نویز چقدر باید باشد؟
۶. فرق ستاره غیرفعال و تکرارکننده فعال در یک شبکه فیبر نوری چیست؟
۷. پهنای باند موجود در طیفی به پهنای $0.1 \mu\text{m}$ در طول موج $1 \mu\text{m}$ چقدر است؟
۸. می خواهیم یکسری تصاویر کامپیوتری اسکن شده را روی یک رشته فیبر نوری بفرستیم. وضوح هر تصویر 480×640 پیکسل، و هر پیکسل ۲۴ بیت است. تصاویر با سرعت ۶۰ صفحه بر ثانیه اسکن می شوند. پهنای باند مورد نیاز چقدر است؟ اگر از باند $1.30 \mu\text{m}$ استفاده کنیم، به چند میکرون از طول موج نیاز داریم؟
۹. آیا قضیه نایکونیست برای فیبرهای نوری هم صادق است، یا فقط برای کابلهای مسی کاربرد دارد؟
۱۰. در شکل ۲-۶، باند سمت چپ از بقیه باندها باریکتر است. چرا؟

۱۱. یک آنتن زمانی بهترین بهره را دارد که قطر آن معادل طول موج امواج رادیویی باشد. قطر قابل قبول آنتنها بین ۱ تا ۵ متر است. این قطر معادل کدام طیف فرکانسی است؟
۱۲. محوشدگی چندمسیره زمانی به حداکثر می‌رسد که دو موج با اختلاف فاز 180° درجه وارد گیرنده شوند. برای به حداکثر رسیدن محوشدگی چندمسیره در یک لینک مایکروویو 1-GHz بطول 50-km این مقدار چقدر باید باشد؟
۱۳. پرتو لیزری بقطر 1 mm روی آشکارسازی بقطر 1 mm که روی پشت‌بامی در فاصله 100 m قرار دارد، نشانه گرفته شده است. حداکثر انحراف زاویه‌ای (بر حسب درجه) چقدر باید باشد، تا پرتو لیزری هدف را گم نکند؟
۱۴. ماهواره پروژۀ ایریدیوم به شش کمربند بدور زمین تقسیم شده‌اند. در ارتفاعی که این ماهواره‌ها قرار دارند، دورۀ گردش مداری ۹۰ دقیقه است. متوسط زمان پاس‌کاری یک فرستندۀ زمینی بین دو ماهواره چقدر است؟
۱۵. ماهواره‌ای در مدار زمین ثابت با صفحه استوای زمین زاویه ϕ می‌سازد. آیا برای فردی که روی زمین در مدار ϕ درجه شمالی ایستاده، این ماهواره در آسمان ثابت بنظر می‌رسد؟ اگر نه، حرکت آنرا توضیح دهید.
۱۶. قبل از سال ۱۹۸۴ (وقتی هر ایستگاه پایانی با کد سه رقمی ناحیه و سه رقم اول شماره تلفن مشخص می‌شد) در سیستم تلفن چند کد ایستگاه پایانی می‌توانست وجود داشته باشد؟ کد ناحیه با عددی بین ۲ تا ۹ شروع می‌شد، رقم دوم می‌توانست ۰ یا ۱ باشد، و رقم سوم محدودیتی نداشت. دو رقم اول کد محلی نیز بایستی بین ۲ تا ۹، و رقم سوم می‌توانست هر عددی باشد.
۱۷. فقط با اطلاعاتی که در اینجا بدست آوردید، آیا می‌توانید بگویند حداکثر شماره تلفنهایی که (بدون تغییر در روش شماره‌گذاری یا اضافه کردن تجهیزات) می‌توان در ایالات متحده نصب کرد، چه تعداد است؟ آیا این تعداد شماره قابل دستیابی است؟ هر دستگاه فکس یا کامپیوتر را نیز یک تلفن در نظر بگیرید، و فرض کنید هر مشترک فقط یک دستگاه تلفن دارد.
۱۸. یک سیستم ساده تلفن را که در آن دو ایستگاه پایانی و یک ایستگاه بین‌شهری بوسیله خطوط دو-طرفه همزمان 1-MHz به هم متصل شده‌اند، در نظر بگیرید. هر دستگاه تلفن بطور متوسط در هر روز کاری ۸ ساعته ۴ تماس برقرار می‌کند، و زمان متوسط هر تماس ۶ دقیقه است. ده درصد تماسها راه دور هستند (یعنی از ایستگاه بین‌شهری رد می‌شوند). حداکثر تعداد شماره‌هایی که هر ایستگاه پایانی می‌تواند پشتیبانی کند، چقدر است؟ مدارها را 4-kHz در نظر بگیرید.
۱۹. یک شرکت تلفن منطقه‌ای ۱۰ میلیون مشترک دارد، که بوسیله زوج-ناپایده به ایستگاه مرکزی متصل شده‌اند. متوسط طول مدارهای پایانی ۱۰ کیلومتر است. ارزش مس موجود در مدارهای پایانی این سیستم چقدر است؟ سطح مقطع سیمها را دایره‌ای بقطر 1 mm، چگالی مس را 9.0 gr/cm^3 و ارزش هر کیلوگرم مس را 3 دلار در نظر بگیرید.
۲۰. یک خط لوله نفت سیستمی یکطرفه است، یا دو-طرفه ناهمزمان، یا دو-طرفه همزمان، یا هیچکدام؟
۲۱. قیمت میکروپروسسورهای سریع آنقدر کاهش یافته، که می‌توان در هر دستگاه مودم یک میکروپروسسور قرار داد. این کار چه تأثیری روی مقابله با خطاهای خطوط تلفن دارد؟
۲۲. دیاگرام فلکی شکل ۲-۲۵ چهار نقطه داده در مختصات $(1, 1)$ ، $(1, -1)$ ، $(-1, 1)$ و $(-1, -1)$ دارد. مودمی با این پارامترها در 1200 baud به چه سرعتی (bps) می‌تواند دست یابد؟
۲۳. مودمی با دیاگرام فلکی شبیه شکل ۲-۲۵ دارای نقاط داده‌ای در مختصات $(0, 1)$ و $(0, 2)$ است. این مودم

- از مدولاسیون فاز استفاده می‌کند، یا مدولاسیون دامنه؟
۲۴. در یک دیاگرام فلکی تمام نقاط روی دایره‌ای به مرکز مبدأ مختصات واقع شده‌اند. مدولاسیون این مودم چیست؟
۲۵. یک مودم دو-طرفه همزمان QAM-64 از چند فرکانس استفاده می‌کند؟
۲۶. در یک سیستم ADSL که از DMT استفاده می‌کند، 3/4 کانالهای موجود به لینک دریافت اختصاص داده شده است. هر کانال از مدولاسیون QAM-64 استفاده می‌کند. ظرفیت لینک دریافت چقدر است؟
۲۷. در سیستم LMDS چهار قطاعی شکل ۲-۳۰، هر قطاع یک کانال اختصاصی 36-Mbps دارد. طبق تئوری صف، اگر 50% کانال پر باشد، زمان انتظار در صف معادل زمان بار شدن است. در چنین شرایطی، بار شدن یک صفحه وب 5-KB چقدر طول خواهد کشید؟ بار شدن این صفحه روی یک خط ADSL با سرعت 1 Mbps چقدر طول می‌کشد؟ با یک مودم 56-kbps چقدر؟
۲۸. ده سیگنال، که هر کدام به پهنای باند 4000 Hz نیاز دارند، با استفاده از FDM روی یک کانال مالتی پلکس شده‌اند. حداقل پهنای باند مورد نیاز این کانال چقدر است؟ پهنای باندهای محافظ را 400 Hz در نظر بگیرید.
۲۹. چرا زمان نمونه برداری PCM در $125 \mu\text{sec}$ ثابت شده است؟
۳۰. درصد سرآیند یک کاربر T1 (درصدی از 1.544 Mbps که بکار داده‌های کاربر نمی‌آید) چقدر است؟
۳۱. حداکثر نرخ داده یک کانال بدون نویز 4-kHz را با استفاده از تکنیکهای زیر مقایسه کنید:
(الف) کدگذاری آنالوگ (مثلاً، QPSK) با 2 bits/sample
(ب) سیستم T1 PCM
۳۲. اگر یک سیستم T1 دچار لغزش شود، برای سنکرون شدن مجدد از اولین بیت هر فریم استفاده می‌کند. برای سنکرون شدن مجدد با احتمال خطای 0.001، چند فریم باید بررسی شود؟
۳۳. فرق بخش دمودلانور یک مودم با بخش دکودر یک کدک چیست (و آیا اساساً فرقی دارند)؟ توجه داشته باشید که هر دوی اینها سیگنالهای آنالوگ را به دیجیتال تبدیل می‌کنند.
۳۴. سیگنالی بصورت دیجیتال روی یک کانال بدون نویز 4-kHz (با یک نمونه در هر $125 \mu\text{sec}$) فرستاده می‌شود. با هر یک از روشهای کدگذاری زیر چند بیت در ثانیه ارسال می‌شود:
(الف) استاندارد CCITT 2.048 Mbps
(ب) سیستم DPCM با مقدار نسبی سیگنال 4-bit
(ج) مدولاسیون دلتا
۳۵. یک سیگنال سینوسی کامل با دامنه A با استفاده از مدولاسیون دلتا (با x samples/sec) کد شده است. خروجی $+1$ معادل $+A/8$ تغییر در سیگنال ورودی، و خروجی -1 معادل $-A/8$ تغییر در سیگنال ورودی است. بیشترین فرکانسی که این سیستم می‌تواند بدون خطای تجمعی تعقیب کند، چقدر است؟
۳۶. ساعت‌های SONET خطایی معادل 1 در 10^9 دارند. چه مدت طول می‌کشد، تا این اختلاف باندازه 1 بیت شود؟ عوارض جانبی این پدیده چیست؟
۳۷. در شکل ۲-۳۷، نرخ داده کاربر OC-3 از 148.608 Mbps شروع شده است. نشان دهید این عدد چگونه از پارامترهای SONET OC-3 بدست آمده است.
۳۸. سیستم SONET برای انطباق با سرعت‌های پائین تر از STS-1 از روشی بنام انشعابات مجازی (Virtual Tributaries - VT) استفاده می‌کند. یک VT عبارتست از یک بار جزئی، که می‌توان آنرا به همراه بارهای

- جزئی دیگر در یک فریم STS-1 قرار داد. VT1.5 از ۳ ستون، VT2 از ۴ ستون، VT3 از ۶ ستون و VT6 از ۱۲ ستون فریم STS-1 استفاده می‌کنند. کدام VT با هر یک از سرویسهای زیر منطبق است؟
- (الف) سرویس DS-1 (1.544 Mbps)
- (ب) سرویس European CEPT-1 (2.048 Mbps)
- (ج) سرویس DS-2 (6.312 Mbps)
۴۹. تفاوت بنیادی سونیچینگ مداری با سونیچینگ بسته‌ای چیست؟
۴۰. بهنای باند کاربر یک اتصال OC-12c چقدر است؟
۴۱. سه شبکه سونیچینگ بسته‌ای هر کدام ۳ گره دارند. شبکه اول دارای توپولوژی ستاره (با سونیچ مرکزی) است، شبکه دوم حلقه (دوطرفه) است، و شبکه سوم اتصالات کامل داخلی دارد (یعنی هر گره مستقیماً به تمام گره‌های دیگر متصل است). بهترین، بدترین و متوسط پرش (hop) در ارتباط از یک نقطه به نقطه دیگر در هر یک از این شبکه‌ها چیست؟
۴۲. زمان تأخیر ارسال یک پیام x -bit در مسیری با k پرش در یک شبکه سونیچینگ مداری و یک شبکه سونیچینگ بسته‌ای (با بار کم) را با یکدیگر مقایسه کنید. زمان برقراری مدار را s ثانیه، زمان تأخیر در هر پرش را d ثانیه، اندازه هر بسته را p بیت، و نرخ انتقال داده را b pbs در نظر بگیرید. در چه شرایطی تأخیر ارسال شبکه سونیچینگ بسته‌ای کمتر است؟
۴۳. فرض کنید می‌خواهیم x بیت اطلاعات را در یک شبکه سونیچینگ بسته‌ای با k پرش، بصورت بسته‌هایی با p بیت داده و h بیت سرآیند (با این فرض که $x \gg p + h$) منتقل کنیم. نرخ انتقال داده خطوط b pbs و b و زمان تأخیر انتشار در آنها قابل صرفنظر کردن است. چه مقداری از p تأخیر کلی را به حداقل می‌رساند؟
۴۴. در یک سیستم تلفن همراه با سلولهای شش ضلعی، استفاده از باندهای فرکانسی مشابه در سلولهای مجاور ممنوع است. اگر ۸۴۰ باند فرکانسی داشته باشیم، در هر سلول از چند فرکانس می‌توان استفاده کرد؟
۴۵. طرح کلی سلولهای یک شبکه تلفن همراه بندرت مانند شکل ۲-۴۱ منظم است؛ حتی شکل هر سلول نیز منظم نیست. یک علت برای این وضعیت بیاورید.
۴۶. برای پوشش دادن به شهری با مساحت 120 km^2 ، به چه تعداد سلول PCS با قطر 100 m نیاز داریم؟ (تخمین بزنید.)
۴۷. گاهی هنگام عبور یک کاربر تلفن همراه از سلولی به سلول دیگر، با وجود اینکه تمام فرستنده‌ها و گیرنده‌ها بخوبی کار می‌کنند، ارتباط ناگهان قطع می‌شود. چرا؟
۴۸. کیفیت صدای D-AMPS بسیار پایتتر از GSM است. آیا این بخاطر اجبار D-AMPS در حفظ سازگاری با AMPS است (در حالیکه GSM چنین محدودیتی ندارد)؟ یا علت دیگری دارد؟
۴۹. حداکثر تعداد کاربران همزمان در یک سلول D-AMPS را محاسبه کنید. آیا چنین محاسبه‌ای برای GSM هم امکان دارد؟ علت را توضیح دهید.
۵۰. فرض کنید سه ایستگاه A ، B و C در یک سیستم CDMA (با توالیهای چیب شکل ۲-۴۵ ب) همزمان اقدام به ارسال بیت‌های 0 می‌کنند. توالی چیب حاصله چیست؟
۵۱. در بحث متعامد بودن بردارهای توالی چیب CDMA، گفتیم که اگر $S \cdot T = 0$ ، آنگاه $S \cdot \bar{T} = 0$ ثابت کنید.
۵۲. اجازه دهید متعامد بودن بردارهای توالی چیب CDMA را به روشی دیگر بیان کنیم: هر بیت در یک جفت توالی یا یکسان هستند، یا نیستند. متعامد بودن بردارها را با استفاده از اصطلاحات یکسان بودن و یکسان بودن توضیح دهید.

۵۳. یک گیرنده CDMA توالی $(+1 +1 -3 +1 -1 -3 +1 +1)$ را دریافت می‌کند. با فرض توابع چیب شکل ۲-۴۵ (ب)، تعیین کنید کدام ایستگاه‌ها، چه بیت‌هایی را ارسال کرده‌اند؟
۵۴. شبکه تلفن در بخش انتهایی دارای توپولوژی ستاره است، که در آن تمام انشعابات به ایستگاه پایانی ختم می‌شوند. در حالیکه در تلویزیون کابلی، یک کابل مشترک مانند ماری بین مشترکین مختلف خزیده است. فرض کنید در آینده در شبکه‌های کابلی بجای کابل‌های مسی از فیبر نوری 10 Gbps استفاده شود. آیا با چنین سیستمی می‌توان مدل شبکه تلفن (یک خط مستقل از هر مشترک به ایستگاه مرکزی) را شبیه‌سازی کرد؟ اگر پاسخ مثبت است، هر فیبر چند کاربر تلفن می‌تواند داشته باشد؟
۵۵. سیستمهای تلویزیون کابلی معمولاً دارای ۱۰۰ کانال تجاری هستند، که بطور متناوب برنامه و آگهی پخش می‌کنند. این سیستم بیشتر شبیه TDM است یا FDM؟
۵۶. یک شرکت کابلی تصمیم می‌گیرد به ۵۰۰۰ مشترک خود سرویس اینترنت ارائه دهد. این شرکت از کابل‌های کواکسیال استفاده می‌کند، که هر کابل می‌تواند تا 100 Mbps روی کانال دریافت از اینترنت ظرفیت داشته باشد. شرکت مزبور برای جذب مشتریان تصمیم می‌گیرد که تا دریافت 2 Mbps را برای هر مشترک تضمین کند. توضیح دهید این شرکت برای رسیدن به هدف فوق چه کاری باید انجام دهد؟
۵۷. با توجه به تخصیص فرکانس نشان داده شده در شکل ۲-۴۸ و اطلاعات داده شده در متن کتاب، یک سیستم کابلی چه مقدار (Mbps) از ظرفیت را به ارسال به اینترنت و چه مقدار را به دریافت از اینترنت اختصاص می‌دهد؟
۵۸. اگر سیستم کابلی کاملاً بیکار باشد، کاربر با چه سرعتی می‌تواند اطلاعات را دریافت کند؟
۵۹. مالتی پلکس کردن استریمهای متعدد STS-1 (که به آنها انشعاب گفته می‌شود) نقش مهمی در سیستم SONET بازی می‌کند. یک مالتی پلکسر 3:1 سه ورودی STS-1 را در یک خروجی STS-3 مالتی پلکس می‌کند. اینکار بصورت بایت به بایت انجام می‌شود، یعنی سه بایت اول خروجی بترتیب بایتهای اول انشعابهای 1، 2 و 3 هستند؛ سه بایت دوم خروجی بترتیب بایتهای دوم انشعابهای 1، 2 و 3؛ و الی آخر. برنامه‌ای بنویسید که این مالتی پلکسر 3:1 را شبیه‌سازی کند. برنامه شما باید پنج روال داشته باشد: روال اصلی (که چهار روال دیگر را اجرا می‌کند)، یک روال برای هر یک از انشعابهای STS-1 (مجموعاً سه روال)، و یکی برای مالتی پلکسر. هر روال انشعاب یک فریم STS-1 را از فایلی بطول ۸۱۰ بایت خوانده، و این فریمها را (بایت به بایت) به روال مالتی پلکسر می‌فرستد. روال مالتی پلکسر این بایتهای را خوانده، و یک فریم STS-3 را (بایت به بایت) روی خروجی استاندارد (stdout) می‌نویسد. برای ارتباط بین پردازشها از پایپ (pipe) استفاده کنید.

لایه پیوند داده



در این فصل اصول طراحی لایه دوم، لایه پیوند داده (data link layer)، را بررسی خواهیم کرد، و طی آن با الگوریتمهای لازم برای دستیابی به یک ارتباط قابل اطمینان و کارا بین دو کامپیوتر همسایه (در لایه پیوند داده) آشنا خواهیم شد. منظور از دو کامپیوتر همسایه، کامپیوترهایی هستند که یک کانال ارتباطی سیم-مانند (کابل کواکسیال، خط تلفن، و یا ارتباط بیسیم) بین آنها برقرار است. خصلت بنیادی یک کانال «سیم-مانند» اینست که بیت‌ها دقیقاً با همان نظمی که فرستاده می‌شوند، در گیرنده دریافت شوند.

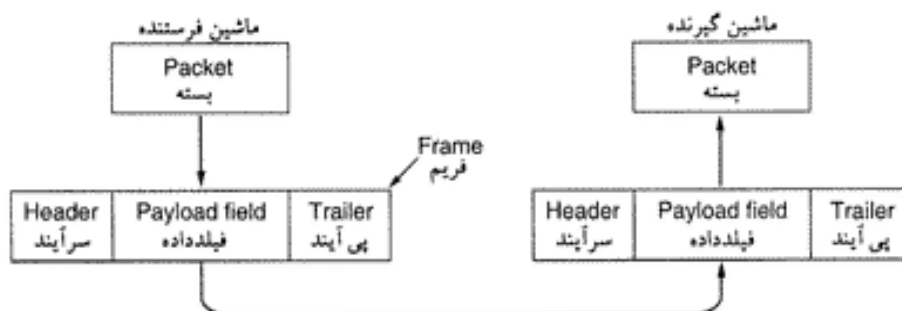
شاید در نگاه اول این خصلت آنقدر ساده و ابتدایی بنظر برسد، که فکر کنید چه نیازی به الگوریتم و نرم‌افزار هست: ماشین A بیت‌ها را می‌فرستد، و ماشین B آنها را می‌گیرد. متأسفانه، مسئله بهین سادگی نیست، چون مدارهای مخابراتی پر از نویز و خطا هستند. علاوه بر آن ظرفیت کانالهای مخابراتی نامحدود نیست، و بین ارسال و دریافت بیت‌ها یک تأخیر زمانی نیز وجود دارد. این محدودیت‌ها تأثیر جدی روی کارایی سیستمهای انتقال داده می‌گذارند. پروتکل‌های مخابراتی (که موضوع اصلی این فصل هستند) باید تمام این ملاحظات را در نظر بگیرند. بعد از آشنایی با نکات کلیدی در طراحی لایه پیوند داده، پروتکل‌های این لایه را بررسی خواهیم کرد. برای شروع خصلت خطاهای کانالهای مخابراتی، منشأ آنها و نحوه کشف و رفع این خطاها را بررسی کرده، و سپس پروتکل‌های لازم برای حل آنها را مورد مطالعه قرار می‌دهیم. در پایان، صحت این مدلها را بررسی کرده، و چند نمونه از پروتکل‌های واقعی لایه پیوند داده ارائه خواهیم کرد.

۱-۳ ملاحظات طراحی لایه پیوند داده

لایه پیوند داده وظایف خاصی دارد که باید انجام دهد. این وظایف عبارتند از:

۱. ارائه سرویسهای مشخص به لایه شبکه.
۲. مدیریت خطاهای انتقال.
۳. تنظیم جریان داده‌ها (بگونه‌ایکه گیرنده‌های گنذ زیر بمباران فرستنده‌های سریع غرق نشوند).

برای رسیدن به این اهداف، لایه پیوند داده بسته‌های رسیده از لایه شبکه را گرفته و آنها بصورت فریم (frame) در می‌آورد. هر فریم سه قسمت دارد: سرآیند (header)، داده اصلی، و پی‌آیند (trailer)؛ شکل ۱-۳ را ببینید. مدیریت فریمها کلیدی‌ترین وظیفه لایه پیوند داده است، که در بخشهای آینده بتفصیل درباره آن صحبت خواهیم کرد

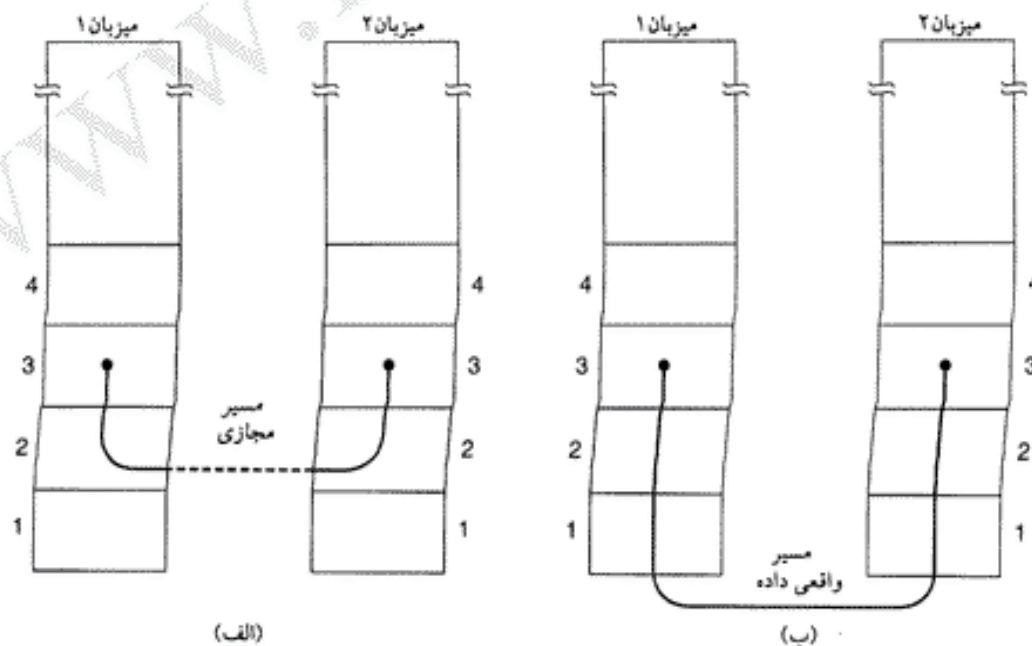


شکل ۳-۱. رابطه بسته و فریم.

با اینکه این فصل منحصرأ درباره لایه پیوند داده و پروتکل های آن است، اصولی که در اینجا خواهید دید (مانند کنترل خطا و کنترل جریان داده) در لایه های دیگر (مانند لایه انتقال) نیز کاربرد دارند. در حقیقت، در بسیاری از شبکه ها این کارکردها را فقط در لایه های بالاتر (نه در لایه پیوند داده) می توانید پیدا کنید. اما صرف نظر از اینکه آنها را کجا می توان پیدا کرد، اصول کار یکسان است و اهمیتی ندارد که در این فصل آنها را بررسی کنیم یا فصلهای دیگر. تنها مزیت لایه پیوند داده آنست که این تکنیکها در این لایه ساده تر و واضح ترند، بنابراین بخوبی می توان آنها را مطالعه کرد.

۱-۱-۳ سرویسهایی که به لایه شبکه داده می شود

وظیفه لایه پیوند داده ارائه سرویس به لایه شبکه است. مهمترین این وظایف عبارتست از انتقال داده ها از لایه شبکه ماشین مبدأ به لایه شبکه ماشین مقصد. در لایه شبکه ماشین مبدأ چیزی هست بنام پرومیس، که تعدادی بیت را به لایه پیوند داده می دهد تا به مقصدی خاص منتقل کند. وظیفه لایه پیوند داده ماشین مبدأ انتقال این بیت ها به ماشین مقصد، و رساندن آنها بدست لایه شبکه مقابل است؛ شکل ۳-۲ (الف) را ببینید. البته مسیری که این بیت ها



شکل ۳-۲. (الف) ارتباط مجازی. (ب) ارتباط واقعی.

واقعاً طی می‌کنند، مانند شکل ۳-۲ (ب) است، ولی ساده‌ترست تصور کنیم دو پروتکل در لایه پیوند داده آنها را بین خود رد و بدل می‌کنند. بهمین دلیل در این فصل همه جا از مدل شکل ۳-۲ (الف) استفاده خواهیم کرد. لایه پیوند داده را می‌توان بگونه‌ای طراحی کرد که سرویسهای مختلفی ارائه کند، که این سرویسها از سیستمی به سیستم دیگر متفاوت است. معقولترین این سرویسها عبارتند از:

۱. سرویس غیرمتصل بدون تصدیق دریافت (unacknowledged connectionless).
۲. سرویس غیرمتصل با تصدیق دریافت (acknowledged connectionless).
۳. سرویس اتصال-گرا با تصدیق دریافت (acknowledged connection-oriented).

اجازه دهید این سرویسها را یکی یکی بررسی کنیم.

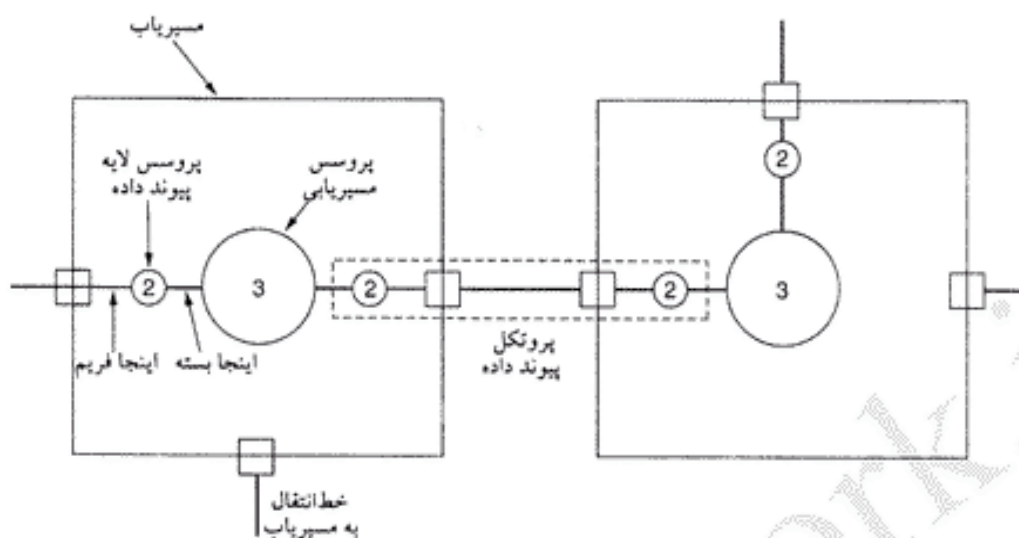
در سرویس غیرمتصل بدون تصدیق دریافت ماشین مبدأ فریمهای مستقلی را به ماشین مقصد می‌فرستد، بدون اینکه منتظر تصدیق دریافت آنها از طرف ماشین مقصد بماند. هیچ اتصال منطقی بین دو ماشین برقرار نمی‌شود، پس نیازی به قطع اتصال هم نیست. اگر فریمی در اثر نویز خط از بین برود، هیچ کوششی برای تشخیص این موضوع و مقابله با آن در لایه پیوند داده صورت نمی‌گیرد. این سرویس برای مواقعی مناسب است که نرخ خطا بسیار پائین باشد، و در این حالت مقابله با خطا به لایه‌های بالاتر واگذار می‌شود. این سرویس برای ترافیک زمان-واقعی (مانند سرویس صدا)، که در آن دیر رسیدن بدتر از نرسیدن است، نیز مناسب است. در اغلب LAN ها نیز لایه پیوند داده از سرویسهای غیرمتصل بدون تصدیق دریافت استفاده می‌کند.

سرویس بعدی که قابلیت اعتماد بیشتری دارد، سرویس غیرمتصل با تصدیق دریافت است. در این سرویس نیز هیچ اتصال منطقی بین مبدأ و مقصد وجود ندارد، ولی دریافت فریمها از سوی ماشین مقصد تصدیق می‌شود. بدین ترتیب، فرستنده می‌تواند پی ببرد که آیا فریمها بدرستی دریافت شده‌اند یا خیر. اگر فریمی در مدت زمان معین به مقصد نرسد، می‌توان آنرا دوباره ارسال کرد. این سرویس برای کانالهای غیر قابل اعتماد (مانند سیستمهای بیسیم) مناسب است.

لازمست تأکید کنیم که توجه به تصدیق دریافت در لایه پیوند داده فقط برای بهینه‌سازی سیستم است و هیچ الزامی در آن نیست، چون این کار را همیشه می‌توان در لایه شبکه انجام داد. اگر تصدیق دریافت در زمان مشخص از راه نرسد، فرستنده می‌تواند بسته را دوباره ارسال کند. مشکل اینجاست که فریمها معمولاً طول مشخصی دارند، در حالیکه بسته‌های که لایه شبکه می‌فرستد چنین نیست. اگر بسته‌ای به، مثلاً، ۱۰ فریم شکسته شود، و ۲۰ درصد این فریمها در راه گم شوند، زمان ارسال بسته بسیار طولانی خواهد شد. اما اگر برای هر فریم تصدیق دریافت درخواست شود، این کار سریعتر می‌شود. در کانالهای قابل اعتماد مانند فیبر نوری، بار اضافی چنین پروتکل‌های سختگیرانه‌ای در لایه پیوند داده غیر ضروری است، اما در محیطهای ذاتاً پرنویز مانند بیسیم ارزشش را دارد.

بهترین سرویسی که لایه پیوند داده می‌تواند به لایه شبکه بدهد، سرویس اتصال-گرا (connection-oriented) است. در این سرویس قبل از شروع ارسال داده از مبدأ به مقصد، یک اتصال بین آنها برقرار می‌شود. هر فریمی که روی این اتصال فرستاده می‌شود شماره‌گذاری شده است، و لایه پیوند داده دریافت آنها را نیز تضمین می‌کند. همچنین تضمین می‌شود که هر فریم فقط یک بار (و به همان ترتیب ارسال) دریافت شود. اما در سرویسهای غیرمتصل، می‌توان انتظار داشت که بسته‌ای چندین بار ارسال (و در نتیجه چندین بار هم دریافت) شود. سرویس اتصال-گرا استریم قابل اعتمادی از بیت‌ها را در اختیار لایه شبکه می‌گذارد.

ارسال داده‌ها در سرویس اتصال-گرا سه مرحله دارد. در مرحله اول اتصال برقرار شده، و متغیرهای لازم (برای شمارش فریمها، و اینکه کدام فریمها دریافت شده‌اند و کدامها خیر) ست می‌شوند. در مرحله دوم، فریمها منتقل می‌شوند. و در مرحله آخر، اتصال قطع شده و منابع آن (متغیرها و بافرها) آزاد می‌شود.



شکل ۳-۳. محل فعالیت لایه پیوند داده.

اجازه دهید یک مثال بزنیم: یک زیرشبکه WAN متشکل از چند مسیریاب که با خطوط نقطه-به-نقطه تلغز به یکدیگر متصل شده‌اند، را در نظر بگیرید. وقتی یک فریم به مسیریاب می‌رسد، سخت‌افزار (با استفاده از تکنیکهایی که در همین فصل خواهید دید) آنرا از نظر خطا چک می‌کند، و سپس به نرم‌افزار لایه پیوند داده (که می‌تواند روی چیپهای کارت شبکه قرار داشته باشد) تحویل می‌دهد. نرم‌افزار لایه پیوند داده فریم را چک می‌کند تا مطمئن شود همان چیزیست که باید باشد، و اگر چنین بود، قسمت داده اصلی آنرا به نرم‌افزار مسیریابی می‌دهد. نرم‌افزار مسیریابی مسیر خروجی مناسب را تعیین کرده، و بسته را به لایه پیوند داده پس می‌دهد تا ارسال شود (شکل ۳-۳ را ببینید).

نرم‌افزارهای مسیریابی معمولاً حوصله بسته‌هایی که مدام گم می‌شوند را ندارند، و دوست دارند بسته‌ها درست و مرتب روی خطوط نقطه-به-نقطه تحویل شوند. این دیگر بر عهده پروتکل لایه پیوند داده است که خطوط پرنویز و غیر قابل اعتماد را بصورتی مطمئن (یا نسبتاً مطمئن) در آورد. با اینکه در شکل ۳-۳ نرم‌افزار لایه پیوند داده (در هر مسیریاب) در دو نقطه دیده می‌شود، اما این در واقع یک پروسس واحد است که (به کمک جدول‌ها و ساختمان داده‌های مختلف) تمام کارها را انجام می‌دهد و تمام خطوط را کنترل می‌کند.

۲-۱-۳ فریم‌بندی

لایه پیوند داده به لایه شبکه سرویس می‌دهد، و خود نیز از سرویسهای لایه فیزیکی استفاده می‌کند. چیزی که لایه فیزیکی می‌گیرد، استریمی است از بیت‌ها که باید آنرا به طرف مقابل تحویل دهد. هیچ تضمینی وجود ندارد که این استریم سالم و عاری از خطا به مقصد برسد. تعداد بیت‌های رسیده می‌تواند کمتر، مساوی یا بیشتر از بیت‌های ارسال شده باشد، و یا حتی مقدار برخی از آنها تغییر کرده باشد. این بر عهده لایه پیوند داده است که خطاها را کشف کرده، و در صورت لزوم آنها را برطرف کند.

یکی از روشهای متداول اینست که استریم بیت‌ها در لایه پیوند داده به چند فریم شکسته شده، و برای هر فریم جمع تطبیقی (checksum) محاسبه شود. (الگوریتمهای جمع تطبیقی را در همین فصل خواهید دید.) وقتی فریمها به مقصد می‌رسند، جمع تطبیقی آنها مجدداً محاسبه شده و با جمع تطبیقی مبدأ (که به انتهای فریم ضمیمه شده) مقایسه می‌شود. اگر این دو یکی نباشند، لایه پیوند داده متوجه می‌شود که خطایی در فریم رخ داده، و به

سراغ روشهای مقابله با خطا می رود (که یکی از این روشها می تواند دور انداختن فریم، و درخواست ارسال مجدد آن باشد).

شکستن استریم بیت ها به فریم (که به آن فریم بندی - framing - گفته می شود) از آنچه در نگاه اول بنظر می رسد، مشکلتر است. یکی از روشهای فریم بندی می تواند انداختن فاصله زمانی در نقاطی از استریم بیت ها باشد (مانند فاصله انداختن بین کلمات متن). ولی در شبکه ها پندرت زمانبندی وجود دارد، و امکان دارد این فاصله ها از بین بروند و یا بیشتر شوند.

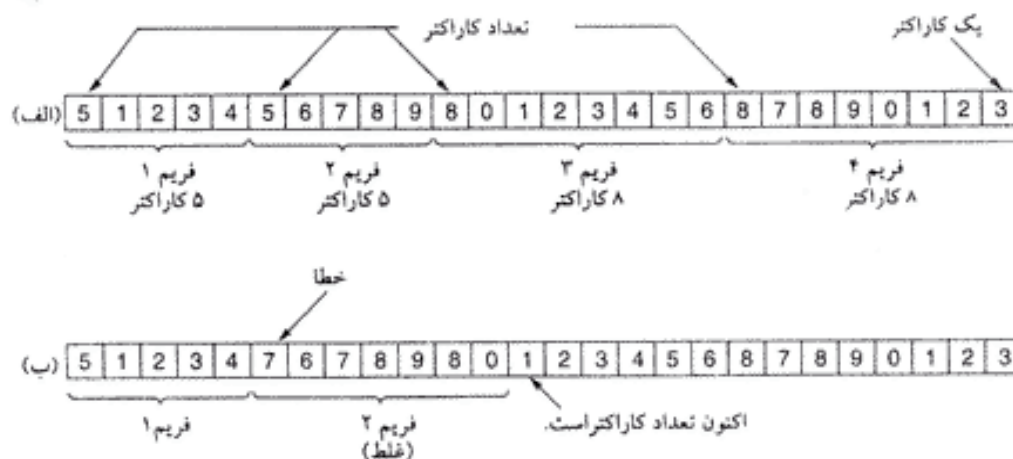
از آنجائیکه تکیه بر زمانبندی برای تعیین ابتدا و انتهای فریمها بسیار خطرناک است، روشهای دیگری برای اینکار ابداع شده، که در این قسمت با چهار تا از آنها آشنا می شوید:

۱. شمارش کاراکترها.
۲. بایت های پرچم، با لاگذاری بایت.
۳. پرچمهای شروع و پایان، با لاگذاری بیت.
۴. حالت های غیرمجاز گذگاری لایه فیزیکی.

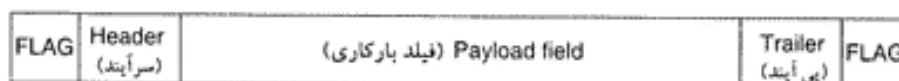
در اولین روش فریم بندی تعداد کاراکترهای فریم در یکی از فیلدهای سرآیند آن نوشته می شود. وقتی این فریم به مقصد می رسد، لایه پیوند داده می تواند به کمک این فیلد ابتدا و انتهای فریم را مشخص کند. در شکل ۳-۴ (الف) چهار فریم با تعداد کاراکترهای ۵، ۵، ۸ و ۸ را می بینید.

اشکال این روش آنست که فیلد تعداد کاراکترها نیز می تواند دچار خطا شود. برای مثال در شکل ۳-۴ (ب)، فیلد تعداد کاراکترها در فریم دوم از ۵ به ۷ تبدیل شده است، و ماشین مقصد دیگر قادر نیست فریمهای بعدی را بدرستی بخواند (چون قادر نیست ابتدای آنها را تشخیص دهد). حتی اگر جمع تطبیقی اشتباه باشد و ماشین مقصد متوجه باشد که خطایی رخ داده، باز هم تشخیص نقطه شروع بعدی برای آن غیرممکن است. درخواست ارسال مجدد نیز کمکی نمی کند، چون ماشین مقصد نمی داند فریمها تا کجا درست بوده، و اشتباه از کجا رخ داده است و نیاز به ارسال مجدد دارد.

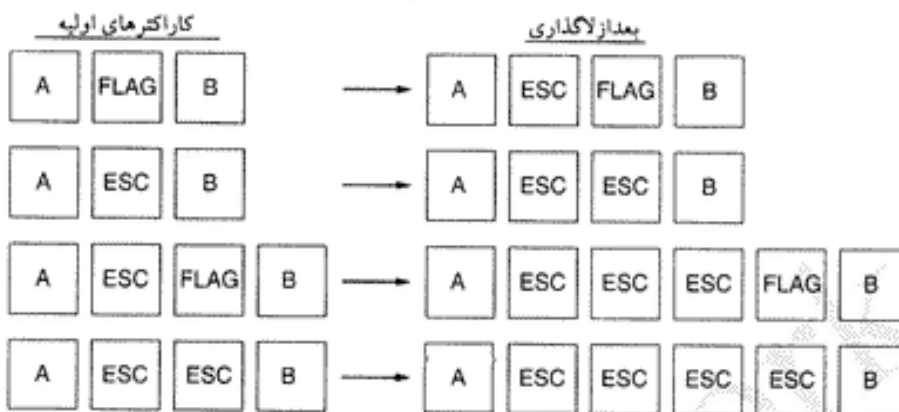
مشکل سنکرون شدن مجدد مبدأ و مقصد بعد از بروز خطا در روش دوم فریم بندی (بایت های پرچم، با لاگذاری بایت) حل شده است، بدین ترتیب که هر فریم با تعدادی بایت خاص شروع و پایان می یابد. در گذشته، بایتهای شروع و پایان متفاوت بودند، ولی در سالهای اخیر از بایتهای یکسانی بعنوان بایت پرچم (flag byte) در



شکل ۳-۴. استریم کاراکترها. (الف) بدون خطا. (ب) با خطا.



(الف)



(ب)

شکل ۳-۵. (الف) تعیین ابتدا و انتهای فریم با استفاده از بایت پرچم. (ب) چهار توالی

بایت فیل و بعد از لاگذاری بایت

شروع و پایان فریم استفاده می شود. این بایتها را در شکل ۳-۵ (الف) با عنوان FLAG ملاحظه می کنید. در این روش اگر گیرنده همزمانی خود با فرستنده را از دست بدهد، فقط کافیسیت با جستجوی بایت پرچم انتهای فریم فعلی را پیدا کند. دو بایت پرچم که پشت سر هم بیایند، بمعنای پایان یک فریم و شروع فریم بعدی هستند. یکی از مشکلات جدی این روش آنست که طرح بیت بایت پرچم می تواند در داده های اصلی نیز وجود داشته باشد (بویژه اگر اطلاعات از نوع برنامه های اجرایی یا اعداد اعشاری باشد). این وضعیت گیرنده را به اشتباه خواهد انداخت. یکی از راه های حل این وضعیت آنست که پروتکل لایه پیوند داده در سمت فرستنده قبل از هر توالی بیت پرچم که در داده اصلی ظاهر می شود، یک بایت گریز (escape byte) خاص قرار دهد. لایه پیوند داده مقصد این بایتها را حذف کرده، و داده های اصلی را به لایه شبکه تحویل می دهد. به این تکنیک لاگذاری بایت (byte stuffing) یا لاگذاری کاراکتر (character stuffing) گفته می شود. با این روش بایت پرچم بآسانی قابل تشخیص است، چون قبل از آن بایت گریز وجود ندارد.

اما حالا سوال دیگری پیش می آید: اگر در وسط داده اصلی طرحی مشابه بایت گریز وجود داشت، چه اتفاقی می افتد؟ جواب اینست که قبل از این بایت هم یک بایت گریز قرار داده می شود؛ عبارت دیگر دو بایت گریز پشت سر هم یعنی یک بایت گریز در داده اصلی. در شکل ۳-۵ (ب) چند نمونه از حالتی که می تواند پیش آید، آورده شده است. در هر مورد آن چیزی که گیرنده می گیرد، دقیقاً مشابه آن چیزیست که فرستنده ارسال کرده است. تکنیک لاگذاری بایت که در شکل ۳-۵ نشان داده شده، شکل ساده شده آن چیزیست که در پروتکل PPP (یکی از مهمترین پروتکل های ارتباط با اینترنت در کامپیوترهای شخصی) مورد استفاده قرار می گیرد. (بعدها در همین فصل درباره PPP صحبت خواهیم کرد.)

یکی از معایب بزرگ فریم بندی بایت های پرچم با لاگذاری بایت وابستگی شدید آن به کاراکترهای ۸-بیتی است، و همانطور که می دانید تمام کدها ۸-بیتی نیستند (برای مثال، در استاندارد یونی کد از کاراکترهای ۱۶-بیتی استفاده می شود). فرض ۸-بیتی بودن کاراکترها در مکانیزم فریم بندی یکی از مشکلات جدی آن محسوب می شود، به همین دلیل روش جدیدی که در آن طول کاراکتر می تواند متغیر باشد، ابداع شده است.

در این روش جدید طول کاراکترها اهمیتی ندارد، و فریمها می توانند تعداد بیتهای دلخواه داشته باشند. طرز کار این تکنیک جدید چنین است: هر فریم با طرح بیت خاصی (01111110 - که در واقع یک بایت پرچم است) شروع می شود. هر گاه لایه پیوند داده در سمت فرستنده پنج 1 پشت سر هم در داده اصلی دید، بطور خودکار یک 0 بعد از آن قرار می دهد. این روش، که به آن لاگذاری بیت (bit stuffing) گفته می شود، بسیار شبیه لاگذاری بایت است. وقتی گیرنده پنج 1 متوالی ببیند که یک 0 پشت سر آنها آمده، بطور خودکار این 0 را حذف می کند. لاگذاری بیت نیز مانند لاگذاری بایت بکلی از دید لایه شبکه در هر دو کامپیوتر پنهان (شفاف) است. اگر در داده کاربر طرح بیت 01111110 وجود داشته باشد، لایه پیوند داده فرستنده آنرا به 011111010 تبدیل می کند، و در سمت گیرنده این 0 اضافی حذف شده و طرح بیت 01111110 به لایه بالاتر تحویل داده می شود. به یک مثال در شکل ۳-۶ توجه کنید.

(الف) 011011111111111111110010

(ب) 01101111101111101111010010

بیت های لاگذاری

(ج) 011011111111111111110010

شکل ۳-۶. لاگذاری بیت. (الف) داده اولیه. (ب) دادهها بصورتی که روی خط فیزیکی ارسال می شود. (ج) دادهها بصورتی که در گیرنده دریافت می شود.

در روش لاگذاری بیت نیز محدوده فریم با استفاده از پرچمهای شروع و پایان مشخص می شود، و گیرنده می تواند از آنها برای سنکرون شدن با فرستنده استفاده کند.

آخرین روش فریم بندی فقط در شبکه هایی قابل بکارگیری است که در کدگذاری لایه فیزیکی آنها نوعی افزونگی (redundancy) وجود داشته باشد. برای مثال در برخی از شبکه های LAN هر بیت داده با دو بیت فیزیکی نمایش داده می شود: بیت 1 با زوج بالا-پائین، و بیت 0 با زوج پائین-بالا. بدین ترتیب هر بیت داده دارای نوعی تغییر ولتاژ است، که تشخیص آنرا برای گیرنده ساده تر می کند. در چنین شبکه هایی زوج بالا-بالا و پائین-پائین برای دادهها استفاده نمی شود، و می توان از آنها برای مشخص کردن محدوده فریمها سود برد.

لازم به ذکر است که در بسیاری از پروتکل های لینک داده برای اطمینان بیشتر از ترکیب روش شمارش کاراکترها با یکی دیگر از تکنیکهای گفته شده استفاده می شود. در این روش، انتهای فریم با استفاده از فیلد تعداد کاراکترها مشخص می شود، ولی فقط زمانی مورد قبول قرار می گیرد که جمع تطبیقی فریم نیز معتبر بوده و در این نقطه طرح بیت پایان فریم وجود داشته باشد. اگر چنین نباشد، گیرنده طرح بیت پایان فریم را در نقاط دیگر جستجو خواهد کرد.

۳-۱-۳ کنترل خطا

بعد از حل مسئله ابتدا و انتهای فریمها، نوبت به مسئله بعدی می رسد: چگونه می توان تمام فریمها را سالم و با ترتیب صحیح به مقصد رساند؟ فرض کنید فرستنده فقط فریمها را می فرستد و کاری ندارد که آنها به مقصد می رسند یا خیر. این وضعیت برای سرویسهای غیر متصل بدون تصدیق دریافت خوب است، ولی برای سرویسهای قابل اعتماد (مانند سرویس اتصال-گرا یا تصدیق دریافت) مسلماً خوب نیست.

یک سرویس قابل اعتماد باید بنحوی از رسیدن بستهها به مقصد و آنچه در آنجا اتفاق می افتد، مطلع شود. معمولاً در این موارد پروتکل درخواست می کند که یک فریم کنترلی خاص (که محتوی تصدیق یا عدم تصدیق

دریافت صحیح فریمهاست) به فرستنده باز پس فرستاده شود. اگر فرستنده تصدیق مثبت دریافت کند، مطمئن می‌شود که فریم به سلامت به مقصد رسیده است. اما تصدیق منفی نشان می‌دهد که اوضاع روبراه نیست، و فریم باید مجدداً فرستاده شود.

مشکل دیگر اینجاست که گاهی (در اثر اشکالات سخت‌افزاری) یک فریم بکلی گم و ناپدید می‌شود. در این حالت گیرنده هیچ عکس‌العملی نشان نمی‌دهد، چون اساساً چیزی نگرفته که عکس‌العمل نشان دهد. بروشنی پیداست که در این حالت پروتکل سمت فرستنده نا ابد منتظر دریافت تصدیق از گیرنده می‌شود، تصدیقی که هرگز نخواهد رسید.

این مشکل را می‌توان با تعبیه یک تایمر در لایه پیوند داده حل کرد. وقتی فرستنده فریمی را می‌فرستد، تایمر را هم راه‌اندازی می‌کند. زمانی که این تایمر اندازه می‌گیرد آنقدر طولانی هست که بتوان با اطمینان گفت «فریم باید به مقصد رسیده، و تصدیق دریافت آن برگشته باشد». اگر همه چیز خوب پیش رفته باشد، معمولاً قبل از اینکه زمان تایمر به انتها برسد، تصدیق دریافت فریم به فرستنده برمی‌گردد (و تایمر ریست می‌شود).

اگر فریم یا پاسخ آن در راه گم شوند، تایمر در انتهای زمان مقرر اخطار می‌دهد؛ و ساده‌ترین راه حل همانا ارسال مجدد فریم است. اما ارسال چندباره فریمها این خطر را در بر دارد که چند تا از این فریمهای یکسان به مقصد برسند و به لایه شبکه تحویل داده شوند. برای اجتناب از این وضعیت، فرستنده به هر فریم یک شماره ترتیبی می‌دهد تا گیرنده بتواند فریمهای مشابه و تکراری را تشخیص دهد.

با تمام این تمهیدات (تایمر و شماره ترتیبی فریمها) می‌توان مطمئن بود که از هر فریم یک (و فقط یک) نسخه به لایه شبکه می‌رسد - و این یکی از مهمترین وظایف لایه پیوند داده است. در ادامه این فصل خواهید دید که لایه پیوند داده چگونه این وظیفه را انجام می‌دهد.

۴-۱-۳ کنترل جریان

یکی دیگر از مسائل مهم در طراحی لایه پیوند داده (و لایه‌های بالاتر) اینست که با فرستنده‌هایی که سریعتر از توان دریافت گیرنده مبادرت به ارسال اطلاعات می‌کنند، چه باید کرد؟ اگر کامپیوتر طرف فرستنده قویتر از گیرنده باشد (و یا بار کاری کمتری داشته باشد)، این وضعیت براحتی می‌تواند پیش بیاید. در این حالت گیرنده در سیلاب فریمهای ارسالی از فرستنده غرق می‌شود. حتی اگر کانال ارتباطی کاملاً عاری از خطا باشد، لحظه‌ای می‌رسد که گیرنده دیگر قادر به پردازش فریمهای ارسال شده نیست، و برخی از آنها را از دست می‌دهد. روشن است که باید کاری برای جلوگیری از این وضعیت کرد.

دو رهیافت برای مقابله با این وضعیت بکار گرفته می‌شود. در رهیافت اول، که کنترل جریان بر اساس بازخور (feedback-based flow control) نام دارد، این گیرنده است که آمادگی خود را برای دریافت اطلاعات بیشتر به فرستنده اعلام می‌کند (یا حداقل اعلام می‌کند در چه وضعیتی است). در رهیافت دوم، کنترل جریان بر اساس نرخ (rate-based flow control)، پروتکل مکانیزمی دارد که بدون استفاده از بازخور گیرنده نرخ ارسال اطلاعات را محدود می‌کند. در این فصل با روشهای کنترل جریان بر اساس بازخور آشنا می‌شوید، ولی از آنجائیکه رهیافت دوم هرگز در لایه پیوند داده کاربرد ندارد، توضیح درباره آنرا به فصل ۵ موقوف می‌کنیم.

انواع مختلفی از کنترل جریان بر اساس بازخور وجود دارد، ولی همه آنها اصول مشترکی دارند: پروتکل قواعد تعریف شده‌ای دارد که زمان ارسال فریم بعدی را مشخص می‌کند. طبق این قواعد فرستنده نمی‌تواند فریم بعدی را بفرستد، مگر اینکه (بطور صریح یا ضمنی) اجازه گیرنده را دریافت کرده باشد. مثلاً، وقتی اتصال برقرار می‌شود، گیرنده می‌تواند به فرستنده بگوید: «اکنون می‌توانی ۱۱ فریم بفرستی، ولی بعد از آن تا اجازه نداده‌ام چیزی نفرست.»

۲-۳ کشف و تصحیح خطا

همانطور که در فصل ۲ دیدید، سیستم تلفن سه بخش عمده دارد: سونیچها، ترانکها، و حلقه‌های محلی. در اکثر کشورها توسعه یافته دو بخش اول تماماً دیجیتال هستند. اما قسمت اعظم حلقه‌های محلی کماکان آنالوگ است، که بایستی با صرف هزینه‌های هنگفت در آینده به دیجیتال تبدیل شود. با اینکه در بخش دیجیتال خطا پندرت روی می‌دهد، نرخ آن در حلقه‌های محلی آنالوگ همچنان بالاست. علاوه بر آن، مخابرات بیسیم نیز سرعت گسترش می‌یابد، که نرخ خطا در این قبیل سیستمها چندین برابر کانالهای فیبر نوری است. نتیجه اخلاقی: فعلاً تا مدتها باید با خطاهای انتقال در سیستمهای مخابراتی بسازیم. در این قسمت خواهید دید چگونه.

خصلت خطا به منبع آن بستگی دارد؛ برای مثال، در سیستمهای رادیویی خطا بصورت فورانی (burst) رخ می‌دهد، نه تکی. این نوع خطا مزایا و معایبی دارد. توجه داشته باشید که کامپیوترها اطلاعات خود را بصورت بسته‌ای ارسال می‌کنند. اگر هر بسته داده ۱۰۰۰ بیت و نرخ خطا نیز ۱ در ۱۰۰۰ باشد، می‌توان انتظار داشت که (در حالت غیر فورانی) تقریباً تمام بسته‌ها با خطا به مقصد برسند. اما اگر خطا بصورت فورانه‌ای ۱۰۰ بیته رخ دهد، بطور متوسط فقط یک یا دو بسته را خراب خواهد کرد. عیب بزرگ خطاهای فورانی آنست که کشف و تصحیح خطا در آنها بسیار دشوارتر است.

۱-۲-۳ گداهای تصحیح خطا

طراحان شبکه دو استراتژی کلی برای مقابله با خطاها توسعه داده‌اند. یک راه اضافه کردن اطلاعات پراکنده به هر بلوک از داده‌هاست، بطوریکه گیرنده بتواند داده واقعی را از آن استنتاج کند. در روش دیگر فقط آنقدر اطلاعات اضافی به داده اصلی اضافه می‌شود که گیرنده از وقوع یا عدم وقوع خطا آگاهی یابد، و در صورت لزوم تکرار ارسال را خواستار شود. استراتژی اول گداهای تصحیح خطا (error-correcting codes) و استراتژی دوم گداهای کشف خطا (error-detecting codes) نام دارند. به کاربرد گداهای تصحیح خطا اغلب تصحیح پیشگیرانه خطا نیز گفته می‌شود.

هر یک از این تکنیکها جایگاه خاص خود را دارند. در کانالهای قابل اطمینان، مانند فیبر نوری، مقرون بصرفه‌تر است که از گداهای کشف خطا استفاده کرده و بسته‌های معدودی را که خراب می‌شوند، دوباره ارسال کنیم. اما در کانالهایی مانند لینکهای بیسیم که پر از خطا هستند، بهتر است از تکنیکهای تصحیح خطا استفاده کرده و اجازه دهیم گیرنده خود داده واقعی را بدست آورد (چون با احتمال زیاد ارسال مجدد بسته‌ها هم عاری از خطا نخواهد بود). برای مقابله با خطاها، ابتدا باید بدانیم خطا واقعاً چیست. معمولاً، یک فریم m بیت داده اصلی (یعنی، پیام) r بیت داده پراکنده (یا اطلاعات چک‌کننده) دارد، که در مجموع n بیت می‌شود ($n = m + r$). به این واحد n بیتی (داده‌های اصلی و پراکنده) اغلب کلمه‌گد n بیتی گفته می‌شود.

دو کلمه‌گد 10001001 و 10110001 را در نظر بگیرید: براحتی می‌توان مشخص کرد که این دو کلمه چند اختلاف دارند. در این مورد ۳ بیت اختلاف وجود دارد. برای تعیین تعداد اختلاف‌ها می‌توان دو کلمه‌گد را با هم XOR (انحصاری) کرد، و تعداد 1 ها را شمرد:

```
10001001
10110001
00111000
```

به تعداد اختلافهای دو کلمه‌گد فاصله همینگ (Hamming distance) گفته می‌شود (Hamming, 1950). اهمیت این فاصله در آنجاست که می‌توان ثابت کرد برای تبدیل شدن اتفاقی دو کلمه با فاصله d ، بایستی d خطای تک‌بیتی روی دهد.

در اکثر سیستمهای انتقال، تمامی 2^m حالت ممکنه داده اصلی مجاز است، ولی بدلیل روش محاسبه بیت های افزونگی، تمام 2^n حالت کلمه کُد مجاز نیست. با توجه به الگوریتم محاسبه بیت های افزونگی، می توان لیستی از تمام حالت های مجاز کلمه کُد بدست آورد، و از این لیست دو کلمه ای که کمترین فاصله همینگ را دارند، پیدا کرد. این فاصله فاصله همینگ الگوریتم یا کُد مورد نظر است.

خصوصیات تصحیح خطا یا کشف خطای یک کُد به فاصله همینگ آن بستگی دارد. برای کشف d خطا، به کُدی با فاصله همینگ $d + 1$ نیاز داریم، چون با چنین کُدی هیچ d خطای تک بیتی وجود ندارد که بتواند یک کلمه کُد مجاز را به کلمه کُد مجاز دیگر تبدیل کند. اگر گیرنده کلمه کُد غیر مجازی دریافت کرد، می تواند با اطمینان بگوید که خطایی رخ داده است. بهمین ترتیب، برای تصحیح d خطا، به کُدی با فاصله $d + 1$ نیاز داریم، چون در این حالت کلمات کُد چنان از هم فاصله دارند که حتی با بروز d خطا، کلمه کُد خراب شده هنوز نزدیکترین فاصله را با کلمه کُد اصلی دارد، و تشخیص آن براحتی ممکن است.

بعنوان نمونه ای از کُد های کشف خطا، کُدی با یک بیت توازن (parity bit) را در نظر بگیرید. این بیت توازن بگونه ای انتخاب می شود که تعداد بیت های 1 کلمه کُد همواره زوج (یا فرد) شود. برای مثال، اگر بخواهیم کلمه 1011010 را با توازن زوج (even parity) ارسال کنیم، یک بیت 0 به انتهای آن اضافه می کنیم (10110100)؛ اما اگر بخواهیم همین کلمه را با توازن فرد (odd parity) ارسال کنیم، باید یک بیت 1 به انتهای آن اضافه کنیم (10110101). کُدی با یک بیت توازن دارای فاصله همینگ 2 است، چون هر خطای تک بیتی کلمه کُدی با توازن اشتباه تولید می کند. این کُد می توان یک خطا در هر کلمه را آشکار کند.

بعنوان یک نمونه ساده از کُد های تصحیح خطا، کُدی را در نظر بگیرید که فقط چهار کلمه کُد مجاز دارد:

0000000000, 0000011111, 1111100000, 1111111111

فاصله همینگ این کُد 5 است، بنابراین می تواند دو خطا را تصحیح کند. اگر گیرنده کلمه کُدی بصورت 0000000111 دریافت کند، می داند که کلمه اصلی باید 0000011111 بوده باشد. اما اگر سه خطا کلمه 0000000000 را به 0000000111 تبدیل کرده باشد، دیگر نمی توان آنرا بدرستی تصحیح کرد.

فرض کنید می خواهیم کُدی با m بیت داده اصلی و r بیت افزونگی طراحی کنیم که بتواند تمام خطاهای تک بیتی را تصحیح کند. هر یک از 2^m پیام مجاز دارای n کلمه کُد غیر مجاز است که با آن 1 فاصله دارد (این را می توان بسادگی از معکوس کردن هر یک از بیت های کلمه کُد n بیتی فهمید). بنابراین هر یک از 2^m پیام مجاز به طرحی اختصاصی با $n + 1$ بیت نیاز دارد. از آنجائیکه تعداد ترکیبات ممکنه کلمه کُد 2^n است، بایستی داشته باشیم: $2^n \leq (n + 1)2^m$. با قرار دادن $n = m + r$ در این رابطه، داریم: $2^n \leq (m + r + 1)2^m$. با داشتن m ، از این رابطه حداقل بیت های افزونگی لازم (r) برای تصحیح خطاهای تک بیتی بدست می آید.

همینگ در یکی از مقالات خود (1950) روشی برای بدست آوردن این حداقل معرفی کرد. وی بیت های کلمه کُد را از چپ براسست شماره گذاری کرد. بیت هایی که توانایی از 2 هستند (1، 2، 4، 8، 16، و غیره)، بیت های چک کننده اند؛ سایر بیت ها (3، 5، 6، 7، 9، و غیره) بیت های پیام (m) هستند. هر بیت چک کننده توازن مجموعه ای از بیت (از جمله خودش) را زوج (یا فرد) می کند. هر بیت می تواند در بیش از یک مجموعه توازن محاسبه شود. برای دیدن اینکه کدام بیت های چک کننده در محاسبه توازن بیت داده ای در موقعیت k دخالت دارند، k را بصورت مجموع توانهای 2 می نویسیم. برای مثال، $11 = 1 + 2 + 8$ ، و $29 = 1 + 4 + 8 + 16$. هر بیت فقط با بیت های چک کننده ای که در موقعیت های بدست آمده از مجموع توانهای 2 قرار دارند، چک می شود (مثلاً، بیت موقعیت 11 فقط با بیت های چک کننده 1، 2، و 8 چک می شود).

وقتی یک کلمه کُد به گیرنده می رسد، گیرنده یک شمارنده را 0 می کند. سپس تمام بیت های چک کننده (k) را

کاراکتر	ASCII	بیت های چک کننده
H	1001000	00110010000
a	1100001	10111001001
m	1101101	11101010101
m	1101101	11101010101
i	1101001	01101011001
n	1101110	01101010110
g	1100111	01111001111
	0100000	10011000000
c	1100011	11111000011
o	1101111	10101011111
d	1100100	11111001100
e	1100101	00111000101

ترتیب انتقال بیت ها

شکل ۳-۷. استفاده از کد همینگ برای تصحیح خطاهای فورانی.

از نظر توازن چک می کند ($k = 1, 2, 4, 8, \dots$). اگر توازن k درست نباشد، گیرنده k را به شمارنده اضافه می کند. اگر پس از پایان این عملیات شمارنده همچنان 0 باشد، کلمه کُد صحیح تلقی و قبول می شود. اگر شمارنده 0 نباشد، حتماً شماره بیت خطا را نشان می دهد. برای مثال، اگر توازن بیت های چک کننده 1، 2، و 8 اشتباه باشد، بیت 11 غلط است، چون این تنها بیتی است که با بیت های چک کننده 1، 2، و 8 چک می شود. در شکل ۳-۷ چند کاراکتر آسکی ۷-بیتی را که با کد همینگ ۱۱-بیتی کُد شده اند، می بینید. فراموش نکنید که داده های اصلی در موقعیتهای 3، 5، 6، 7، 9، 10، و 11 قرار دارند.

کدهای همینگ فقط می توانند خطاهای تک بیتی را تصحیح کنند. با این حال روشی وجود دارد که اجازه می دهد تا این کُد خطاهای فورانی را نیز تصحیح کند. در این روش k کلمه کُد متوالی بصورت ماتریس (یک کلمه کُد در هر سطر) چیده می شوند. معمولاً، این کلمات تک به تک (از چپ بر راست) ارسال می شوند. برای تصحیح خطاهای فورانی، بایستی داده ها را بصورت ستونی (باز هم از چپ بر راست) ارسال کرد. وقتی k بیت اول (ستون اول) ارسال شد، نوبت به ستون دوم (و سپس ستونهای بعدی) می رسد (شکل ۳-۷ را ببینید). وقتی این فریم به گیرنده رسید، ماتریس از نو (ستون به ستون) ساخته می شود. اگر یک خطای فورانی به طول k رخ داده باشد، حداکثر یک بیت در هر کلمه کُد تغییر خواهد کرد، و از آنجائیکه کُد همینگ می تواند یک خطا را تصحیح کند، تمام بلوک قابل تصحیح خواهد بود. در این روش برای مصون کردن k بیت داده در مقابل خطاهای فورانی با طول k (یا کمتر)، از k^2 بیت چک کننده استفاده شده است.

۲-۲-۳ کدهای کشف خطا

کدهای کشف خطا در لینکهای بیسیم، که در مقایسه با سیم مسی و فیبر نوری بطور وحشتناکی نویزی هستند، کاربرد گسترده ای دارد. بدون این کدها شاید اساساً نتوان چیزی روی این لینکها رد و بدل کرد. اما در سیمهای مسی و فیبرهای نوری نرخ خطا بسیار کمتر است، و تشخیص خطا و ارسال مجدد بسته هایی که (ندرتاً) خراب می شوند، کاملاً کفایت می کند.

بعنوان مثال، کانالی را در نظر بگیرید که نرخ خطا در آن 1 در 10^6 و خطاها غیر فورانی هستند؛ اندازه هر بلوک را هم 1000 بیت فرض می کنیم. برای داشتن ویژگی تصحیح خطا، هر بلوک 1000 بیتی به 10 بیت چک کننده نیاز دارد، بعبارت دیگر برای ارسال 1 Mb داده باید 10 kb اطلاعات افزونگی (بیت های چک کننده) را نیز به همراه آن

بفرستیم. اما برای کشف خطا فقط یک بیت توازن در هر بلوک کافیهست. در این روش بار اضافی کشف خطا + ارسال مجدد یک بلوک خراب برای 1 Mb داده فقط 2001 بیت است، که در مقایسه با 10,000 بیت کُد همینگ بسیار کمتر است.

اگر در هر بلوک از یک بیت توازن برای کشف خطا استفاده کنیم و یک خطای فورانی رخ دهد، احتمال اینکه بتوانیم خطا را کشف کنیم فقط 50٪ است، که بهیچوجه قابل قبول نیست. اما با تشکیل ماتریسی با n ستون و k سطر (که در بالا توضیح دادیم) اوضاع بنحو قابل توجهی بهتر خواهد شد. در این روش برای هر ستون یک بیت توازن محاسبه، و در آخرین سطر ماتریس نوشته می شود. هنگام ارسال نیز این ماتریس بصورت ستونی فرستاده می شود. گیرنده بعد از دریافت کل ماتریس، تمام بیت های توازن را چک می کند؛ و اگر هر یک از این بیت ها غلط باشد، ارسال مجدد ماتریس را درخواست می کند. این کار تا زمانی که ماتریس بطور کامل و بدون خطای توازن به دست گیرنده برسد، تکرار خواهد شد.

روش فوق می تواند خطاهای فورانی با طول حداکثر n بیت را آشکار کند، چون در این حالت فقط یک بیت در هر ستون تغییر خواهد کرد. اما اگر یک خطای فورانی با طول $n + 1$ رخ دهد بگونه ای که فقط بیت اول و آخر را تغییر دهد (و سایر بیت ها تغییر نکنند)، نمی توان آنرا کشف کرد، زیرا بیت اول و آخر در یک سطر قرار می گیرند و توازن این سطر بدون تغییر خواهد ماند. (یک خطای فورانی الزاماً بمعنای معکوس شدن تمام بیت ها نیست؛ فقط می توان از معکوس شدن بیت اول و آخر مطمئن بود.) اگر طول خطای فورانی خیلی زیاد باشد یا تعدادی خطای فورانی کوتاه و پشت سر هم رخ دهد، احتمال اینکه یکی از ستونها تصادفاً صاحب توازن درست شود، 50٪ است، بنابراین احتمال اینکه چنین بلوکی (به اشتباه) صحیح تلقی شود، 2^{-n} خواهد بود.

با اینکه روش فوق در مواردی کفایت می کند، اما در عمل از روش دیگری استفاده می شود: کُد چندجمله ای (polynomial code)، که به CRC (چک افزونگی چرخه ای - Cyclic Redundancy Check) نیز معروفست. در کدهای چندجمله ای مبنای این است که هر رشته یک چندجمله ایست با ضرایب 0 و 1. با این فرض، یک فریم k -بیتی معادلت با عبارتی k جمله ای، با ضرایب x^{k-1} تا x^0 . این چندجمله ای از درجه $k - 1$ است. بالارزترین بیت (متهی الیه سمت چپ) ضریب x^{k-1} است، بیت بعدی ضریب x^{k-2} ، والی آخر. برای مثال، رشته 110001 دارای 6 بیت است بنابراین نشاندهنده یک شش جمله ایست با ضرایب 1، 0، 0، 0، 1، و 1، که می توان آنرا چنین نوشت: $x^5 + x^4 + x^0$.

محاسبات چندجمله ایها در مدول 2 (و طبق قوانین جبر میدان) انجام می شود. در جمع و تفریق 2 بر 1 نادیده گرفته می شود، بعبارت دیگر شبیه XOR است. برای مثال،

$$\begin{array}{r} 10011011 \\ + 11001010 \\ \hline 01010001 \end{array} \quad \begin{array}{r} 00110011 \\ + 11001101 \\ \hline 11111110 \end{array} \quad \begin{array}{r} 11110000 \\ - 10100110 \\ \hline 01010110 \end{array} \quad \begin{array}{r} 01010101 \\ - 10101111 \\ \hline 11111010 \end{array}$$

تقسیم درست مانند تقسیم باینری است، با این تفاوت که تفریق ها در مدول 2 (مانند بالا) انجام می شود. در هنگام استفاده از روش کُد چندجمله ای، فرستنده و گیرنده بایستی از قبل بر سر یک چندجمله ای مولد (generator polynomial)، که آنرا $G(x)$ می نامیم، توافق کنند. بالارزترین (چپ ترین) و کم ارزشترین (راست ترین) بیت های چندجمله ای مولد باید 1 باشد. برای محاسبه مجموع چک (checksum) یک فریم m -بیتی (که چندجمله ای متناظر با آن $M(x)$ است)، این فریم باید طولانیتر از چندجمله ای مولد باشد. ایده آنست که یک مجموع چک به انتهای فریم اصلی چسبانده شود، بگونه ای که فریم حاصله بر $G(x)$ قابل تقسیم باشد. اگر

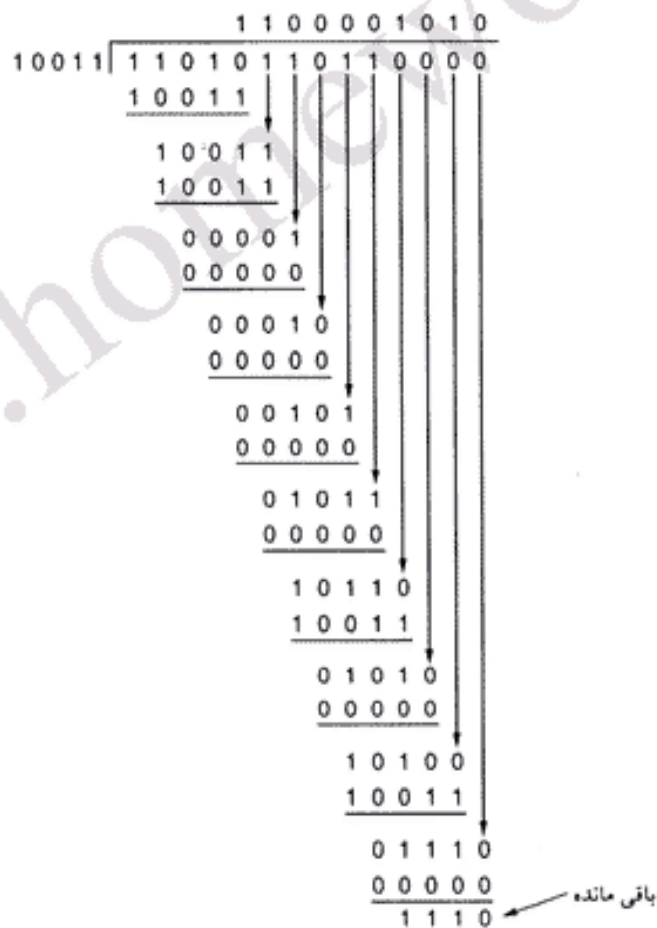
تقسیم این فریم بر $G(x)$ در سمت گیرنده باقیمانده آورد، معلوم می شود که خطایی رخ داده است. الگوریتم محاسبه مجموع چک چنین است:

۱. فرض می کنیم چندجمله ای $G(x)$ از درجه r است. r بیت 0 به سمت راست فریم اضافه می کنیم تا تعداد بیت های آن به $m + r$ برسد. این چندجمله ای معادل $x^m M(x)$ خواهد شد.
 ۲. رشته $x^m M(x)$ را (در مدول 2) بر $G(x)$ تقسیم می کنیم.
 ۳. باقیمانده را (که همیشه r بیت یا کمتر دارد) از $x^m M(x)$ کم می کنیم (این تفریق هم در مدول 2 انجام می شود). حاصل تفریق همان فریم مورد نظر (فریم اولیه + مجموع چک) است، که آنرا $T(x)$ می نامیم.
- در شکل ۸-۳ طرز محاسبه مجموع چک برای فریم 1101011011 را با مولد $G(x) = x^4 + x + 1$ ملاحظه می کنید.

فریم: 1101011011

مولد: 10011

پیام بعد از اضافه شدن ۴ بیت: 11010110110000



فریم اضافه شده: 11010110111110

شکل ۸-۳ محاسبه مجموع چک کد چندجمله ای.

همانطور که براحتی معلوم می شود، $T(x)$ بر $G(x)$ (در مدول 2) بخش پذیر است (چون وقتی باقیمانده تقسیم را از مقسوم کم کنیم، عدد حاصله بطور حتم بر مقسوم علیه بخش پذیر خواهد بود). بطور مثال، اگر 210,278 را (در مبنای 10) بر 10,941 تقسیم کنیم، باقیمانده 2399 می شود که اگر آنرا از 210,278 کم کنیم، آنچه باقی می ماند (207,879) بر 10,941 بخش پذیر خواهد بود.

حال اجازه دهید قدرت این روش را بررسی کنیم. این روش چه نوع خطاهایی را می تواند کشف کند؟ فرض کنید خطایی رخ داده، و بجای $T(x)$ رشته $T(x) + E(x)$ به گیرنده رسیده است، بطوریکه هر بیت 1 در $E(x)$ متناظر با یک بیت تغییر یافته است (بعبارت دیگر، اگر در این عبارت، $E(x)$ ، k بیت 1 وجود داشته باشد، k خطای تکبیتی رخ داده است). خطای فورانی نیز عبارتست از خطایی که با یک بیت 1 شروع و ختم شود، و بین آنها هر ترکیبی از 0 و 1 می تواند وجود داشته باشد.

وقتی این فریم به مقصد می رسد، گیرنده آنرا بر $G(x)$ تقسیم می کند (بعبارت دیگر $[T(x)+E(x)]/G(x)$ را محاسبه می کند). از آنجائیکه $T(x)/G(x) = 0$ ، این تقسیم معادل $E(x)/G(x)$ است. همانطور که می بینید، اگر خطای رخ داده دقیقاً طرحی شبیه $G(x)$ نداشته باشد، بطور مسلم آشکار خواهد شد.

فرض کنید یک خطای تکبیتی رخ داده است، یعنی، $E(x) = x^i$ (که i بیت خطاست). اگر $G(x)$ بیش از دو جمله داشته باشد، $E(x)$ هرگز بر آن بخش پذیر نخواهد بود - پس، این روش می تواند تمام خطاهای تکبیتی را آشکار کند.

اگر دو خطای تکبیتی جدا از هم رخ دهد، بطوریکه $E(x) = x^i + x^j$ (که در آن $i > j$)، می توان $E(x)$ را به صورت $x^j(x^{i-j} + 1)$ تجزیه کرد. اگر $G(x)$ بر x بخش پذیر نباشد، شرط کافی برای اینکه تمام خطاهای دوبیتی قابل کشف باشد آن است که $E(x)$ عبارت $x^k + 1$ را (برای تمام k های کوچکتر از $i - j$) بخش نکند. چندجمله ای های ساده و از درجه پائینی می شناسیم که می توان با آنها فریمهای نسبتاً طولیل را محافظت کرد. مثلاً، چندجمله ای $x^{15} + x^{14} + 1$ هیچ عبارت $x^k + 1$ برای تمام k های کوچکتر از 32,768 بخش نمی کند.

اگر تعداد خطاهای رخ داده عددی فرد باشد، تعداد جملات $E(x)$ نیز فرد خواهد بود (برای مثال، تعداد جملات $x^2 + 1 + x^4$ فرد است، ولی $x^2 + 1$ چنین نیست). جالبست بدانید که هیچ چندجمله ای با تعداد جملات فرد وجود ندارد که (در مدول 2) بر $x + 1$ بخش پذیر باشد. بدین ترتیب اگر $G(x)$ را طوری انتخاب کنیم که بر $x + 1$ بخش پذیر باشد، می توانیم هر خطایی که تعداد بیت های تغییر کرده فرد باشد را کشف کنیم.

برای اثبات اینکه هیچ چندجمله ای فرد وجود ندارد که بر $x + 1$ بخش پذیر باشد، فرض کنید $E(x)$ چندجمله ای فردیست که چنین خاصیتی دارد (بر $x + 1$ بخش پذیر است). اگر از $x + 1$ فاکتور بگیریم، $E(x)$ بصورت $Q(x)(x + 1)$ در می آید. حال $E(1) = (1 + 1)Q(1)$ را محاسبه می کنیم. از آنجائیکه (در مدول 2) $1 + 1 = 0$ ، $E(1) = 0$ باید باشد. اما اگر تعداد جملات $E(x)$ فرد باشد، قرار دادن 1 بجای x در آن همیشه نتیجه 1 می دهد. بنابراین فرض ما نمی تواند درست باشد، و هیچ چندجمله ای فرد بر $x + 1$ بخش پذیر نیست.

بالاخره، و از همه مهمتر، یک کد چندجمله ای با r بیت چک کننده تمام خطاهای فورانی با طول کمتر یا مساوی r را آشکار می کند. یک خطای فورانی با طول k را می توان با $(x^k + x^{k-1} + \dots + 1)$ نشان داد، که در آن i نقطه شروع خطای فورانی از سمت راست فریم است. اگر مولد $G(x)$ دارای جمله x^0 باشد، بر x^k بخش پذیر نخواهد بود؛ بنابراین اگر درجه عبارت داخل پرائتر از درجه $G(x)$ کمتر باشد، باقیمانده تقسیم هرگز نمی تواند 0 شود.

اگر طول خطای فورانی $r + 1$ باشد، باقیمانده تقسیم بر $G(x)$ صفر می شود فقط و فقط اگر طرح بیت خطا با $G(x)$ یکسان باشد. طبق تعریف بیت های اول و آخر خطای فورانی باید 1 باشند، بنابراین یکسان بودن آنها به $r - 1$ بیت میانی بستگی دارد. اگر تمام ترکیبات این $r - 1$ بیت را یکسان فرض کنیم، احتمال بروز این وضعیت $\frac{1}{2^{r-1}}$ خواهد بود.

همچنین می توان نشان داد که اگر طول خطای فورانی از $1 + 2^r$ بزرگتر باشد یا چند خطای فورانی کوتاهتر رخ دهد، احتمال کشف نشدن خطا (با فرض یکسان بودن تمام ترکیبات ممکنه) $\frac{1}{2^r}$ است.

برخی از چند جمله ایها بصورت استاندارد بین المللی در آمده اند، که از میان آنها می توان به چند جمله ای زیر (که در IEEE 802 از آن استفاده می شود) اشاره کرد:

$$1 + x + x^2 + x^4 + x^5 + x^7 + x^8 + x^{10} + x^{11} + x^{12} + x^{16} + x^{22} + x^{23} + x^{26} + x^{32}$$

از ویژگیهای جالب این چند جمله ای آن است که هر نوع خطای فورانی با طول 32 یا کمتر، و خطاهای فورانی که تعداد بیتهای تغییر کرده فرد باشد، را آشکار می کند.

با اینکه بنظر می رسد محاسبه مجموع چک و تست آن کار پیچیده ای باشد، پترسون و براون (1961) نشان دادند که می توان این کار را با یک مدار شیفت رجیستر (shift register) ساده بصورت سخت افزاری انجام داد. در واقع، این مدار در تمام کارتهای شبکه تعبیه شده است، و بسیاری از خطوط نقطه به نقطه هم از آن استفاده می کنند.

برای مدتهای مدید تصور بر آن بود که فریمهایی که مجموع چک آنها محاسبه می شود دارای طرح بیت تصادفی هستند، و تمام الگوریتمهای محاسبه مجموع چک نیز فرض را بر این می گذاشتند. اما بررسی دقیق داده های واقعی نشان داده که این فرض بکلی اشتباه است. در نتیجه، خطاهایی که (تحت شرایط خاص) کشف نشده می مانند شایعتر از آن چیزیست که قبلاً تصور می شد (Partridge et al., 1995).

۳-۳ چند پروتکل ساده لینک داده

برای آشنایی با پروتکل های لایه پیوند داده، در این قسمت سه پروتکل را (که بتدریج پیچیده تر می شوند) بررسی خواهیم کرد. برای خوانندگان علاقمند، شبیه ساز این پروتکلها (و پروتکلهایی که در آینده خواهید دید) را در سایت وب کتاب قرار داده ایم (<http://www.prenhall.com/tanenbaum>). اما قبل از اینکه سراغ این پروتکلها برویم، اجازه دهید چند تا از فرض هایی را که درباره مدل ارتباطی زیربنایی داشته ایم، توضیح دهیم. اول اینکه فرض کرده ایم در لایه های فیزیکی، لایه پیوند داده و شبکه پروسسهایی هستند که مستقل از یکدیگرند، و ارتباط آنها از طریق رد و بدل کردن پیام صورت می گیرد. در بسیاری از موارد، پروسسهای لایه فیزیکی و لینک داده در پردازنده کارت شبکه اجرا می شوند، و پروسسهای لایه شبکه در CPU اصلی کامپیوتر. البته پیاده سازیهای دیگری نیز محتمل است (مثلاً، تمام پروسسهای لایه فیزیکی، لینک داده و شبکه در پردازنده کارت شبکه اجرا شوند، یا همگی آنها را CPU اصلی اجرا کند). در هر حال، مستقل دانستن این پروسسهای بحث درباره آنها را بسیار ساده تر می کند، و تأکید است بر مستقل بودن لایه ها.

فرض کلیدی دیگر اینست که، ماشین A با استفاده از یک سرویس اتصال-گرای قابل اعتماد استریم طویلی از داده ها را به ماشین B می فرستد. بعدها این حالت که B هم همزمان به A داده بفرستد، را بررسی خواهیم کرد. علاوه بر آن، فرض کرده ایم ماشین A منبع بی پایانی از داده ها دارد که ارسال کند، و هرگز منتظر آمدن داده ها نخواهد شد. بعبارت دیگر، هر گاه لایه پیوند داده A درخواست داده کند، لایه شبکه بلافاصله اجابت می کند. (بعدها این فرض را هم کنار خواهیم گذاشت).

فرض دیگر ما اینست که این کامپیوترها هرگز از کار نمی افتند؛ یعنی، پروتکلهای ما فقط با خطاهای مخابراتی سروکار دارند، نه هنگ کردن کامپیوتر و مسائلی از این قبیل.

دیگر اینکه، تا آنجا که به لایه پیوند داده مربوط است، بسته ای که به لایه شبکه داده می شود، داده خالص است، و باید تا آخرین بیت به آن تحویل شود. این که بخشی از این داده ها سرآیند بسته هستند و لایه شبکه آنها را دور می ریزد، به خودش مربوط است نه به لایه پیوند داده.

وقتی لایه پیوند داده بسته ای از لایه شبکه می گیرد تا ارسال کند، آنرا فریم بندی کرده و سرآیند (header) و پی آیند (trailer) های لازم را به آن می چسباند (شکل ۳-۱ را ببینید). بنابراین هر فریم از سه بخش تشکیل می شود: قسمتی از بسته ای که لایه شبکه فرستاده، یک سرآیند شامل اطلاعات کنترل، و یک پی آیند شامل مجموع چک فریم. سپس این فریم به لایه پیوند داده ماشین مقصد فرستاده می شود. فرض ما بر این است که روال های کتابخانه ای *to_physical_layer* (برای ارسال) و *from_physical_layer* (برای دریافت) از قبل وجود دارند. مجموع چک نیز (مثلاً، با استفاده از کدهای چند جمله ای) توسط سخت افزار محاسبه (و به انتهای فریم اضافه) می شود، بنابراین لازم نیست لایه پیوند داده نگران آن باشد.

در ابتدا، لازم نیست گیرنده کاری انجام دهد؛ فقط منتظر می ماند تا اتفاقی بیفتد. در مثالهای این فصل، فرض کرده ایم که لایه پیوند داده این کار را با روالی بنام *wait_for_event(&event)* انجام می دهد. این روال فقط وقتی به پایان می رسد (و کنترل را به برنامه اصلی برمی گرداند) که اتفاقی افتاده باشد (یعنی، یک فریم دریافت شده باشد). اینکه چه اتفاقی افتاده است، را متغیر *event* مشخص می کند؛ و این که چه اتفاقهایی می تواند بیفتد، به تعریف پروتکل بستگی دارد. توجه داشته باشید که در دنیای واقعی لایه پیوند داده (مانند این مثالها) در یک حلقه بی انتها منتظر رسیدن فریمها نمی ماند، بلکه با استفاده از وقفه (*interrupt*) به آنها رسیدگی می کند. با این حال برای اجتناب از پیچیدگی مطلب، فرض کرده ایم که لایه پیوند داده هیچ کار دیگری جز رسیدگی به کانال ارتباطی ما ندارد.

وقتی یک فریم به گیرنده می رسد، سخت افزار مجموع چک آنرا محاسبه می کند. اگر این مجموع چک اشتباه باشد (یعنی خطایی رخ داده)، به لایه پیوند داده اطلاع داده می شود (*event = cksum_err*). اگر فریم بدرستی دریافت شده باشد، باز هم به اطلاع لایه پیوند داده می رسد (*event = frame_arrival*). در این حالت لایه پیوند داده با استفاده از تابع *from_physical_layer* فریم را گرفته، اطلاعات کنترلی موجود در سرآیند آنرا چک می کند، و اگر همه چیز مرتب باشد، سرآیند را جدا کرده و بخش اصلی داده را به لایه شبکه تحویل می دهد.

تحت هیچ شرایطی سرآیند فریم به لایه شبکه تحویل نمی شود، و برای این کار دلیل خوبی وجود دارد پروتکل های لینک داده و شبکه باید کاملاً از یکدیگر مستقل باشند. مستقل بودن پروتکل های این دو لایه باعث می شود که بتوان هر کدام از این پروتکلها را تغییر داد، بدون اینکه نیاز باشد پروتکل های لایه دیگر تغییر کند (البته تحت هر شرایطی نحوه تعامل و ارتباط لایه ها نباید تغییر کند). جدا و مستقل بودن لایه ها تا حد زیادی طراحی آنها را ساده می کند، چون می توان بدون نگرانی از اتفاقاتی که در لایه های دیگر می افتد، روی طراحی عملکردهای همان لایه تمرکز کرد.

در شکل ۳-۹ مقداری تعریف (به زبان C) می بینید، که در طراحی پروتکل های این قسمت به آنها نیاز داریم. در اینجا پنج ساختار داده تعریف شده است: *boolean*، *seq_nr*، *packet*، *frame_kind* و *frame*. ساختار *boolean* از نوع شمارشی (*enum*) است، و می تواند دو مقدار *true* و *false* بگیرد. ساختار *seq_nr* از نوع عدد صحیح بدون علامت (*unsigned int*) تعریف شده، و برای شماره گذاری فریمها از آن استفاده خواهیم کرد. شماره گذاری فریمها از 0 تا *MAX_SEQ* (که بسته به نیاز هر پروتکل تعریف می شود) انجام می گیرد. *packet* (بسته) واحدی از اطلاعات است که بین لایه شبکه و لایه پیوند داده (روی یک ماشین، یا روی ماشین های جداگانه) رد و بدل می شود. در مدل ما هر بسته همیشه حاوی *MAX_PKT* بایت داده است، ولی به واقعیت نزدیکتر است که طول بسته را متغیر در نظر بگیریم.

هر *frame* از چهار فیلد تشکیل شده: *seq*، *kind*، *ack* و *info* - که سه تای اول اطلاعات کنترلی هستند، و آخری همان داده هائیکه باید منتقل شود. به مجموعه فیلدهای کنترلی سرآیند فریم (*frame header*) گفته می شود.

```

#define MAX_PKT 1024  /* determines packet size in bytes */

typedef enum {false, true} boolean;  /* boolean type */
typedef unsigned int seq_nr;  /* sequence or ack numbers */
typedef struct {unsigned char data[MAX_PKT];} packet;  /* packet definition */
typedef enum {data, ack, nak} frame_kind;  /* frame_kind definition */

typedef struct {  /* frames are transported in this layer */
    frame_kind kind;  /* what kind of a frame is it? */
    seq_nr seq;  /* sequence number */
    seq_nr ack;  /* acknowledgement number */
    packet info;  /* the network layer packet */
} frame;

/* Wait for an event to happen; return its type in event. */
void wait_for_event(event_type *event);

/* Fetch a packet from the network layer for transmission on the channel. */
void from_network_layer(packet *p);

/* Deliver information from an inbound frame to the network layer. */
void to_network_layer(packet *p);

/* Go get an inbound frame from the physical layer and copy it to r. */
void from_physical_layer(frame *r);

/* Pass the frame to the physical layer for transmission. */
void to_physical_layer(frame *s);

/* Start the clock running and enable the timeout event. */
void start_timer(seq_nr k);

/* Stop the clock and disable the timeout event. */
void stop_timer(seq_nr k);

/* Start an auxiliary timer and enable the ack_timeout event. */
void start_ack_timer(void);

/* Stop the auxiliary timer and disable the ack_timeout event. */
void stop_ack_timer(void);

/* Allow the network layer to cause a network_layer_ready event. */

```

```
void enable_network_layer(void);

/* Forbid the network layer from causing a network_layer_ready event. */
void disable_network_layer(void);

/* Macro inc is expanded in-line: Increment k circularly. */
#define inc(k) if (k < MAX_SEQ) k = k + 1; else k = 0
```

شکل ۳-۹. تعریف های مورد نیاز برای پروتکل هایی که در این فصل می نویسیم. این تعریف ها در فایل بنام *protocol.h* قرار داده می شوند.

فیلد *kind* می گوید که آیا داده ای در فریم وجود دارد یا خیر، چون برخی از پروتکلها فریمهای بدون داده را از فریمهایی که داده دارند، تمیز می دهند. فیلدهای *seq* و *ack* بترتیب برای شماره ترتیبی فریم و تصدیق دریافت مورد استفاده قرار می گیرند؛ بعداً در این باره بیشتر توضیح خواهیم داد. داده اصلی (بسته) در فیلد *info* فریم قرار دارد؛ فریمهای کنترلی نیز وجود دارند که اساساً در آنها فیلد *info* وجود ندارد. در پروتکل های واقعی که طول فیلد *info* می تواند متغیر باشد، نیازی به تمایز بین فریمهای داده و فریمهای کنترلی نیست (چون فریم کنترلی فریمی است که طول فیلد *info* در آن 0 است).

در اینجا لازم است تفاوت فریم (*frame*) و بسته (*packet*) را مجدداً یادآور شویم. لایه شبکه با گرفتن پیام از لایه انتقال و اضافه کردن سرآیند، آنرا بصورت بسته در می آورد. سپس این بسته به لایه پیوند داده تحویل می شود، که در آنجا در فیلد *info* یک فریم قرار داده شده و برای ارسال آماده می شود. وقتی این فریم به لایه پیوند داده ماشین مقصد رسید، بسته از فیلد *info* استخراج شده و به لایه شبکه تحویل می شود. این فرآیند بکلی شفاف است، و لایه های شبکه در دو ماشین متقابل تصور می کنند که مستقیماً در حال تبادل بسته اند.

در برنامه شکل ۳-۹ چند روال (تابع) نیز تعریف شده است. اینها روالهای کتابخانه ای هستند، که فقط کاری که انجام می دهند برای ما مهم است (و اصلاً اهمیتی ندارد این کار را چگونه انجام می دهند). روال *wait_for_event* (همانطور که قبلاً گفتیم) در یک حلقه بی انتها به انتظار می ماند تا اتفاقی بیفتد. روالهای *to_network_layer* و *from_network_layer* بترتیب برای ارسال بسته به لایه شبکه و برای گرفتن بسته از این لایه بکار می روند (اینها واسطه لایه های ۲ و ۳ هستند). برای تبادل اطلاعات با لایه فیزیکی نیز از روالهای *to_physical_layer* و *from_physical_layer* استفاده می شود (اینها واسطه لایه های ۱ و ۲ هستند).

در پروتکل های واقعی فرض بر اینست که کانال ارتباطی نامطمئن است، و احتمال این هست که فریمها در راه از بین بروند. برای مقابله با چنین وضعیتی، لایه پیوند داده همزمان با ارسال هر فریم، باید یک تایمر را راه اندازی کند. اگر بعد از مدتی معین پاسخی از طرف مقابل نرسید، تایمر مزبور لایه پیوند داده را (با استفاده از یک وقفه) مطلع می کند.

در پروتکل های ما، در چنین وضعیتی روال *wait_for_event* مقدار *event = timeout* برمی گرداند. روالهای *start_timer* و *stop_timer* نیز بترتیب تایمر را روشن و خاموش می کنند (البته سپری شدن زمان منقضی - *timeout* - فقط زمانی اتفاق می افتد که تایمر روشن باشد). یک تایمر را می توان (قبل از منقضی شدن آن) با اجرای مجدد روال *start_timer* ریست کرد.

روالهای *start_ack_timer* و *stop_ack_timer* تایمر دیگری را (برای ایجاد فریمهای تصدیق دریافت در شرایطی خاص) کنترل می کنند.

از روالهای `enable_network_layer` و `disable_network_layer` در پروتکل‌های پیچیده‌تر استفاده می‌شود (ما در پروتکل‌های ساده این قسمت از این روالها استفاده نخواهیم کرد، چون فرض کرده‌ایم که لایه شبکه همیشه می‌تواند به لایه پیوند داده بسته تحویل دهد). وقتی لایه پیوند داده لایه شبکه را فعال می‌کند (`enable_network_layer`)، لایه شبکه اجازه دارد آماده شدن بسته داده را با یک وقفه به لایه پیوند داده اطلاع دهد (ایسن کار با `event = network_layer_ready` انجام می‌شود). اگر لایه شبکه غیرفعال باشد (`disable_network_layer`)، اجازه چنین کاری را ندارد. با استفاده دقیق و بجای از روالهای `enable_network_layer` و `disable_network_layer`، لایه پیوند داده می‌تواند مطمئن شود که (هنگام پر شدن بافر) در سیلاب بسته‌های ارسالی از لایه شبکه غرق نخواهد شد.

شماره ترتیبی فریم همیشه بین 0 تا `MAX_SEQ` (و از جمله خود این دو عدد) است، که البته `MAX_SEQ` در پروتکل‌های مختلف می‌تواند متفاوت باشد. شماره ترتیبی فریمها معمولاً یکی یکی اضافه می‌شود، و وقتی به `MAX_SEQ` رسید، دوباره 0 خواهد شد - این کار بر عهده ماکرو `inc` گذاشته شده است. برای سرعت بخشیدن به اجرای این عملیات، `inc` بصورت ماکرو (`macro`) تعریف شده است. [کامپایلر با دیدن یک ماکرو، دستور معادل را جایگزین آن می‌کند، و مانند تابع آنرا فراخوانی نمی‌کند. م.] همانطور که بعداً خواهیم دید، سرعت اجرای پروتکلها یکی از عوامل کلیدی در کارایی شبکه است، و استفاده از ماکرو (بجای تابع) تأثیر زیادی بر بهبود این کارایی دارد. همچنین، از آنجائیکه `MAX_SEQ` در پروتکل‌های مختلف مقادیر متفاوتی دارد، تعریف `inc` بصورت ماکرو امکان می‌دهد تا بدون هیچ مشکلی از آن در پروتکل‌های مختلف استفاده کنیم.

تعاریف شکل ۳-۹ بخشی از پروتکل‌هاییست که در قسمتهای آینده خواهیم نوشت. البته می‌توانستیم آنها را در ابتدای هر پروتکل نیز بیاوریم، ولی با جمع کردن آنها در یک فایل کدهای آینده بسیار ساده‌تر خواهند شد. در زبان C، این کار با استفاده از دستور `#include` و نوشتن نام فایل تعاریف (در اینجا `protocol.h`) انجام می‌شود.

۱-۳-۳ پروتکل یکطرفه نامقید

در اولین مثال یک پروتکل بسیار ساده را در نظر می‌گیریم، که در آن داده‌ها فقط در یک جهت منتقل می‌شوند. لایه شبکه در فرستنده و گیرنده آماده کار هستند، زمان پردازش را می‌توان نادیده گرفت، از نظر بافر هیچ کمبودی وجود ندارد، و مهمتر از همه اینکه کانال ارتباطی بین دو لایه پیوند داده کامل و بدون نقص است، و هیچ خطایی در آن رخ نمی‌دهد. این پروتکل غیرواقعی (که شاید «اتوپیا» - پروتکل آرمانی - مناسبترین نام برای آن باشد) را در شکل ۳-۱۰ ملاحظه می‌کنید.

```
/* Protocol 1 (utopia) provides for data transmission in one direction only, from
sender to receiver. The communication channel is assumed to be error free
and the receiver is assumed to be able to process all the input infinitely quickly.
Consequently, the sender just sits in a loop pumping data out onto the line as
fast as it can. */
```

```
typedef enum {frame_arrival} event_type;
#include "protocol.h"

void sender1(void)
{
```



```

frame s; /* buffer for an outbound frame */
packet buffer; /* buffer for an outbound packet */

while (true) {
    from_network_layer(&buffer); /* go get something to send */
    s.info = buffer; /* copy it into s for transmission */
    to_physical_layer(&s); /* send it on its way */
} /* Tomorrow, and tomorrow, and tomorrow,
    Creeps in this petty pace from day to day
    To the last syllable of recorded time.
    - Macbeth, V, v */
}

void receiver1(void)
{
    frame r;
    event_type event; /* filled in by wait, but not used here */

    while (true) {
        wait_for_event(&event); /* only possibility is frame_arrival */
        from_physical_layer(&r); /* go get the inbound frame */
        to_network_layer(&r.info); /* pass the data to the network layer */
    }
}

```

شکل ۳-۱۰. پروتکل یکطرفه نامفید.

این پروتکل دارای دو روال مجزا است: فرستنده و گیرنده. فرستنده در لایه پیوند داده ماشین مبدأ، و گیرنده در لایه پیوند داده ماشین مقصد اجرا می شود. در اینجا از شماره ترتیبی فریمها و تصدیق دریافت استفاده ای نمی شود، بنابراین به MAX_SEQ هم نیازی نیست. تنها رویداد قابل انتظار $frame_arrival$ (رسیدن صحیح و سالم فریم) است.

فرستنده یک حلقه بی انتهای $while$ است که داده ها را با حداکثر توان به بیرون پمپ می کند. بدنه این حلقه سه کار انجام می دهد: آوردن یک بسته از لایه شبکه (که همیشه آماده خدمت است)، ایجاد فریم خروجی با استفاده از متغیر s ، و فرستادن فریم. این پروتکل فقط از فیلد $info$ فریم استفاده می کند، چون فیلدهای دیگر مربوط به کنترل جریان و خطا هستند، که طبق فرض ما چنین محدودیتهایی در اینجا وجود ندارد.

گیرنده هم بهمان اندازه ساده است: انتظار بی پایان برای دریافت فریمی که همیشه سالم و بی نقص است. وقتی یک فریم از راه رسید، روال $wait_for_event$ کنترل را به برنامه اصلی برمی گرداند، و متغیر $event$ را به $frame_arrival$ ست می کند (که بهر حال استفاده ای از آن نمی شود). با فراخوانی روال $from_physical_layer$ ، فریم تازه از راه رسیده از بافر سخت افزاری برداشته شده و در متغیر r قرار داده می شود (تا گیرنده می تواند آنرا بردارد). در پایان، بخش داده این فریم (فیلد $info$) به لایه شبکه فرستاده شده، و لایه پیوند داده به انتظار فریم بعدی می نشیند.

۲-۳-۳ پروتکل توقف-انتظار یکطرفه

حال اجازه دهید غیرواقعی ترین بخش از پروتکل ۱ را کنار بگذاریم: نامحدود بودن توانایی این پروتکل در دریافت بسته‌ها از لایه شبکه، و پردازش فریمهای ورودی (که به معنای نامحدود بودن بافر لایه پیوند داده در سمت گیرنده است). اما کانال ارتباطی را همچنان بدون خطا، و ارتباط را یکطرفه فرض کرده ایم.

مهمترین مشکلی که با آن روبرو هستیم، این است که چگونه از غرق شدن گیرنده در سیلاب فریمهایی (که پردازش آنها از توان وی خارج است) جلوگیری کنیم. واضحست که، اگر گیرنده برای اجرای روالهای *from_physical_layer* و *to_network_layer* به زمان Δt نیاز داشته باشد، سرعت متوسط ارسال فرستنده بایستی از یک فریم بر Δt کمتر باشد. همچنین اگر فرض کنیم که سخت‌افزار گیرنده بطور خودکار عمل بافر کردن فریمها را انجام نمی‌دهد، فرستنده نباید قبل از برداشته شدن یک فریم از بافر لایه فیزیکی (که توسط روال *from_physical_layer* انجام می‌شود)، فریم بعدی را ارسال کند چون در غیر اینصورت فریم قبلی از بین می‌رود (به این حالت روهم‌نویسی - overrun - می‌گویند).

در شرایط خاصی (مانند ارتباط سنکرون، و گیرنده‌ای که تنها وظیفه آن گرفتن اطلاعات از خط ورودی است)، با ایجاد تأخیر در قسمت فرستنده پروتکل ۱ و کند کردن آن می‌توان به اهداف فوق دست یافت. اما بسیار محتملتر است که یک لایه پیوند داده مجبور باشد چندین خط ورودی را پردازش کند، که در این حالت فاصله زمانی دریافت فریمها و پردازش آنها می‌تواند بسیار متغیر باشد. اگر طراحان شبکه بتوانند بدترین حالت گیرنده را محاسبه کنند، می‌توانند فرستنده را آنقدر کند کنند که روهم‌نویسی هرگز اتفاق نیفتد. اما این روش بسیار محافظه کارانه است و بنحو بسیار بدی پهنای باند را تلف می‌کند، مگر اینکه تفاوت بهترین و بدترین حالت چندان زیاد نباشد (یعنی تفاوت پاسخهای لایه پیوند داده ناچیز باشد).

راه حل بهتر این معضل، برگرداندن بازخور (feedback) از گیرنده به فرستنده است. بعد از تحویل بسته به لایه شبکه، گیرنده یک فریم کوچک (که لازم نیست معنی خاصی هم داشته باشد) به فرستنده می‌فرستد، که در واقع مجوز ارسال فریم بعدی محسوب می‌شود. فرستنده بعد از ارسال یک فریم، آنقدر منتظر می‌ماند تا این فریم کوچک (که در واقع همان تصدیق دریافت - acknowledgement - است) از راه برسد. استفاده از بازخور گیرنده برای اطلاع به فرستنده (و دادن مجوز ارسال فریمهای بعدی) یکی از نمونه‌های کنترل جریان (flow control)، که قبلاً به آن اشاره کردیم، است.

پروتکلهایی که در آنها فرستنده قبل از ارسال فریم بعدی منتظر تصدیق دریافت فریم قبلی از گیرنده می‌ماند، به پروتکلهای توقف-انتظار (stop-and-wait) معروفند. در شکل ۳-۱۱ یک نمونه از پروتکلهای توقف-انتظار را ملاحظه می‌کنید.

/* Protocol 2 (stop-and-wait) also provides for a one-directional flow of data from sender to receiver. The communication channel is once again assumed to be error free, as in protocol 1. However, this time, the receiver has only a finite buffer capacity and a finite processing speed, so the protocol must explicitly prevent the sender from flooding the receiver with data faster than it can be handled. */

```
typedef enum {frame_arrival} event_type;
"*.locotorp" edulcni#
```

```

void sender2(void)
{
    frame s;                /* buffer for an outbound frame */
    packet buffer;          /* buffer for an outbound packet */
    event_type event;       /* frame_arrival is the only possibility */

    while (true) {
        from_network_layer(&buffer); /* go get something to send */
        s.info = buffer;             /* copy it into s for transmission */
        to_physical_layer(&s);      /* bye-bye little frame */
        wait_for_event(&event);     /* do not proceed until given the go ahead */
    }
}

void receiver2(void)
{
    frame r, s;             /* buffers for frames */
    event_type event;       /* frame_arrival is the only possibility */
    while (true) {
        wait_for_event(&event); /* only possibility is frame_arrival */
        from_physical_layer(&r); /* go get the inbound frame */
        to_network_layer(&r.info); /* pass the data to the network layer */
        to_physical_layer(&s);    /* send a dummy frame to awaken sender */
    }
}

```

شکل ۳-۱۱. پروتکل توقف-انتظار یکطرفه.

با اینکه این پروتکل یکطرفه (simplex) است (یعنی ما فقط در یک جهت ارسال می‌کنیم)، اما فریمها می‌توانند در هر دو جهت رفت و آمد کنند. برای این منظور لازم است کانال ارتباطی ما چنین قابلیت داشته باشد؛ البته یک کانال دوطرفه ناهمزمان (half-duplex) هم کفایت می‌کند، چون فرستنده و گیرنده در آن واحد اقدام به فرستادن فریمها نمی‌کنند: فرستنده یک فریم می‌فرستد، گیرنده پاسخ می‌دهد، فرستنده فریم بعدی را می‌فرستد، گیرنده پاسخ می‌دهد، و الی آخر.

در اینجا هم (مانند پروتکل ۱) فرستنده همان سه کار قبلی را انجام می‌دهد: آوردن یک بسته از لایه شبکه، ایجاد فریم خروجی، و فرستادن آن. اما برخلاف پروتکل ۱، قبل از ادامه کار (آوردن بسته بعدی و ارسال آن در قالب یک فریم) باید منتظر رسیدن فریم تصدیق دریافت از گیرنده بماند. نیازی نیست که لایه پیوند داده فرستنده فریم دریافتی را بررسی کند، چون فقط یک احتمال وجود دارد: این فریم همیشه تصدیق دریافت گیرنده است. تنها تفاوت *receiver2* با *receiver1* این است که بعد از تحویل بسته به لایه شبکه و قبل از ورود به حالت انتظار، *receiver2* یک فریم تصدیق دریافت به فرستنده باز پس می‌فرستد. از آنجائیکه فقط خود فریم مهم است و نه محتویات آن، گیرنده هیچ داده‌ای در فیلد *info* این فریم قرار نمی‌دهد.

۳-۳-۳ پروتکل یکطرفه برای کانالهای نوپزدار

اکنون به یک حالت واقعی تر می پردازیم: کانالهایی که نویز دارند. فریمها می توانند با خطا به مقصد برسند، و یا بکلی گم شده و اصلاً به مقصد نرسند. با این حال، فرض می کنیم که اگر فریمی با خطا به مقصد رسید، سخت افزار لایه فیزیکی با محاسبه جمع تطبیقی متوجه خطا می شود. پروتکل ما در یک حالت به اشتباه عمل خواهد کرد: خطای رخ داده آنقدر شدید باشد که جمع تطبیقی تصادفاً درست از کار در آید (اتفاقی که بسیار نامحتمل است). در نگاه اول با یک تغییر کوچک در پروتکل ۲ (اضافه کردن یک تایمر) می توان آن را با وضعیت جدید تطبیق داد. فرستنده می تواند در هر زمانی یک فریم بفرستد، ولی گیرنده فقط وقتی فریم تصدیق دریافت را برمی گرداند که این فریم را بدرستی دریافت و پردازش کرده باشد. اگر فریم ناقص به مقصد برسد، دور انداخته خواهد شد. بعد از مدتی تایمر فرستنده به انتها می رسد، و چون هنوز تصدیق دریافت گیرنده را نگرفته، مجدداً اقدام به ارسال فریم می کند. این ماجرا تا زمانی که فریم به سلامت به مقصد برسد، تکرار خواهد شد. اما طرح بالا یک مشکل اساسی دارد. قبل از خواندن ادامه کتاب، کمی فکر کنید و ببینید می توانید متوجه اشکال آن شوید.

برای درک مشکل، بیاد بیاورید که وظیفه برقراری یک کانال ارتباطی عاری از خطا بین لایه های شبکه بر عهده لایه پیوند داده است. لایه شبکه ماشین A یک سری بسته به لایه پیوند داده می دهد، و باید مطمئن باشد که این بسته ها به همان ترتیبی که ارسال شده اند بدست لایه شبکه ماشین B خواهند رسید. بویژه، لایه شبکه ماشین B هیچ راهی ندارد تا بفهمد که یک بسته گم شده یا تکراریست. به همین دلیل لایه پیوند داده ماشین B باید تضمین کند که هیچ بسته ای گم نمی شود، و یا تکراری نیست. سناریوی زیر را در نظر بگیرید:

۱. لایه شبکه ماشین A بسته ۱ را به لایه پیوند داده می دهد. این بسته صحیح و سالم به لایه پیوند داده ماشین B می رسد، و تحویل لایه شبکه می شود. ماشین B یک فریم تصدیق دریافت به A می فرستد.
۲. فریم تصدیق دریافت ماشین B در راه از بین می رود، و هرگز به A نمی رسد. اگر فقط فریمهای داده گم می شدند و این اتفاق برای فریمهای کنترلی نمی افتاد، زندگی چقدر شیرین تر بود! ولی متأسفانه کانالهای مخابراتی اهل تبعیض نیستند!
۳. تایمر لایه پیوند داده A به انتها می رسد، و چون هیچ فریم تصدیق دریافتی بدستش نرسیده، (باشتابه) تصور می کند که فریم به مقصد نرسیده (یا خراب شده)، پس آنرا دوباره می فرستد.
۴. فریم تکراری (در کمال صحت و سلامت) به لایه پیوند داده B می رسد، و این لایه هم (بی خبر از همه جا) آنرا به لایه شبکه تحویل می دهد. تصور کنید که اگر A در حال ارسال یک فایل به B باشد، تکراری بودن بخشی از آن چه فاجعه ای پدید خواهد آورد. همانطور که می بینید، پروتکل ما یک شکست کامل است.

چیزی که ما به آن احتیاج داریم، وسیله ایست که بتوان فریمهای تکراری را از فریمهایی که برای اولین بار دریافت می شوند، تشخیص داد. راه حل واضح این مشکل آن است که فرستنده یک شماره ترتیبی (sequence number) در سرآیند فریمهایی که می فرستد، قرار دهد. گیرنده می تواند این شماره را چک کرده، و فریمهای تکراری را دور بیندازد.

از آنجائیکه سرآیند یک فریم باید حتی الامکان کوچک باشد، سؤالی که پیش می آید اینست که: حداقل تعداد بیتهای لازم برای فیلد شماره ترتیبی چندناست؟ در پروتکل ما تنها ابهام در فریم m و فریم بعدی آن یعنی $m + 1$ است. اگر فریم m خراب شود یا از بین برود، گیرنده فریم تصدیق دریافت آنرا بر نمی گرداند، پس فرستنده سعی می کند آنرا دوباره بفرستد. همین که این فریم سالم به مقصد رسید، گیرنده فریم تصدیق دریافت را به فرستنده پس می فرستد. همین جاست که مشکل بروز می کند: اگر فریم تصدیق دریافت صحیح و سالم به فرستنده بفرستد،

فرستنده فریم بعدی (یعنی $m + 1$) را می‌فرستد، در غیر اینصورت فریم m را خواهد فرستاد. برای ارسال فریم $m + 2$ ، فرستنده باید قبلاً تصدیق دریافت فریم $m + 1$ را گرفته باشد. اما این بدان معناست که فریم m به سلامت به مقصد رسیده و تصدیق دریافت آن هم بدرستی به فرستنده برگشت داده شده است (چون در غیر اینصورت فرستنده فریم $m + 1$ را هم نمی‌فرستاد، چه رسد به فریم $m + 2$). بنابراین، تنها ابهامی که می‌تواند وجود داشته باشد، بین یک فریم و فریم بعدی آن است. برای تشخیص این دو هم یک شماره ترتیبی یک‌بیتی (0 یا 1) کافیست. عبارت دیگر، در هر لحظه گیرنده باید بدنبال شماره بعدی باشد. اگر فریمی با شماره اشتباه دریافت شد، گیرنده آنرا تکراری تلقی کرده و دور می‌اندازد. اما اگر شماره ترتیبی فریم درست بود، به لایه شبکه تحویل داده می‌شود. با این توصیف فیلد شماره ترتیبی باید در جدول 2 افزایش داده شود (عبارت دیگر، 1 به 0 تبدیل می‌شود، و 0 به 1).

پروتکلی با این مشخصات را در شکل ۳-۱۲ ملاحظه می‌کنید. به پروتکل‌هایی که فرستنده برای ارسال فریم بایستی منتظر یک تصدیق دریافت مثبت بماند، PAR (تصدیق دریافت مثبت با ارسال مجدد - Positive Acknowledgement with Retransmission) یا ARQ (درخواست تکرار خودکار - Automatic Repeat reQuest) نیز گفته می‌شود. این پروتکل هم، مانند پروتکل ۲، فقط در یک جهت داده می‌فرستد. تفاوت پروتکل ۳ با دو تای قبلی اینست که، روالهای فرستنده و گیرنده متغیری دارند که مقدار آن حتی در زمانی که لایه پیوند داده به حالت انتظار می‌رود، دست نخورده باقی می‌ماند. فرستنده باید شماره ترتیبی فریم بعدی که می‌خواهد بفرستد، را بداند: `next_frame_to_send`؛ و گیرنده هم باید شماره ترتیبی فریم بعدی که باید منتظر آن باشد، را بداند: `frame_expected`.

```

/* Protocol 3 (par) allows unidirectional data flow over an unreliable channel. */
#define MAX_SEQ 1 /* must be 1 for protocol 3 */
typedef enum {frame_arrival, cksum_err, timeout} event_type;
#include "protocol.h"
void sender3(void)
{
    seq_nr next_frame_to_send; /* seq number of next outgoing frame
*/
    frame s; /* scratch variable */
    packet buffer; /* buffer for an outbound packet */
    event_type event;
    next_frame_to_send = 0; /* initialize outbound sequence
numbers */
    from_network_layer(&buffer); /* fetch first packet */
    while (true) {
        s.info = buffer; /* construct a frame for transmission
*/
        s.seq = next_frame_to_send; /* insert sequence number in frame
*/
        to_physical_layer(&s); /* send it on its way */
        start_timer(s.seq); /* if answer takes too long, time out

```

```

*/
    wait_for_event(&event);          /* frame_arrival, cksum_err, timeout
*/
    if (event == frame_arrival) {
        from_physical_layer(&s);      /* get the acknowledgement */
        if (s.ack == next_frame_to_send) {
            stop_timer(s.ack);        /* turn the timer off */
            from_network_layer(&buffer); /* get the next one to send */
            inc(next_frame_to_send);  /* invert next_frame_to_send */
        }
    }
}
}

void receiver3(void)
{
    seq_nr frame_expected;
    frame r, s;
    event_type event;
    frame_expected = 0;
    while (true) {
        wait_for_event(&event);      /* possibilities: frame_arrival,
cksum_err */
        if (event == frame_arrival) { /* a valid frame has arrived. */
            from_physical_layer(&r);  /* go get the newly arrived frame
*/
            if (r.seq == frame_expected) { /* this is what we have been
waiting for. */
                to_network_layer(&r.info); /* pass the data to the network
layer */
                inc(frame_expected);     /* next time expect the other
sequence nr */
            }
            s.ack = 1 - frame_expected; /* tell which frame is being acked
*/
            to_physical_layer(&s);     /* send acknowledgement */
        }
    }
}
}
}

```

شکل ۳-۱۲. پروتکل تصدیق دریافت مثبت با ارسال مجدد (PAR).

فرستنده، بعد از ارسال فریم، تایمر را راه می‌اندازد. اگر تایمر از قبل در حال اجرا باشد، این کار آنرا ریست کرده و آماده کار بعدی می‌کند. فاصله زمانی تایمر باید بگونه‌ای انتخاب شود که وقت کافی برای سه رویداد بدست دهد: رسیدن فریم به گیرنده، پردازش آن در گیرنده (در بدترین حالت)، و برگشت فریم تصدیق دریافت به فرستنده. فقط پس از سپری شدن این زمان است که می‌توان مطمئن شد فریم (یا تصدیق دریافت آن) به مقصد نرسیده، و فرستنده باید دوباره آنرا بفرستد. اگر فاصله زمانی تایمر کم انتخاب شود، تعداد دفعاتی که فرستنده فریم تکراری می‌فرستد افزایش یافته، و (با اینکه این کار تأثیر منفی روی گیرنده ندارد) به کارایی سیستم لطمه می‌زند. بعد از ارسال فریم و راه انداختن تایمر، فرستنده منتظر یک اتفاق هیجان‌انگیز می‌ماند. البته فقط سه احتمال برای چنین اتفاقی وجود دارد: فریم تصدیق دریافت صحیح و سالم از راه برسد، فریم تصدیق دریافت با خطا وارد شود، تایمر منقضی شود (زمان مشخص شده به انتها برسد). در حالت اول، فرستنده بسته دیگری از لایه شبکه گرفته، و در بافر خود (buffer) قرار می‌دهد. سپس، شماره ترتیبی فریم (next_frame_to_send) را بالا می‌برد. اما در دو حالت دیگر (خراب شدن فریم تصدیق دریافت یا نرسیدن آن)، بافر و شماره ترتیبی هیچکدام تغییری نمی‌کند، بنابراین در هیچ حالتی فریم تکراری فرستاده نخواهد شد. وقتی یک فریم به گیرنده می‌رسد، گیرنده شماره ترتیبی آن را چک می‌کند، تا از تکراری نبودن آن مطمئن شود. این فریم فقط در صورت تکراری نبودن به لایه شبکه تحویل داده می‌شود. بدین ترتیب، فریمهای تکراری و خراب به لایه شبکه نخواهند رسید.

۴.۳ پروتکل‌های پنجره لغزنده

در پروتکل‌های قبل، فریمهای داده فقط در یک جهت ارسال می‌شدند. اما در عمل باید بتوانیم در هر دو جهت انتقال داده داشته باشیم. یکی از راههای داشتن یک کانال دوطرفه همزمان (full-duplex) استفاده از دو کانال یکطرفه (simplex) در دو جهت مخالف است، که هر کدام فقط در یک جهت داده می‌فرستند (و البته فریمهای تصدیق دریافت می‌کنند). اما این روش چیزی جز اتلاف پهنای باند (و پول) نیست. روش بهتر استفاده از یک کانال واحد برای ارسال داده در دو جهت است (مگر نه اینکه در پروتکل‌های ۲ و ۳ در هر دو جهت فریم ارسال کردیم). از آنجائیکه در این مدل، فریمهای داده و تصدیق دریافت در هر دو جهت می‌توانند فرستاده شوند، باید کاری کنیم که گیرنده بتواند آنها را از یکدیگر تشخیص دهد. برای این منظور می‌توانیم از فیلد *info* در سرآیند فریمها استفاده کنیم. با اینکه ترکیب فریمهای داده و تصدیق دریافت روی یک مدار واحد (بجای دو مدار جداگانه) یک قدم به جلو محسوب می‌شود، اما باز هم می‌توان کارایی سیستم را بهبود بخشید. وقتی گیرنده یک فریم داده دریافت می‌کند، بجای اینکه بلافاصله یک فرم کنترلی پس بفرستد، منتظر می‌ماند تا بسته بعدی را برای ارسال از لایه شبکه بگیرد. تصدیق دریافت فریم قبلی در فیلد *ack* فریم داده‌ای که اکنون می‌خواهد فرستاده شود، قرار داده می‌شود، و در واقع فریم تصدیق دریافت از فریم داده سواری مجانی (piggyback) می‌گیرد (و به همین نام هم خوانده می‌شود). یکی از مزایای تکنیک سواری مجانی نسبت به ارسال مستقل فریمهای تصدیق دریافت، استفاده بهینه‌تر از پهنای باند موجود است: فیلد *ack* فقط چند بیت از سرآیند را اشغال می‌کند، در حالیکه یک فریم مستقل برای خود سرآیند، جمع تطبیقی، و تصدیق دریافت دارد. علاوه بر آن، هر چه تعداد فریمهایی که در یک جهت فرستاده می‌شوند کمتر باشد، گیرنده بهتر می‌تواند به کارهای دیگرش (پردازش فریمهای رسیده و خالی کرن بافرها) برسد، و این هم به بهبود کارایی سیستم کمک می‌کند. در پروتکلی که در این قسمت می‌نویسیم، فیلد سواری مجانی فقط یک بیت به سرآیند فریم اضافه می‌کند (و بندرت پیش می‌آید که مقدار آن از چند بیت بیشتر شود).

اما سواری مجانی هم خالی از اشکال نیست. برای مثال، در این حالت لایه پیوند داده گیرنده چقدر باید منتظر بسته از لایه شبکه خود شود؟ اگر خیلی منتظر بماند، تایمر فرستنده به انتها رسیده و فریم را تکرار می‌کند، که این نقض غرض (از ارسال فریم تصدیق دریافت) است. اگر لایه پیوند داده قدرت پیشگویی داشت، می‌توانست زمان دریافت بسته بعدی از لایه شبکه را پیشگویی کند، و آنوقت می‌توانست تصمیم بگیرد که منتظر این بسته بماند یا بلافاصله فریم تصدیق دریافت را به فرستنده بفرستد. اما متأسفانه لایه پیوند داده نمی‌تواند آینده را پیشگویی کند، پس باید روش ساده‌تری (انتظار بمدت ثابت، مثلاً چند میلی‌ثانیه) پیدا کنیم. اگر در این فاصله بسته‌ای از لایه شبکه رسید، لایه پیوند داده تصدیق دریافت را سوار آن می‌کند؛ در غیر اینصورت یک فریم مستقل تصدیق دریافت به سمت فرستنده ارسال می‌کند.

پروتکل‌هایی که در این قسمت خواهید دید، به کلاس پروتکل‌های پنجره لغزنده (sliding window) تعلق دارند، که فقط از نظر کارایی، پیچیدگی و بافر با هم تفاوت دارند. در این پروتکلها، مانند سایر پروتکل‌های پنجره لغزنده، هر فریم خروجی یک شماره ترتیبی (از 0 تا یک حداکثر) دارد. حداکثر شماره فریمها معمولاً $2^n - 1$ است، بنابراین در یک فیلد n -بیتی بخوبی جا می‌شود. پروتکل پنجره لغزنده توقف-انتظار از $n = 1$ استفاده می‌کرد، اما در پروتکل‌های دیگر این عدد می‌تواند بیشتر باشد.

ایده اصلی در تمام پروتکل‌های پنجره لغزنده این است که، فرستنده در هر لحظه از زمان لیستی از شماره‌های ترتیبی متناظر با فریم‌هایی که می‌تواند ارسال کند، در اختیار دارد. اصطلاحاً گفته می‌شود که این فریمها در پنجره ارسال (sending window) قرار دارند. گیرنده هم یک پنجره دریافت (receiving window) دارد که متناظر است با فریم‌هایی که مجاز به دریافت آنهاست. الزامی نیست که پنجره ارسال و پنجره دریافت حد پائین و بالای مشابه داشته باشند، یا حتی هم‌اندازه باشند. در برخی پروتکلها اندازه این پنجره‌ها ثابت است، ولی در پروتکل‌های دیگر می‌تواند کوچک یا بزرگ شود.

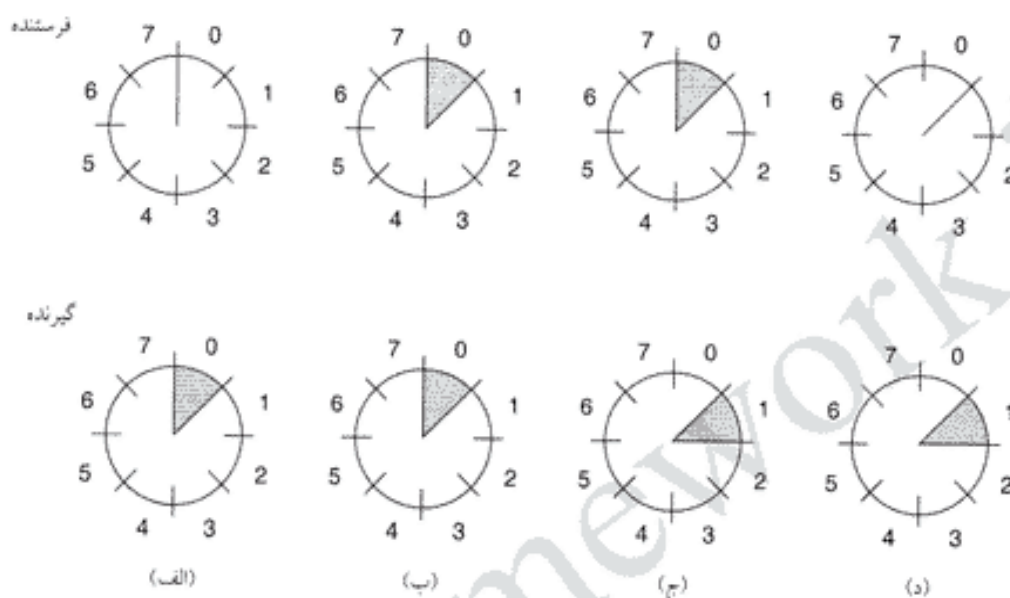
با اینکه این پروتکلها آزادی عمل بیشتری به لایه پیوند داده در ارسال و دریافت فریمها می‌دهند، لازم است مجدداً تأکید کنیم که تحویل بسته‌ها به لایه شبکه مقصد باید با همان ترتیبی صورت گیرد که در ماشین مبدأ تحویل لایه پیوند داده شده‌اند. (و یک بار دیگر خاطر نشان می‌کنیم که، لایه فیزیکی یک کانال ارتباطی ساده است که فریمها را به همان ترتیبی که به آن داده شده، منتقل می‌کند.)

شماره‌های موجود در پنجره ارسال شماره فریم‌هایست که باید ارسال شوند، و یا ارسال شده‌اند ولی هنوز تصدیق دریافت آنها برنگشته است. وقتی یک بسته جدید از لایه شبکه می‌رسد، لایه پیوند داده بالاترین شماره ترتیبی موجود را به آن می‌دهد، و لبه بالایی پنجره ارسال را یکی زیاد می‌کند. وقتی یک تصدیق دریافت وارد شد، لبه پائینی پنجره ارسال نیز بالا برده می‌شود. بدین ترتیب پنجره ارسال همیشه شامل لیست فریم‌هایست که دریافت آنها هنوز توسط گیرنده تصدیق نشده است. در شکل ۳-۱۳ یک نمونه را ملاحظه می‌کنید.

از آنجائیکه امکان گم یا خراب شدن فریم‌هایی که در حال حاضر در پنجره ارسال قرار دارند، همیشه وجود دارد، فرستنده باید آنها را برای ارسال مجدد (احتمالی) در حافظه نگه دارد. بنابراین اگر حداکثر اندازه این پنجره n باشد، فرستنده باید بافری باندازه n فریم تصدیق نشده داشته باشد. اگر پنجره ارسال از حداکثر پیش‌بینی شده بزرگتر شود، لایه پیوند داده باید گرفتن بسته از لایه شبکه را تا زمان آزاد شدن بافر متوقف کند.

پنجره دریافت در گیرنده متناظر است با فریم‌هایی که گیرنده مجاز به دریافت آنهاست. هر فریمی که خارج از این پنجره قرار گیرد، بدون هیچ توضیحی دور انداخته خواهد شد. وقتی فریمی که شماره ترتیبی آن معادل لبه پائین پنجره دریافت است، از راه می‌رسد، به لایه شبکه تحویل شده و پس از ایجاد تصدیق دریافت آن، پنجره دریافت یک واحد می‌چرخد. بر خلاف پنجره فرستنده، پنجره گیرنده همیشه به همان اندازه اولیه می‌ماند. توجه

کنید که پنجره دریافت ۱ بمعنای اینست که لایه پیوند داده فریمها را فقط به ترتیب می‌پذیرد، ولی در پنجره‌های بزرگتر الزاماً چنین نیست (با این حال، لایه شبکه همیشه داده‌ها را به ترتیب صحیح تحویل لایه پیوند داده می‌دهد).



شکل ۳-۱۳. یک پنجره لغزنده یک واحدی، با شماره ترتیبی ۳-بیتی. (الف) در شروع کار. (ب) بعد از ارسال اولین فریم. (ج) بعد از آنکه اولین فریم دریافت شد. (د) بعد از آنکه فرستنده اولین تصدیق دریافت را گرفت.

در شکل ۳-۱۳ حداکثر اندازه پنجره ۱ است. در لحظه اول لبه‌های پائین و بالای پنجره ارسال یکی هستند، ولی با گذشت زمان موقعیت آنها طبق شکل تغییر می‌کند.

۱-۴-۳ پروتکل پنجره لغزنده ۱-بیتی

قبل از پرداختن به حالت کلی، اجازه دهید ابتدا پروتکلی با پنجره لغزنده ۱-بیتی را مورد بررسی قرار دهیم. در واقع این پروتکل نوعی پروتکل توقف-انتظار است، چون فرستنده قبل از گرفتن تصدیق دریافت فریم فرستاده شده، فریم بعدی را ارسال نخواهد کرد.

در شکل ۳-۱۴ پروتکل پنجره لغزنده ۱-بیتی را ملاحظه می‌کنید. مانند پروتکل‌های قبلی، این پروتکل هم با تعریف متغیرها شروع می‌شود. متغیر *next_frame_to_send* فریم بعدیست که فرستنده باید بفرستد. در طرف مقابل هم، متغیر *frame_expected* فریمی را نشان می‌دهد که گیرنده منتظر آن است. در هر دو طرف، تنها حالت‌های مجاز فقط ۰ یا ۱ است.

```
/* Protocol 4 (sliding window) is bidirectional. */
#define MAX_SEQ 1 /* must be 1 for protocol 4 */
typedef enum {frame_arrival, cksum_err, timeout} event_type;
#include "protocol.h"
void protocol4 (void)
```

```

{
    seq_nr next_frame_to_send;           /* 0 or 1 only */
    seq_nr frame_expected;              /* 0 or 1 only */
    frame r, s;                         /* scratch variables */
    packet buffer;                       /* current packet being sent */
    event_type event;
    next_frame_to_send = 0;              /* next frame on the outbound
stream */
    frame_expected = 0;                  /* frame expected next */
    from_network_layer(&buffer);         /* fetch a packet from the network
layer */
    s.info = buffer;                     /* prepare to send the initial frame */
    s.seq = next_frame_to_send;          /* insert sequence number into
frame */
    s.ack = 1 - frame_expected;          /* piggybacked ack */
    to_physical_layer(&s);               /* transmit the frame */
    start_timer(s.seq);                  /* start the timer running */
    while (true) {
        wait_for_event(&event);          /* frame_arrival, cksum_err, or
timeout */
        if (event == frame_arrival) {    /* a frame has arrived
undamaged. */
            from_physical_layer(&r);     /* go get it */
            if (r.seq == frame_expected) /* handle inbound frame
stream. */
                to_network_layer(&r.info); /* pass packet to network
layer */
                inc(frame_expected);     /* invert seq number expected
next */
            }
            if (r.ack == next_frame_to_send) { /* handle outbound frame
stream. */
                stop_timer(r.ack);       /* turn the timer off */
                from_network_layer(&buffer); /* fetch new pkt from network
layer */
                inc(next_frame_to_send); /* invert sender's sequence
number */
            }
        }
        s.info = buffer;                 /* construct outbound frame */
        s.seq = next_frame_to_send;      /* insert sequence number into it

```

```

*/
    s.ack = 1 - frame_expected;          /* seq number of last received
frame */
    to_physical_layer(&s);                /* transmit a frame */
    start_timer(s.seq);                   /* start the timer running */
}
}

```

شکل ۳-۱۴. پروتکل پنجره لغزنده ۱-بینی.

در شرایط عادی، یکی از دو طرف پیش دستی کرده و اولین فریم را می‌فرستد. بعبارت دیگر، فقط یکی از دو لایه پیوند داده روالهای *to_physical_layer* و *start_timer* را خارج از حلقه اصلی برنامه اجرا می‌کند. اگر در پیشامدی نادر هر دو طرف بطور همزمان شروع به ارسال اولین فریم کنند، وضعیت عجیبی پیش می‌آید، که بعداً آنرا توضیح خواهیم داد. ماشین شروع‌کننده اولین بسته را از لایه شبکه گرفته، یک فریم از آن می‌سازد، و سپس ارسال می‌کند. وقتی این فریم (با هر فریم دیگری) به طرف مقابل رسید، لایه پیوند داده گیرنده چک می‌کند که تکراری نباشد (درست مثل پروتکل ۳). اگر این همان فریم موردنظر باشد، به لایه شبکه تحویل شده و پنجره دریافت یک واحد به جلو لغزنده می‌شود.

فیلد تصدیق دریافت حاوی شماره آخرین فریمیست که بدون خطا دریافت شده است. اگر این شماره با شماره فریمی که فرستنده در صدد ارسال آن است یکی باشد، فرستنده می‌فهمد که دیگر نیازی به فریم داخل بافر ندارد و می‌تواند بسته بعدی را از لایه شبکه بگیرد. اگر شماره‌ها یکی نباشند، فرستنده باید به ارسال همان فریم قبلی ادامه دهد. وقتی یک فریم دریافت شود، فریمی نیز پس فرستاده می‌شود.

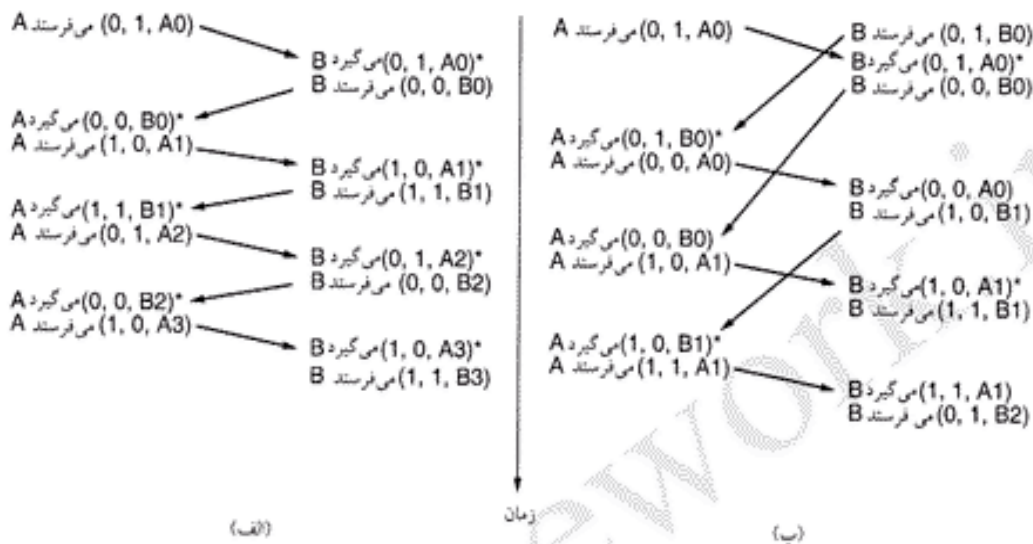
حال اجازه دهید ببینیم پروتکل ۴ در شرایط غیرعادی چگونه رفتار می‌کند. فرض کنید ماشین *A* در صدد ارسال فریم ۰ به کامپیوتر *B* است، و در همان زمان *B* نیز تصمیم می‌گیرد فریم ۰ خود را به *A* بفرستد. در ضمن فرض می‌کنیم که *A* مسابقه را زودتر شروع می‌کند، ولی فاصله زمانی تایمر آن بسیار کوتاه است. در نتیجه، ماشین *A* پشت سر هم فریمهای یکسان، با $seq = 0$ و $ack = 1$ ، به *B* می‌فرستد.

وقتی اولین فریم به کامپیوتر *B* رسید، پذیرفته شده و *frame_expected* به ۱ ست می‌شود؛ تمام فریمهای بعدی رد خواهند شد، چون اینک *B* در انتظار فریمی با شماره ترتیبی ۱ است نه ۰. علاوه بر آن، چون تمام فریمهای تکراری $ack = 1$ و *B* همچنان منتظر تصدیق دریافت فریم ۰ است، ماشین *B* گرفتن بسته از لایه شبکه خود را متوقف خواهد کرد.

بعد از آن که تمام فریمهای تکراری به *B* رسیدند، *B* یک فریم با $seq = 0$ و $ack = 0$ به *A* می‌فرستد. بالاخره یکی از این فریمها سالم به *A* می‌رسد، و باعث می‌شود تا *A* ارسال فریم بعدی را شروع کند. همانطور که می‌بینید، در هیچ شرایطی بسته تکراری به لایه شبکه نمی‌رسد، بسته‌ای گم نمی‌شود، و سیستم قفل نمی‌کند.

اما اگر هر دو طرف در یک لحظه شروع به ارسال اولین فریم کنند، وضعیت عجیبی پیش می‌آید. این مشکل ستکرون شدن را در شکل ۳-۱۵ (ب) مشاهده می‌کنید. (در شکل ۳-۱۵ الف برای مقایسه عملکرد عادی پروتکل ۴ نشان داده شده است.) اگر *B* قبل از ارسال فریمهای خود منتظر دریافت اولین فریم *A* بماند، هیچ مشکلی پیش نمی‌آید و تمام بسته‌ها براحتی پذیرفته می‌شوند (شکل الف). ولی اگر هر دو با هم مخابره اولین فریم را شروع کنند، این فریمها با یکدیگر تصادم کرده، و حالت (ب) پیش می‌آید. در (الف) دریافت هر فریم باعث گرفتن یک بسته از لایه شبکه می‌شود، و هیچ فریم تکراری هم وجود ندارد. اما در (ب) با وجود اینکه هیچ خطایی هم در

کانال وجود ندارد، نصف فریمها تکراری هستند. این وضعیت هنگامی که تایمر یکی از دو طرف بیش از حد کوتاه باشد، نیز پیش می آید (حتی اگر دو طرف همزمان شروع به مخابره فریمهای خود نکرده باشند). در چنین وضعیتی حتی امکان دارد برخی از فریمها سه یا چهار بار تکرار شوند.



شکل ۳-۱۵. دو سناریوی پروتکل ۴. (الف) حالت عادی. (ب) حالت غیرعادی. اعداد داخل پرانتز از چپ به راست عبارتند از: seq، شماره بسته. بسته‌هایی که پذیرفته و به لایه شبکه تحویل می‌شوند، با * مشخص شده‌اند.

۲-۴-۳ پروتکل «N تا به عقب برگرد»

تا اینجا بطور ضمنی فرض کرده بودیم که زمان رسیدن فریم به مقصد بعلاوه زمان برگشت فریم تصدیق ناچیز است. اما گاهی این فرض آشکارا نادرست است. در چنین مواردی زمان طولانی رفت و برگشت فریم می‌تواند تأثیر چشمگیری روی کارایی مصرف پهنای باند داشته باشد. بعنوان مثال، یک کانال ماهواره‌ای با پهنای باند 50 kbps و زمان تأخیر رفت و برگشت 500 msec را در نظر بگیرید. فرض کنید می‌خواهیم با این لینک ماهواره‌ای و با استفاده از پروتکل ۴ فریمهای 1000 بیتی ارسال کنیم. در لحظه $t = 0$ فرستنده مخابره اولین فریم را شروع می‌کند، و در $t = 20$ msec کار ارسال فریم به پایان می‌رسد. اما در بهترین شرایط (و با فرض اینکه در گیرنده نیز هیچ تأخیری وجود ندارد)، تا $t = 270$ msec این فریم هنوز به گیرنده نرسیده، و تا $t = 520$ msec نیز مسلماً فریم تصدیق دریافت به دست فرستنده نخواهد رسید. این بدان معناست که فرستنده در $500/520$ یا 96% زمان باید متوقف بماند، و نمی‌تواند چیزی بفرستد؛ بعبارت دیگر، از فقط 4% پهنای باند استفاده می‌کند. پیداست که ترکیب تأخیر طولانی در کانال، پهنای باند زیاد، و فریمهای کوچک چیزی جز اتلاف وحشتناک منابع نیست. مشکلی که در بالا دیدید از آنجا ناشی می‌شد که فرستنده برای ارسال یک فریم باید منتظر تصدیق دریافت فریم قبلی بماند. اما اگر این قید را برداریم، می‌توانیم به کارایی بهتری دست پیدا کنیم. در واقع بجای 1 فریم، معمولاً فرستنده می‌تواند w فریم بفرستد و پس از آن منتظر رسیدن فریمهای تصدیق بماند. اگر w طوری انتخاب شود که فرستنده در تمام مدت زمان تأخیر رفت و برگشت در حال ارسال فریم باشد، می‌تواند بدون مشکل از تمام پهنای باند استفاده کند. در مثال بالا w باید حداقل 26 باشد. فرستنده مانند قبل ارسال اولین فریم را در $t = 0$ شروع می‌کند، و زمانی که فریم 26 را می‌فرستد ($t = 520$ msec)، تصدیق فریم 0 را دریافت خواهد کرد. از آن به

بعد نیز فریمهای تصدیق دریافت هر 20 msec از راه می‌رسند، و فرستنده می‌تواند بطور پیوسته به ارسال فریمها ادامه دهد. در تمام زمانها فرستنده 25 یا 26 فریم در بافر خود دارد که هنوز تصدیق دریافت آنها را نگرفته است. به بیان دیگر، اندازه پنجره ارسال حداکثر 26 است.

پنجره ارسال بزرگ فقط زمانی لازم می‌شود که حاصلضرب پهنای باند \times تأخیر رفت و برگشت عددی بزرگ باشد. اگر پهنای باند زیاد باشد، حتی با تأخیر کم نیز فرستنده بسرعت پنجره ارسال را پر می‌کند. اگر تأخیر رفت و برگشت زیاد باشد (مانند کانالهای ماهواره‌ای GEO)، حتی در پهنای باند متوسط نیز پنجره ارسال بزودی پر می‌شود. ظرفیت یک کانال اساساً با حاصلضرب این دو عامل (پهنای باند و تأخیر رفت و برگشت) تعیین می‌شود، و فرستنده برای رسیدن به حداکثر کارایی باید بتواند کانال را بدون وقفه پر کند.

به این تکنیک لوله کشی (pipelining) گفته می‌شود. اگر ظرفیت کانال b bits/sec، اندازه فریم l بیت، و تأخیر رفت و برگشت R sec باشد، زمان لازم برای ارسال هر فریم l/b sec خواهد بود. بعد از ارسال آخرین بیت یک فریم، و قبل از رسیدن این بیت به گیرنده، تأخیری به میزان $R/2$ وجود دارد؛ بازگشت فریم تصدیق دریافت نیز با همین مقدار تأخیر همراه است (که کل تأخیر به R می‌رسد). با پروتکل توقف-انتظار، خط بمدت l/b کار کرده و سپس بمدت R بیکار می‌ماند، که در نتیجه

$$\text{بهره خط} = l / (l + bR)$$

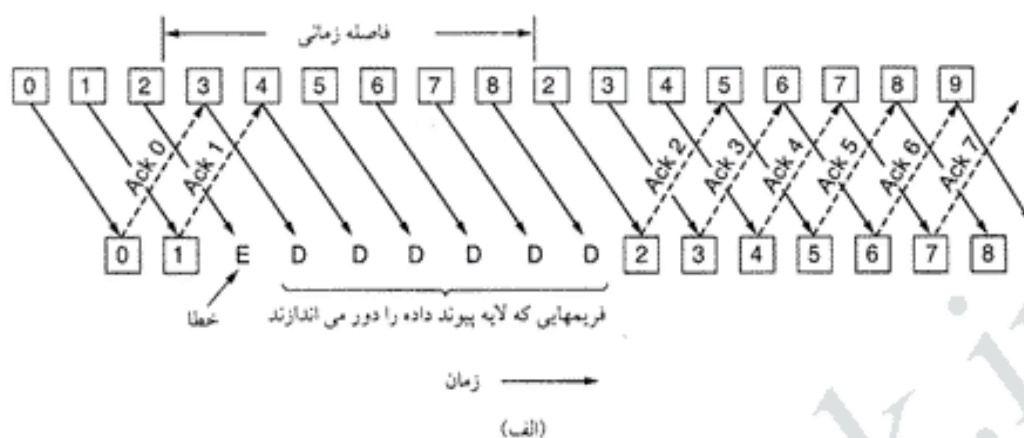
اگر $l < bR$ ، کارایی خط زیر 50% خواهد بود. از آنجائیکه تأخیر رفت و برگشت خطوط انتقال هرگز صفر نیست، با تکنیک لوله کشی می‌توان کاری کرد که خط همیشه مشغول باشد - اما اگر مقدر این تأخیر ناچیز باشد، ارزش پیچیدگی بیشتر پروتکل را ندارد.

لوله کشی فریمها در کانالهای غیرقابل اطمینان می‌تواند منجر به مشکلات جدی شود. اول اینکه، اگر یکی از فریمهای این صف طویل ناپدید یا خراب شود، چه خواهد شد؟ قبل از آنکه حتی فرستنده متوجه این خطا شود، تعداد زیادی از فریمهای بعدی به گیرنده رسیده‌اند. با رسیدن فریم خراب، گیرنده مسلماً آنرا دور می‌اندازد، اما با فریمهای سالم بعدی چه باید بکند؟ بیاد داشته باشید که لایه پیوند داده باید بسته‌ها را با ترتیب صحیح به لایه شبکه تحویل دهد. در شکل ۳-۱۶ این وضعیت (بروز خطا در خط لوله - pipeline) را ملاحظه می‌کنید. اجازه دهید آنرا دقیقتر بررسی کنیم.

برای مقابله با خطا در تکنیک لوله کشی دو رهیافت کلی وجود دارد. در رهیافت اول، که به «N تا به عقب برگرد» معروف است، گیرنده تمام فریمهای بعد از فریم خراب را دور می‌اندازد و هیچ تصدیقی برای آنها برنمی‌گرداند. این استراتژی معادل است با پنجره دریافتی باندازه 1. بعبارت دیگر، لایه پیوند داده در گیرنده هیچ فریمی غیر از آن فریمی که باید به لایه شبکه تحویل دهد، را قبول نمی‌کند. اگر پنجره ارسال فرستنده قبل از انقضای تایمر پر شود، خط لوله شروع به خالی شدن خواهد کرد. پس از آن تایمر فرستنده به انتها رسیده و تمام فریمهای باقی مانده (از فریمی که خراب یا گم شده) را دوباره ارسال خواهد کرد. اگر نرخ خطا در خط زیاد باشد (مانند کانالهای بیسیم)، این رهیافت باعث اتلاف شدید پهنای باند خواهد شد.

در شکل ۳-۱۶ (الف) این حالت را ملاحظه می‌کنید. در این مثال فریمهای 0 و 1 سالم به مقصد رسیده‌اند، ولی فریم 2 خراب شده است. فرستنده تا زمانی که تایمر این فریم منقضی نشود، از خراب شدن آن مطلع نخواهد شد. پس از آنکه فرستنده فهمید فریم 2 سالم به مقصد نرسیده، برمی‌گردد و ارسال فریمها را از این فریم از سر می‌گیرد.

رهیافت دوم مقابله با خطا در تکنیک لوله کشی تکرار انتخابی (selective repeat) نام دارد. در این رهیافت، فریم خراب در گیرنده دور انداخته شده، ولی فریمهای سالم بعدی بافر می‌شوند. وقتی تایمر فریم معیوب منقضی



شکل ۳-۱۶. مقابله با خطا در خط لوله. تأثیر خطا وقتی که (الف) اندازه پنجره دریافت گیرنده ۱ است، و (ب) پنجره دریافت بزرگ است.

شد، فرستنده فقط همین فریم (که قدیمی ترین فریم تصدیق نشده است) را مجدداً ارسال می کند. اگر این فریم سالم به مقصد رسید، گیرنده این فریم و فریمهای بافر شده را بترتیب به لایه شبکه تحویل می دهد. در رهیافت تکرار انتخابی، معمولاً گیرنده برای فریمهای خراب (فریمهایی که جمع تطبیقی آنها اشتباه است، یا خارج از نظم وارد می شوند) نیز یک فریم تصدیق دریافت منفی (negative acknowledgement) - که به فریم NAK معروف است - به فرستنده برمی گرداند. فریم NAK کارایی سیستم را به مقدار زیادی بهبود می بخشد، چون باعث می شود که فرستنده قبل از انقضای تایمر کار ارسال مجدد فریمهای از دست رفته را شروع کند.

در شکل ۳-۱۶ (ب) نیز فریمهای ۰ و ۱ سالم به مقصد رسیده اند، ولی فریم ۲ خراب شده است. وقتی فریم ۳ به گیرنده می رسد، لایه پیوند داده متوجه می شود که یک فریم جا افتاده است، پس یک NAK برای فریم ۲ به فرستنده فرستاده، و فریم ۳ را در بافری که برای این منظور اختصاص یافته، ذخیره می کند. فریمهای ۴ و ۵ نیز پس از رسیدن به گیرنده، بجای تحویل به لایه شبکه، بافر می شوند. با رسیدن NAK فریم ۲، فرستنده بلافاصله این فریم را ارسال می کند، که با رسیدن آن به مقصد، گیرنده می تواند فریم ۲ و فریمهای پس از آن (تا فریم ۵) را به لایه شبکه تحویل دهد (و تصدیق دریافت آنها را به فرستنده برگرداند). حتی اگر NAK فریم ۲ در راه گم شود و به دست فرستنده نرسد، باز هم پس از انقضای تایمر، فرستنده نسبت به ارسال مجدد آن (و فقط همین یک فریم) اقدام خواهد کرد؛ که البته در این میان فقط وقت بیشتری تلف خواهد شد. در واقع، NAK فقط ارسال مجدد

فریمهای معیوب را تسریع می‌کند.

رهیافت تکرار انتخابی خاص پنجره‌های دریافت بزرگتر از 1 است. هر فریمی در داخل پنجره دریافت قرار داشته باشد، می‌تواند بافر شود تا فریمهای قبل از آن دریافت و به لایه شبکه تحویل شوند. البته اگر پنجره دریافت خیلی بزرگ باشد، لایه پیوند داده به حافظه زیادی برای بافر کردن فریمها نیاز دارد.

این دو رهیافت داد و ستدی هستند بین پهنای باند و فضای بافر لایه پیوند داده، که بسته به اهمیت هر یک از این منابع می‌توان یکی از رهیافت‌های «N تا به عقب برگرد» یا «تکرار انتخابی» را انتخاب کرد. در شکل ۳-۱۷ پروتکل لوله‌کشی با پنجره دریافت 1 را ملاحظه می‌کنید؛ در این پروتکل تمام فریمهای بعد از یک فریم معیوب دور انداخته می‌شوند. در این پروتکل برای اولین بار فرض نامحدود بودن بسته‌هایی که لایه شبکه می‌تواند ارائه کند، را نیز کنار گذاشته‌ایم. در اینجا، وقتی لایه شبکه آماده ارسال یک بسته است، رویداد `network_layer_ready` را تحریک می‌کند. با این حال، برای آنکه هیچوقت بیش از `MAX_SEQ` فریم وارد صف ارسال لایه پیوند داده نشود، باید کاری کنیم که این لایه بتواند بطرفی مانع ارسال بسته‌های اضافی از لایه شبکه شود - برای این منظور از روالهای کتابخانه‌ای `enable_network_layer` و `disable_network_layer` استفاده کرده‌ایم.

```
/* Protocol 5 (go back n) allows multiple outstanding frames. The sender may transmit up
to MAX_SEQ frames without waiting for an ack. In addition, unlike in the previous
protocols, the network layer is not assumed to have a new packet all the time. Instead,
the network layer causes a network_layer_ready event when there is a packet to send.
*/
```

```
#define MAX_SEQ 7 /* should be 2^n - 1 */
typedef enum {frame_arrival, cksum_err, timeout, network_layer_ready} event_type;
#include "protocol.h"
```

```
static boolean between(seq_nr a, seq_nr b, seq_nr c)
```

```
{
/* Return true if a <= b < c circularly; false otherwise. */
if (((a <= b) && (b < c)) || ((c < a) && (a <= b)) || ((b < c) && (c < a)))
return(true);
else
return(false);
}
```

```
static void send_data(seq_nr frame_nr, seq_nr frame_expected, packet buffer[])
```

```
{
/* Construct and send a data frame. */
frame s; /* scratch variable */

s.info = buffer[frame_nr]; /* insert packet into frame */
```

```

s.seq = frame_nr;                /* insert sequence number into frame */
s.ack = (frame_expected + MAX_SEQ) % (MAX_SEQ + 1); /* piggyback ack */
to_physical_layer(&s);          /* transmit the frame */
start_timer(frame_nr);          /* start the timer running */
}

void protocol5(void)
{
    seq_nr next_frame_to_send;    /* MAX_SEQ > 1; used for outbound
stream */
    seq_nr ack_expected;          /* oldest frame as yet unacknowledged */
    seq_nr frame_expected;        /* next frame expected on inbound
stream */
    frame r;                      /* scratch variable */
    packet buffer[MAX_SEQ + 1];    /* buffers for the outbound stream */
    seq_nr nbuffered;             /* # output buffers currently in use */
    seq_nr i;                     /* used to index into the buffer array */
    event_type event;

    enable_network_layer();        /* allow network_layer_ready events */
    ack_expected = 0;              /* next ack expected inbound */
    next_frame_to_send = 0;        /* next frame going out */
    frame_expected = 0;           /* number of frame expected inbound */
    nbuffered = 0;                /* initially no packets are buffered */
    while (true) {
        wait_for_event(&event);    /* four possibilities: see event_type
above */

        switch(event) {
            case network_layer_ready: /* the network layer has a packet to send
*/

                /* Accept, save, and transmit a new frame. */
                from_network_layer(&buffer[next_frame_to_send]); /* fetch new packet */
                nbuffered = nbuffered + 1; /* expand the sender's window */
                send_data(next_frame_to_send, frame_expected, buffer); /* transmit the
frame */

                inc(next_frame_to_send); /* advance sender's upper window
edge */

                break;

            case frame_arrival:      /* a data or control frame has arrived */

```



```

        from_physical_layer(&r);          /* get incoming frame from physical
layer */

        if (r.seq == frame_expected) {
            /* Frames are accepted only in order. */
            to_network_layer(&r.info);    /* pass packet to network layer */
            inc(frame_expected);          /* advance lower edge of receiver's
window */
        }

        /* Ack n implies n - 1, n - 2, etc. Check for this. */
        while (between(ack_expected, r.ack, next_frame_to_send)) {
            /* Handle piggybacked ack. */
            nbuffered = nbuffered - 1; /* one frame fewer buffered */
            stop_timer(ack_expected); /* frame arrived intact; stop timer */

            inc(ack_expected);          /* contract sender's window */
        }
        break;

        case cksum_err: break;          /* just ignore bad frames */

        case timeout:                   /* trouble; retransmit all outstanding
frames */
            next_frame_to_send = ack_expected; /* start retransmitting here */
            for (i = 1; i <= nbuffered; i++) {
                send_data(next_frame_to_send, frame_expected, buffer); /* resend
1 frame */

                inc(next_frame_to_send); /* prepare to send the next one */
            }
        }

        if (nbuffered < MAX_SEQ)
            ;(royal_krowten_elbane
            else
                disable_network_layer();
        }
    }
}

```

شکل ۳-۱۷. پروتکل پنجره لغزنده N تا به عقب برگردد.

توجه کنید که در هر لحظه نباید بیش از MAX_SEQ فریم (و نه $MAX_SEQ + 1$) در صف ارسال وجود داشته باشد - حتی با وجود اینکه تعداد شماره‌های ترتیبی، در $MAX_SEQ + 1$ است: $0, 1, 2, \dots, MAX_SEQ$. برای پی بردن به علت این محدودیت، سناریوی زیر را که در آن $MAX_SEQ = 7$ نظر بگیرید:

۱. فرستنده فریمهای 0 تا 7 (هشت فریم) را ارسال می‌کند.
۲. پس از مدتی، تصدیق دریافت فریم 7 (با استفاده از تکنیک سواری مجانی) به فرستنده برمی‌گردد.
۳. فرستنده 8 فریم بعدی را با همان شماره‌های ترتیبی 0 تا 7 ارسال می‌کند.
۴. حال فریم تصدیق دریافت دیگری برای فریم شماره 7 (به همان صورت سواری مجانی) از راه می‌رسد.

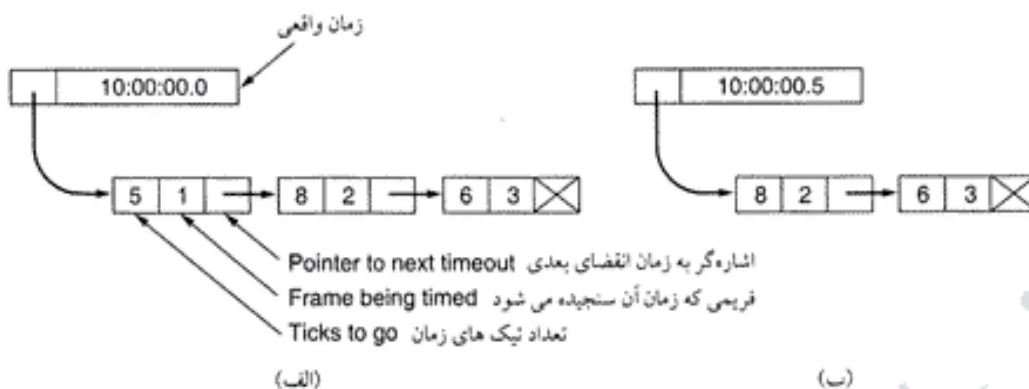
سؤال این است: آیا تمام 8 فریمی که متعلق به دسته دوم بودند، سالم به مقصد رسیده‌اند، یا (با توجه به اینکه تمام فریمهای پس از یک فریم خراب دور انداخته می‌شوند) همگی از بین رفته‌اند؟ اگر دقت کرده باشید، در هر دو حالت گیرنده تصدیق دریافت فریم 7 را پس می‌فرستد، و فرستنده هم راهی برای تشخیص این موضوع ندارد. به همین دلیل حداکثر فریمهایی که در صف ارسال می‌ایستند، باید MAX_SEQ باشد.

با اینکه پروتکل 5 فریمهای پس از یک فریم معیوب را بافر نمی‌کند، اما هنوز به مقداری بافر در سمت فرستنده نیاز داریم. از آنجائیکه فرستنده باید تمام فریمهای داخل پنجره ارسال را تا رسیدن تصدیق دریافت آنها نگه دارد (چون ممکنست لازم شود آنها را دوباره بفرستد)، باید فضای کافی برای بافر کردن این فریمها داشته باشد. در این پروتکل وقتی تصدیق دریافت فریم n از راه برسد، فریمهای $n-1$ ، $n-2$ ، و ... نیز بطور خودکار تصدیق شده محسوب می‌شوند. این ویژگی بویژه اگر فریمهای تصدیق قبلی در راه برگشت به فرستنده از بین بروند، اهمیت می‌یابد. با رسیدن هر فریم تصدیق دریافت، لایه پیوند داده چک می‌کند که کدام بافرها را می‌تواند آزاد کند. وقتی بافر آزاد (و جایی در پنجره ارسال باز) شد، لایه شبکه (که قبلاً متوقف شده) دوباره با رویداد `enable_network_layer` فعال می‌شود و می‌تواند بسته‌های بعدی را به لایه پیوند داده بفرستد.

در پروتکل 5 فرض کرده‌ایم که همیشه ترافیک برگشتی کافی برای سواری مجانی وجود دارد. اگر چنین نباشد، فریمهای تصدیق دریافت را هم نمی‌توان ارسال کرد. در پروتکل 4 چنین فرضی وجود نداشت، و برای هر فریم یک فریم تصدیق دریافت مستقل پس فرستاده می‌شد. در پروتکل آینده این مشکل را هم بنحو جالبی حل خواهیم کرد.

دیدید که در پروتکل 5، فرستنده می‌تواند در هر لحظه تعداد زیادی فریم ارسال شده (ولی هنوز تصدیق نشده) در بافر خود داشته باشد: هر یک از این فریمها یک تایمر مستقل می‌خواهند. این تایمرها را می‌توان بصورت نرم‌افزاری و با استفاده از وقفه‌های ساعت سخت‌افزاری، ایجاد کرد. این تایمرها تشکیل یک لیست پیوندی (linked list) می‌دهند، که هر گره آن سه بخش دارد: زمان باقی مانده تایمر، فریمی که به این تایمر مربوط است، و یک اشاره‌گر به گره بعدی.

در شکل 3-18 (الف) طرز پیاده‌سازی این تایمرها را می‌بینید. فرض کنید که ساعت سیستم هر 100 msec یک تیک (وقفه سخت‌افزاری) می‌فرستد. در لحظه شروع، که زمان واقعی 10:00:00.0 است، سه تایمر برای زمانهای 10:00:00.5، 10:00:01.3 و 10:00:01.9 ست می‌شوند. با هر تیک ساعت سخت‌افزاری، شمارنده تایمری که در رأس لیست قرار دارد، کاهش می‌یابد. وقتی این شمارنده 0 شد، گره مربوطه از لیست حذف می‌شود؛ شکل 3-18 (ب) را ببینید. سازماندهی تایمرها به صورت فوق باعث می‌شود که در هر بار فراخوانی روالهای `start_timer` و `stop_timer` کل لیست اسکن شود، ولی از سوی دیگر کار لازم برای به روز در آوردن تایمرها در هر تیک بسیار ناچیز است. همانطور که می‌بینید، در پروتکل 5 روالهای `start_timer` و `stop_timer` پارامتری می‌گیرند، که نشان می‌دهد زمان کدام فریم بایستی سنجیده شود.



شکل ۳-۱۸. شبیه سازی چند تایمر بوسیله نرم افزار.

۳-۴-۳ پروتکل تکرار انتخابی

اگر نرخ خطا ناچیز باشد، پروتکل ۵ بخوبی کار خواهد کرد، اما در خطوط پرنویز این پروتکل مقدار زیادی از پهنای باند را (برای ارسال مجدد فریمها) به هدر می دهد. استراتژی دیگر مقابله با خطاهای خط اینست که به گیرنده اجازه دهیم فریمهای سالمی که بعد از یک (یا چند) فریم معیوب از راه می رسند، را بافر کنند. چنین پروتکلی دیگر فریمها را به صرف اینکه فریم قبلی آنها خراب بوده یا گم شده، دور نمی اندازد.

در این پروتکل، فرستنده و گیرنده هر دو پنجره ای از شماره های قابل قبول دارند. پنجره ارسال در فرستنده از 0 شروع شده، و می تواند تا MAX_SEQ بالا برود؛ اما اندازه پنجره دریافت همیشه ثابت، و معادل MAX_SEQ است. گیرنده برای هر یک از فریمهایی که در پنجره دریافت قرار می گیرند، یک بافر ثابت دارد، که هر یک از این بافرها دارای بیتی هستند (بنام بیت *arrived*) که مشخص می کند بافر پر است یا خالی. وقتی یک فریم بدست گیرنده می رسد، شماره آن را با تابع *between* چک می کند تا ببیند در داخل پنجره دریافت است یا خارج آن. اگر فریم داخل پنجره باشد و تکراری هم نباشد، پذیرفته شده و در بافر ذخیره می شود (خواه این فریم محتوی بسته ای باشد که باید در نوبت بعد به لایه شبکه داده شود، یا نه). البته این فریم تا زمانی که تمام فریمهای قبلی از آن به لایه شبکه تحویل نشده اند، در لایه پیوند داده می ماند. در شکل ۳-۱۹ پروتکلی که بر اساس این الگوریتم نوشته شده، را ملاحظه می کنید.

```
/* Protocol 6 (selective repeat) accepts frames out of order but passes packets to the
network layer in order. Associated with each outstanding frame is a timer. When the
timer
expires, only that frame is retransmitted, not all the outstanding frames, as in
protocol 5. */
```

```
#define MAX_SEQ 7 /* should be 2^n - 1 */
#define NR_BUFS ((MAX_SEQ + 1)/2)
typedef enum {frame_arrival, cksum_err, timeout, network_layer_ready, ack_timeout}
event_type;
#include "protocol.h"
boolean no_nak = true; /* no nak has been sent yet */
```

```

seq_nr oldest_frame = MAX_SEQ + 1;      /* initial value is only for the simulator
*/
static boolean between(seq_nr a, seq_nr b, seq_nr c)
{
/* Same as between in protocol5, but shorter and more obscure. */
return ((a <= b) && (b < c)) || ((c < a) && (a <= b)) || ((b < c) && (c < a));
}
static void send_frame(frame_kind fk, seq_nr frame_nr, seq_nr frame_expected,
packet buffer[])
{
/* Construct and send a data, ack, or nak frame. */
frame s;                                /* scratch variable */

s.kind = fk;                             /* kind == data, ack, or nak */
if (fk == data) s.info = buffer[frame_nr % NR_BUFS];
s.seq = frame_nr;                         /* only meaningful for data frames */
s.ack = (frame_expected + MAX_SEQ) % (MAX_SEQ + 1);
if (fk == nak) no_nak = false;           /* one nak per frame, please */
to_physical_layer(&s);                   /* transmit the frame */
if (fk == data) start_timer(frame_nr % NR_BUFS);
stop_ack_timer();                         /* no need for separate ack frame */
}
void protocol6(void)
{
seq_nr ack_expected;                      /* lower edge of sender's window */
seq_nr next_frame_to_send;               /* upper edge of sender's window +
1 */
seq_nr frame_expected;                   /* lower edge of receiver's window */
seq_nr too_far;                          /* upper edge of receiver's window + 1
*/

int i;                                    /* index into buffer pool */
frame r;                                  /* scratch variable */
packet out_buf[NR_BUFS];                 /* buffers for the outbound stream */
packet in_buf[NR_BUFS];                  /* buffers for the inbound stream */
boolean arrived[NR_BUFS];                /* inbound bit map */
seq_nr nbuffered;                         /* how many output buffers currently
used */
event_type event;

enable_network_layer();                  /* initialize */
ack_expected = 0;                        /* next ack expected on the inbound

```

```

stream */
    next_frame_to_send = 0;           /* number of next outgoing frame */
    frame_expected = 0;
    too_far = NR_BUFS;
    nbuffered = 0;                   /* initially no packets are buffered */
    for (i = 0; i < NR_BUFS; i++) arrived[i] = false;
    while (true) {
        wait_for_event(&event);      /* five possibilities: see event_type
above */
        switch(event) {
            case network_layer_ready: /* accept, save, and transmit a new
frame */
                nbuffered = nbuffered + 1; /* expand the window */
                from_network_layer(&out_buf[next_frame_to_send % NR_BUFS]); /* fetch
new packet */
                send_frame(data, next_frame_to_send, frame_expected, out_buf); /*
transmit the frame */
                inc(next_frame_to_send); /* advance upper window edge */
                break;

            case frame_arrival:      /* a data or control frame has arrived
*/
                from_physical_layer(&r); /* fetch incoming frame from
physical layer */
                if (r.kind == data) {
                    /* An undamaged frame has arrived. */
                    if ((r.seq != frame_expected) && no_nak)
                        send_frame(nak, 0, frame_expected, out_buf); else
                        start_ack_timer();
                    if (between(frame_expected, r.seq, too_far) &&
(arrived[r.seq%NR_BUFS] == false)) {
                        /* Frames may be accepted in any order. */
                        arrived[r.seq % NR_BUFS] = true; /* mark buffer as
full */
                        in_buf[r.seq % NR_BUFS] = r.info; /* insert data into
buffer */

                        while (arrived[frame_expected % NR_BUFS]) {
                            /* Pass frames and advance window. */
                            to_network_layer(&in_buf[frame_expected % NR_BUFS]);
                            no_nak = true;
                            arrived[frame_expected % NR_BUFS] = false;

```

```

    inc(frame_expected); /* advance lower edge of receiver's
window */

                                inc(too_far); /* advance upper edge of
receiver's window */

                                start_ack_timer(); /* to see if a separate ack is
needed */

                                }
                                }
                                }
    if((r.kind==nak)
between(ack_expected,(r.ack+1)%(MAX_SEQ+1),next_frame_to_send))
    send_frame(data, (r.ack+1) % (MAX_SEQ + 1), frame_expected,
out_buf);

    while (between(ack_expected, r.ack, next_frame_to_send)) {
        nbuffered = nbuffered - 1; /* handle piggybacked ack */
        stop_timer(ack_expected % NR_BUFS); /* frame arrived intact
*/
        inc(ack_expected); /* advance lower edge of sender's
window */
    }
    break;

    case cksum_err:
        if (no_nak) send_frame(nak, 0, frame_expected, out_buf); /* damaged
frame */
        break;

    case timeout:
        send_frame(data, oldest_frame, frame_expected, out_buf); /* we timed
out */
        break;

    case ack_timeout:
        send_frame(ack,0,frame_expected, out_buf); /* ack timer expired;
send ack */
    }
    ;)(reyal_krowten_elbasid esle ;)(reyal_krowten_elbane )SFUB_RN < dereffubn( fi
    }
    }

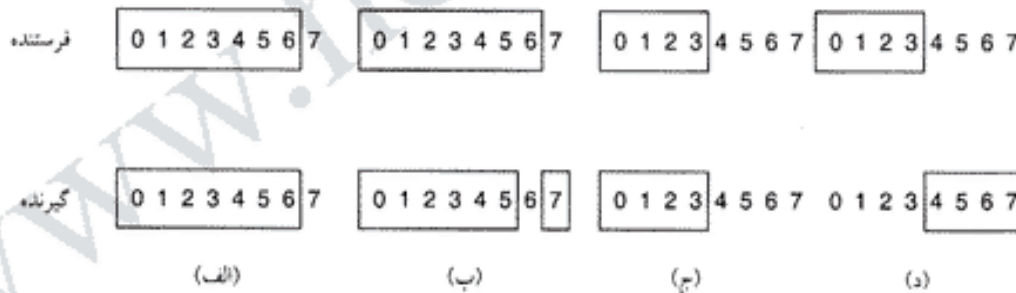
```

شکل ۳-۱۹. پروتکل پنجره لغزنده «تکرار انتخابی».

دریافت نامنظم فریمها مسائلی را بوجود می آورد که در پروتکل های قبلی (که فریمها را فقط بترتیب شماره قبول می کنند) وجود نداشت. با یک مثال بهتر می توان این مشکل را نشان داد. فرض کنید از شماره های ترتیبی سه بیتی استفاده می کنیم، بنابراین فرستنده قبل از توقف برای رسیدن اولین فریم تصدیق دریافت می تواند حداکثر هفت فریم ارسال کند. در لحظه شروع، پنجره های ارسال و دریافت مانند شکل ۳-۲۰ (الف) هستند. فرستنده فریمهای 0 تا 6 را می فرستد. پنجره دریافت گیرنده فقط اجازه پذیرش فریمهایی را می دهد که شماره آنها (منحصراً) بین 0 تا 6 باشد. تمام هفت فریم اول سالم به مقصد می رسند، بنابراین گیرنده دریافت آنها را تصدیق کرده، پنجره دریافت را برای دریافت سری بعدی فریمها (7، 0، 1، 2، 3، 4، 5) بجلو می برد، و تمام بافرها را هم با علامت «خالی» نشانه گذاری می کند؛ شکل ۳-۲۰ (ب) را ببینید.

درست در همین لحظه یک صاعقه به خط تلفن اصابت کرده، و تمام فریمهای تصدیق دریافت را (در راه بازگشت به فرستنده) از بین می برد. پس از مدتی تایمر فریم 0 به انتها رسیده، و فرستنده این فریم را مجدداً ارسال می کند. وقتی این فریم تکراری به گیرنده رسید، گیرنده چک می کند که آیا در داخل پنجره دریافت هست یا خیر. متأسفانه همانطور که در شکل ۳-۲۰ (ب) می بینید، فریم 0 هنوز در داخل پنجره دریافت قرار دارد، بنابراین گیرنده آنرا قبول می کند. از آنجائیکه فریمهای 0 تا 6 دریافت شده اند، گیرنده تصدیق دریافت فریم 6 را سوار فریم بعدی کرده و به فرستنده برمی گرداند.

فرستنده هم خوشحال از اینکه تمام فریمهای فرستاده شده سالم به مقصد رسیده اند، پنجره ارسال را بجلو برده و بلافاصله فریمهای 7، 0، 1، 2، 3، 4 و 5 را می فرستد. وقتی فریم 7 به مقصد رسید، لایه پیوند داده گیرنده آنرا تحویل لایه شبکه می دهد. بلافاصله پس از آن لایه پیوند داده چک می کند که آیا بسته 0 معتبری وجود دارد یا خیر، و چون چنین بسته ای در بافرهای آن موجود است، آنرا به لایه شبکه می دهد؛ در حالیکه می دانیم این همان بسته 0 اول است و نباید دوباره به لایه شبکه داده شود - پس پروتکل ما مرتکب خطا شده است.



شکل ۳-۲۰. (الف) پنجره های ارسال و دریافت هفت تایی در لحظه شروع. (ب) بعد از رسیدن هفت فریم به مقصد، و قبل از بازگشت تصدیق دریافت به فرستنده. (ج) پنجره های ارسال و دریافت چهار تایی در لحظه شروع. (د) بعد از رسیدن چهار فریم به مقصد، و قبل از بازگشت تصدیق دریافت به فرستنده.

منشأ این مشکل آنجاست که بعد از جلو رفتن پنجره دریافت درگیرنده، شماره های معتبر جدید با شماره های قدیمی همپوشانی (overlap) دارد. در نتیجه، فریمهای بعدی می توانند تکراری باشند (اگر تمام فریمهای تصدیق دریافت از بین برود) یا نباشند (اگر تمام فریمهای تصدیق دریافت سالم به فرستنده برسند). گیرنده بیچاره هیچ راهی برای تشخیص این وضعیت ندارد.

راه چاره این معضل آن است که مطمئن شویم بعد از جلو رفتن پنجره گیرنده، با پنجره اصلی همپوشانی نداشته باشد. برای رسیدن به این هدف پنجره دریافت باید از نصف تعداد شماره های ترتیبی تجاوز نکند؛ شکل

۳-۲۰ (پ) و (ت) را ببینید. برای مثال، اگر از شماره‌های ترتیبی ۴-بیتی استفاده کنیم، محدوده ما 0 تا 15 خواهد بود، و فرستنده نباید در هر لحظه بیش از هشت فریم تصدیق نشده در بافر خود داشته باشد. بدین ترتیب، وقتی گیرنده فریمهای 0 تا 7 را گرفته و پنجره دریافت را جلو ببرد، به سری 8 تا 15 می‌رسد که آشکار است دیگر با مشکل قبلی مواجه نخواهد شد. در حالت کلی، اندازه پنجره در پروتکل ۶ بایستی حداکثر $(MAX_SEQ + 1)/2$ باشد. در مثال قبل، پنجره ارسال و دریافت را باید $(7 + 1)/2 = 4$ انتخاب کنیم.

و حالا یک سؤال جالب دیگر: گیرنده چند بافر باید داشته باشد؟ گیرنده در هیچ شرایطی فریمی که شماره آن کمتر از لایه پائین پنجره دریافت یا بیشتر از لایه بالای آن باشد، را قبول نخواهد کرد. در نتیجه، تعداد بافرهای گیرنده باید معادل اندازه پنجره دریافت باشد (نه تعداد شماره‌های ترتیبی). با شماره‌های ترتیبی ۴-بیتی، گیرنده به حداکثر هشت بافر نیاز دارد. وقتی فریم i بدست گیرنده می‌رسد، آنرا در بافری بشماره $i \text{ MOD } 8$ قرار می‌دهد. شاید حدس زده باشید که فریمهای i و $i + 8$ هر دو در یک بافر قرار می‌گیرند، ولی توجه کنید که این دو فریم هرگز در یک پنجره واقع نمی‌شوند (چون برای آنکه چنین اتفاقی بیفتد، اندازه پنجره باید حداقل 9 باشد).

به دلیل مشابه، تعداد تایمرهای فرستنده نیز باید معادل پنجره دریافت باشد، نه تعداد شماره‌های ترتیبی (چون هر تایمر به یک بافر وابسته است و وقتی تایمر به انتها می‌رسد، بافر دوباره ارسال می‌شود). در پروتکل ۵ این پیش‌فرض ضمنی را داشتیم که بار کانال بسیار زیاد است: وقتی یک فریم از راه می‌رسد، تصدیق آن بلافاصله برگردانده نمی‌شود تا سوار فریم بعدی (که از گیرنده به فرستنده می‌رود) شود. اما اگر ترافیک در جهت مخالف کم باشد، فریمهای تصدیق دریافت خیلی معطل خواهند شد. در این روش، وقتی فرستنده MAX_SEQ بسته فرستاد، منتظر می‌ماند و از آنجائیکه ترافیک جهت مقابل کم است، مدت زیادی بیکار خواهد ماند؛ علت فرض زیاد بودن بار کانال نیز همین موضوع است.

در پروتکل ۶ این مشکل نیز برطرف شده است. وقتی یک فریم (که طبق ترتیب مورد انتظار فرستاده شده) از راه رسید، گیرنده یک تایمر کمکی را (با تابع $start_ack_timer$) راه می‌اندازد. اگر در مدتی که این تایمر منقضی می‌شود، فریمی برای ارسال تحویل لایه پیوند داده نشد، یک فریم تصدیق دریافت مستقل برگردانده خواهد شد (رویداد این تایمر $ack_timeout$ نام دارد). با این تمهید دیگر نیازی نیست متکی به ترافیک سنگین دو طرفه باشیم، و پروتکل ۶ حتی می‌تواند بصورت کاملاً یکطرفه هم کار کند. از این تایمر کمکی فقط یکی وجود دارد، و اجرای تابع $start_ack_timer$ آنرا ریست می‌کند. البته ضروری است که فاصله زمانی تایمر کمکی بسیار کوتاهتر از فاصله زمانی تایمرهای فریمهای داده باشد، تا این تایمرها قبل از رسیدن تصدیق دریافت منقضی نشوند.

استراتژی مقابله با خطا در پروتکل ۶ بسیار کارآمدتر از پروتکل ۵ است. اگر گیرنده به هر دلیلی ظن خطا ببرد، یک فریم تصدیق دریافت منفی (NAK) به فرستنده پس می‌فرستد. این NAK صریحاً از فرستنده می‌خواهد که فریم مشخص شده را دوباره ارسال کند. دو حالت وجود دارد که گیرنده باید به بروز خطا مشکوک شود: دریافت یک فریم معیوب، یا دریافت فریمی که انتظار آنرا ندارد. برای اجتناب از تکرار این درخواستها، گیرنده لیستی از فریمهایی که برای آنها NAK فرستاده را نگه می‌دارد. اگر متغیر no_nak در پروتکل ۶ مقدار true داشته باشد، نشان می‌دهد که هنوز برای فریم $frame_expected$ هیچ NAK فرستاده نشده است. خراب یا گم شدن NAK ها مشکلی بوجود نخواهد آورد، فقط زمان ارسال مجدد فریمها را کمی به تأخیر می‌اندازد (چون تایمر فرستنده بهر حال منقضی خواهد شد). اگر پس از NAK فریم خواسته شده از راه برسد، گیرنده مقدار no_nak را به true ست کرده و تایمر کمکی را راه می‌اندازد ($start_ack_timer$). پس از انقضای این تایمر، گیرنده یک فریم ACK به فرستنده پس می‌فرستد، و بدین ترتیب خود را با آن سنکرون می‌کند.

در برخی شرایط، زمان لازم برای سفر فریمهای داده به مقصد، پردازش در آنجا، و بازگشت فریم تصدیق

دریافت (تقریباً) ثابت است. در چنین شرایطی فرستنده می تواند تایمرهای خود را کمی بالاتر از این مقدار تنظیم کند. اما اگر این زمان در حد وسیعی متغیر باشد، فرستنده دو راه در پیش رو دارد: فاصله زمانی تایمرهای خود را خیلی کوچک بگیرد (و ریسک ارسالهای تکراری را بپذیرد)، یا آنرا بسیار بزرگ بگیرد (و بعد از هر خطا مدت زیادی بیکار بماند). هر دوی این گزینه ها تلف کردن پهنای باند است.

اگر ترافیک جهت مخالف کم باشد، زمان برگشت فریمهای تصدیق دریافت نیز نامنظم خواهد بود (گاهی کم و گاهی زیاد). تغییر زمان پردازش فریم در گیرنده هم می تواند مزید علت باشد. در کل، اگر انحراف معیار فاصله زمانی فریم تصدیق دریافت (در مقایسه با کل فاصله زمانی) کوچک باشد، می توان فاصله زمانی تایمرها را «کوچک» در نظر گرفت، که در این حالت NAK سودمندی خود را از دست می دهد. در غیر اینصورت، برای اجتناب از تکرار در ارسال فریمها باید فاصله زمانی تایمرها را «بزرگ» گرفت؛ در این حالت استفاده از NAK می تواند ارسال مجدد فریمهای معیوب و گم شده را تسریع کند.

سؤال دیگری که در همین زمینه پیش می آید این است که: کدام فریم باعث انقضای تایمر شده است؟ پروتکل 5 همیشه *ack_expected* است (یعنی فقط انتظار دریافت تصدیق را دارد)، چون همیشه قدیمی ترین باعث انقضای تایمر می شود. اما در پروتکل 6 راه ساده ای برای تعیین این موضوع وجود ندارد. فرض کنید فریمهای 0 تا 4 فرستاده شده اند، و لیست فریمهای بافر شده در فرستنده از قدیم ترین به جدیدترین (از چپ بر راست) عبارتند از: 01234. حال وضعیت زیر را در نظر بگیرید: فریم 0 منقضی می شود، فرستنده فریم (جدید) 5 را می فرستد، فریم 1 منقضی می شود، و فرستنده فریم (جدید) 6 را می فرستد. در این لحظه، بافر فرستنده حاوی فریمهای 3405126 (از قدیم به جدید) است. اگر در این لحظه تمام فریمهای برگشتی (که فرض می کنیم حامل فریمهای تصدیق دریافت هستند) از بین بروند، هفت فریمی که در بافر قرار دارند، به همین ترتیب منقضی می شوند. برای اجتناب از پیچیدگی بیشتر (که تا همین جا هم باندازه کافی پیچیده هست)، به مدیریت تایمرها نپرداختیم. بجای آن فرض کردیم که متغیر *oldest_frame* در لحظه انقضای تایمر نشان می دهد که کدام فریم منقضی شده است.

۵-۳ ارزیابی پروتکل ها

پروتکل های واقعی (و کُد نرم افزاری آنها) اغلب بسیار پیچیده اند، به همین دلیل تحقیقات زیادی انجام شده تا روشهای ریاضی و مشخصی برای ارزیابی آنها ابداع شود. در این قسمت برخی از این تکنیکها و مدلها را بررسی خواهیم کرد. با اینکه در اینجا برای ارزیابی پروتکل های لایه پیوند داده از این تکنیکها استفاده کرده ایم، ولی آنها را می توان برای لایه های دیگر نیز بکار گرفت.

۱-۵-۳ مدل ماشین حالت محدود

یکی از مفاهیم کلیدی در مدلسازی پروتکلها، ماشین حالت محدود (finite state machine) است. در این تکنیک، هر ماشین پروتکل (protocol machine) - یعنی، فرستنده یا گیرنده - در هر لحظه از زمان در حالتی خاص قرار دارد. این حالت (state) عبارتست از مقدار تمام متغیرها، و شمارنده برنامه (program counter). در اکثر مواقع می توان تعداد زیادی از حالتها را برای آنالیز دسته بندی کرد. برای مثال، گیرنده پروتکل 3 را در نظر بگیرید؛ تمام حالت های ممکن این گیرنده را می توان به دو دسته مهم تقسیم کرد: انتظار برای فریم 0، و انتظار برای فریم 1. تمام حالت های دیگر را می توان مراحل گذار از یکی از این دو حالت به حالت دیگر دانست. معمولاً حالتها را بگونه ای انتخاب می کنند که در آن لحظه ماشین پروتکل در انتظار وقوع رویداد بعدی است (در مثال ما، اجرای روال *wait(event)*). در این لحظه، حالت پروتکل را می توان بطور کامل با دانستن مقدار متغیرهای آن

تعیین کرد. تعدادی حالت‌های چنین ماشینی 2^n است، که در آن n تعداد بیت‌های لازم برای نمایش تمام ترکیبات ممکنه متغیرهای آن است.

حالت کل سیستم نیز عبارتست از ترکیب حالت‌های دو ماشین پروتکل (فرستنده و گیرنده) و حالت کانال. حالت کانال با محتویات آن تعیین می‌شود. اگر باز هم از پروتکل ۳ کمک بگیریم، کانال ما می‌تواند حالت‌های زیر را داشته باشد: فریم 0 یا 1 از فرستنده به گیرنده می‌رود، فریم تصدیق دریافت از گیرنده به فرستنده برمی‌گردد، و یا کانال خالیست. اگر فرض کنیم گیرنده و فرستنده نیز هر کدام فقط دو حالت دارند، کل سیستم دارای ۱۶ حالت مجزا خواهد بود.

همین جا باید نکته کوچکی را درباره کانال توضیح دهیم. وقتی می‌گوئیم «فریم در کانال است»، البته از یک مفهوم مجرد حرف می‌زنیم. آنچه واقعاً منظور ماست، اینست که «فریم احتمالاً به مقصد رسیده، ولی هنوز پردازش نشده است». این فریم تا زمانی که ماشین پروتکل روال *FromPhysicalLayer* را اجرا نکرده و آنرا پردازش نکند، «در کانال می‌ماند».

هر حالت دارای تعدادی گذار (transition) به حالت‌های دیگر است. گذار زمانی روی می‌دهد که رویدادی رخ دهد. در یک ماشین پروتکل، ارسال یک فریم، دریافت یک فریم، انقضای یک تایمر و مانند آن، همگی نمونه‌هایی از گذار هستند. رویدادهایی که می‌تواند در یک کانال رخ دهد، نیز عبارتند از: وارد شدن یک فریم جدید به کانال توسط ماشین پروتکل، برداشته شدن فریم از کانال توسط ماشین پروتکل، و یا گم شدن فریم در اثر نویز. با در دست داشتن توصیف کاملی از ماشین‌های پروتکل و مشخصات کانال، می‌توان سیستم را بصورت نموداری از گره‌ها (حالت‌ها) و خطوط متصل‌کننده (گذارها) نمایش داد.

یکی از حالت‌های مهم در هر سیستم، حالت اولیه (initial state) است. این حالت متناظر است با وضعیت سیستم در لحظه شروع به کار، یا (اگر مناسبتر باشد) کمی پس از آن. از این حالت اولیه می‌توان به کمک توالی گذارها به تمام (یا برخی از) حالت‌های دیگر رسید. با کمک تکنیک‌های نظریه گراف (graph theory) می‌توان مشخص کرد که کدام حالت‌ها قابل دسترسی اند، و کدامها نیستند. با این تکنیک، که به آنالیز دسترسی (reachability analysis) معروفست (Lin et al., 1987)، می‌توان دریافت که آیا یک پروتکل درست عمل می‌کند یا خیر.

مدل ماشین حالت محدود یک پروتکل را می‌توان بردار چهار عضوی (S, M, I, T) دانست، که در آن:

۱. S عبارتست از مجموعه حالت‌هایی که پروسس‌ها و کانال می‌توانند در آن باشند.
۲. M عبارتست از مجموعه فریم‌هایی که می‌توان روی کانال مبادله کرد.
۳. I عبارتست از مجموعه حالت‌های اولیه پروسس‌ها.
۴. T عبارتست از مجموعه گذارهای بین حالت‌ها.

در لحظه شروع، تمام پروسس‌ها در حالت اولیه‌شان هستند. سپس رویدادها شروع به رخ دادن می‌کنند: فریمی برای ارسال آماده می‌شود، تایمرها خاموش می‌شوند، و مانند آن. هر رویداد می‌تواند باعث شود که یک پروسس یا کانال عملی را انجام داده و به حالت دیگر برود. با تعیین دقیق پیامدهای هر حالت، می‌توان گراف دسترسی را رسم و پروتکل را آنالیز کرد.

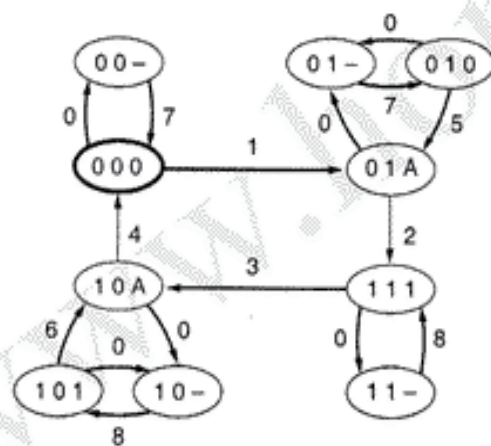
آنالیز دسترسی می‌تواند خطاهای مختلفی را در طراحی پروتکل روشن کند. برای مثال، اگر یک فریم خاص بتواند در حالت خاصی وجود داشته باشد و ماشین حالت محدود نتواند بگوید در این موقعیت چه باید کرد، طراحی ما مشکل دارد (ناقص است). اگر حالت یا حالت‌هایی وجود داشته باشد که نتوان از آن خارج شد (بعبارت دیگر، دریافت فریم سالم دیگر امکانپذیر نباشد)، باز هم پروتکل ما مشکل دارد (بن‌بست). موقعیت دیگری (که

در واقع مشکل چندان بزرگی نیست) آنست که پروتکل ما برای حالتی تجهیز شده باشد که امکان زوی دادن آنها وجود ندارد (گذارهای نامربوط). گراف دسترسی خطاهای دیگر را هم می تواند کشف کند.

در شکل ۳-۲۱ (الف) نمونه ای از یک مدل ماشین حالت محدود را ملاحظه می کنید. این گراف معادل پروتکل ۳ است، که در بالا توضیح دادیم؛ هر ماشین پروتکل دارای دو حالت، و کانال دارای چهار حالت است. در کل ۱۶ حالت ممکن وجود دارد، که البته از حالت اولیه نمی توان به همه آنها رسید. در شکل این حالتی غیر قابل دسترسی را نشان نداده ایم، و برای سادگی کار از خطاهای جمع تطبیقی (checksum) هم چشم پوشیده ایم.

هر حالت با سه حرف SRC مشخص می شود، که در آن S یا 0 است یا 1 (متناظر با فریمی که فرستنده می خواهد بفرستد)؛ R نیز یا 0 است یا 1 (متناظر با فریمی که گیرنده منتظر دریافت آن است)؛ و C می تواند چهار مقدار 0 (فریم 0)، 1 (فریم 1)، A (فریم تصدیق دریافت) یا خالی (-) بگیرد (متناظر با حالتی چهارگانه کانال). در مثال بالا، حالت اولیه با (000) نشان داده شده است: یعنی فرستنده فریم 0 را فرستاده، گیرنده منتظر دریافت فریم 0 است، و فریم 0 اکنون در کانال است.

در شکل ۳-۲۱ که حالت گذار مختلف نشان داده شده است. در گذار 0 کانال محتویات خود را از دست می دهد. در گذار 1 کانال محتویات خود را به گیرنده تحویل می دهد، و گیرنده نیز حالت خود را به 1 (انتظار برای فریم 1) تغییر داده و یک فریم تصدیق دریافت به فرستنده پس می فرستد. گذار 1 همچنین متناظرست با تحویل بسته 0 به لایه شبکه در گیرنده. گذارهای دیگر را در شکل ۳-۲۱ (ب) ملاحظه می کنید. رسیدن فریمی با خطای جمع تطبیقی در اینجا نشان داده نشده، چون این اتفاق در پروتکل ۳ باعث تغییر حالت نمی شود.



(الف)

به لایه شبکه	ارسال شده	فریم پذیرفته شده	فریم نوبت کیست ؟	انتقال
-	(فریم گم شده)	-	-	0
Yes	A	0	R	1
-	1	A	S	2
Yes	A	1	R	3
-	0	A	S	4
No	A	0	R	5
No	A	1	R	6
-	0	(انقضای زمان)	S	7
-	1	(انقضای زمان)	S	8

(ب)

شکل ۳-۲۱. (الف) دیاگرام حالت پروتکل ۳. (ب) گذارها.

در حالت عادی، گذارهای 1، 2، 3 و 4 پشت سرهم و بارها و بارها تکرار می شوند. در هر سیکل دو بسته منتقل می شود، و با این کار فرستنده دوباره به حالت اولیه (ارسال فریم 0) برمی گردد. اگر فریم 0 در کانال از بین برود، سیستم از حالت (000) به حالت (00-) می رود (گذار 0). پس از مدتی تایمر فرستنده به انتها رسیده (گذار 7)، و سیستم به حالت (000) بازمی گردد. از بین رفتن فریم A (تصدیق دریافت) پیچیده تر است، و برای جبران آن به دو گذار نیاز داریم: 7 و 5، یا 6 و 8.

یکی از ویژگیهایی که یک پروتکل با شماره های ترتیبی 1-بیتی باید داشته باشد اینست که هرگز نباید دو فریم فرد متوالی (یا دو فریم زوج متوالی) به گیرنده برسد. در شکل ۳-۲۱ می توان این ویژگی را چنین نشان داد: هیچ

مسیری از حالت اولیه وجود ندارد که طی آن دو گذار متوالی 1 رخ دهد، بدون اینکه بین آنها یک گذار 3 وجود داشته باشد، و بالعکس. از شکل می توان دید که پروتکل 3 این ویژگی را دارد.

الزام دیگر اینست که هیچ مسیری نباید وجود داشته باشد که طی آن فرستنده دو بار تغییر حالت دهد (از 0 به 1 و برگشت دوباره به 0) در حالیکه گیرنده ثابت مانده است. اگر چنین مسیری وجود داشته باشد، بمعنای آن است که دو فریم از بین رفته بدون اینکه گیرنده متوجه شده باشد.

ویژگی مهمتر یک پروتکل عدم وجود بن بست (deadlock) در آن است. بن بست حالتی است که پروتکل (تحت هیچ شرایطی) دیگر قادر به جلو رفتن نباشد (یعنی نتواند بسته ها را لایه شبکه تحویل دهد). در مدل گراف، بن بست به زیرمجموعه ای از حالتها گفته می شود که بتوان از شرایط اولیه به آن رسید و دو ویژگی زیر را داشته باشد:

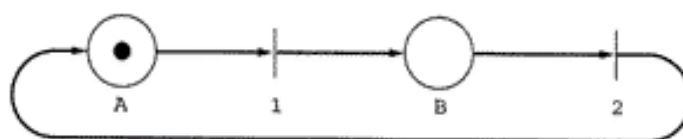
۱. هیچ گذاری برای خروج از این زیرمجموعه وجود نداشته باشد.

۲. هیچ گذاری در داخل زیرمجموعه وجود نداشته باشد که باعث پیشرفت کار شود.

وقتی یک پروتکل وارد بن بست شود، دیگر برای همیشه آنجا می ماند. باز هم از گراف شکل ۳-۲۱ می توان دید که (خوشبختانه) پروتکل 3 هیچ بن بستنی ندارد.

۲-۵-۳ مدل شبکه پتری

ماشین حالت محدود تنها مدل برای ارزیابی پروتکلها نیست. در این قسمت به بررسی تکنیکی کاملاً متفاوت بنام شبکه پتری (Petri net) می پردازیم (Danthine, 1980). یک شبکه پتری دارای چهار عنصر اساسی است: مکان (place)، گذار (transition)، کمان (arc)، و نشانه (token). مکان حالتیست که سیستم (یا بخشی از آن) می تواند در آن باشد. در شکل ۳-۲۲ یک شبکه پتری را با دو مکان A و B می بینید، که با دایره مشخص شده اند. سیستم در حال حاضر در حالت A قرار دارد، که این موضوع با نشانه (نقطه سیاه) در مکان A مشخص شده است. گذار با یک خط افقی یا عمودی مشخص می شود. هر گذار می تواند دارای تعدادی کمان ورودی (input arc) - که از مکانهای ورودی آن می آیند) و تعدادی کمان خروجی (output arc) - که به مکانهای خروجی آن می روند) باشد. گذار فعال به گذاری گفته می شود که حداقل یک نشانه در یکی از ورودیهای آن وجود داشته باشد. گذار فعال می تواند هر گاه که اراده کند، آتش (fire) کرده و یک نشانه را از یکی از ورودیها برداشته و در یکی از خروجیهای خود قرار دهد. اگر تعداد کمانهای ورودی با کمانهای خروجی مساوی نباشد، نشانه ابقا (conserve) نخواهد شد.

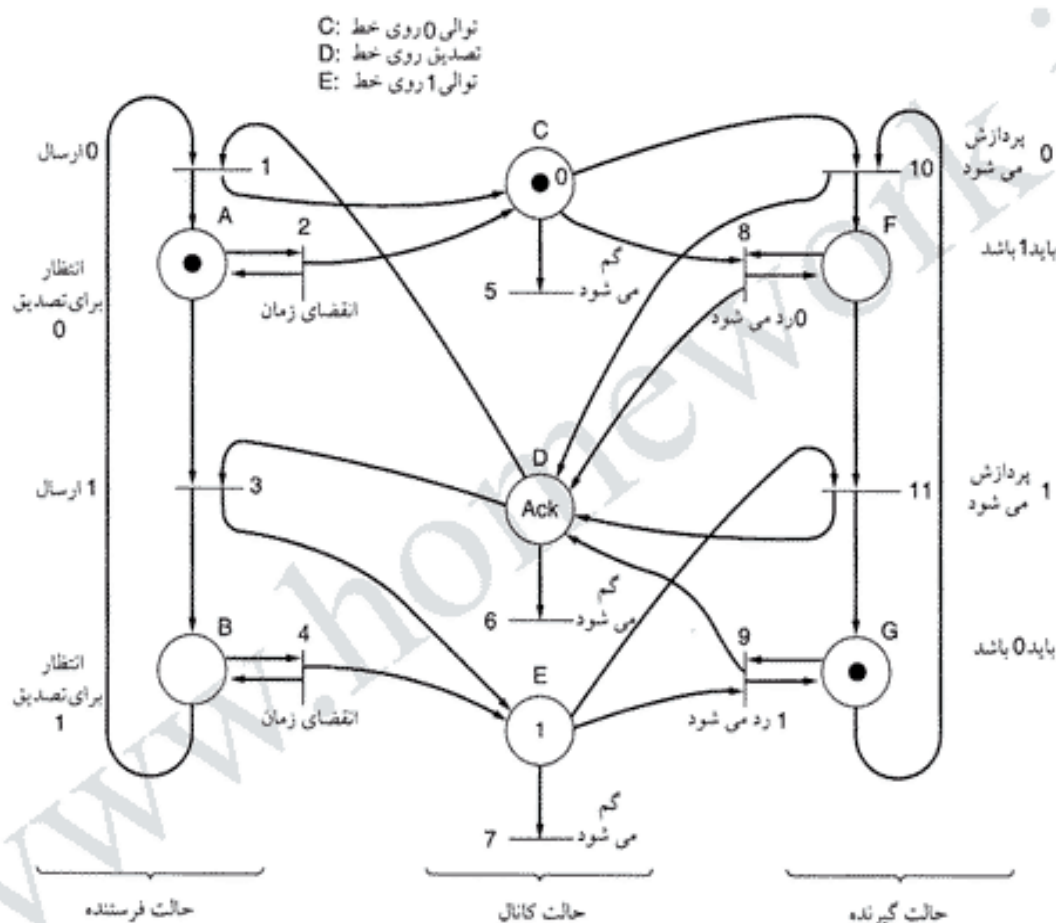


شکل ۳-۲۲. یک شبکه پتری با دو مکان و دو گذار.

اگر دو یا چند گذار فعال وجود داشته باشد، هر کدام از آنها می توانند آتش کنند. اینکه کدام گذار آتش می کند نامشخص است، و همین ویژگیست که شبکه پتری را برای مدلسازی پروتکلها سودمند کرده است. شبکه پتری شکل ۳-۲۲ کاملاً مشخص است و از آن فقط برای مدلسازی پروسههایی که دو فاز بیشتر ندارند، می توان استفاده کرد (مانند رفتار یک نوزاد: خوردن، خوابیدن، خوردن، خوابیدن، و الی آخر). در اینجا هم مانند سایر تکنیکهای مدلسازی جزئیات زائد حذف می شوند.

در شکل ۳-۲۳ مدل شبکه پتری شکل ۳-۱۲ (پروتکل 3) را ملاحظه می کنید. بر خلاف مدل ماشین حالت

محدود، در اینجا حالت های ترکیبی وجود ندارد؛ حالت فرستنده، گیرنده و کانال بطور مجزا و مستقل نمایش داده می شوند. گذارهای 1 و 2 بترتیب عبارتند از ارسال فریم 0 توسط فرستنده در حالت عادی، و بعد از انقضای تایمر. گذارهای 3 و 4 متناظر با ارسال فریم 1 در این دو موقعیت هستند. گذارهای 5، 6 و 7 نیز بترتیب از بین رفتن فریمهای 0، ACK، 1 را نشان می دهند. گذارهای 8 و 9 نشان می دهند که فریمی با شماره ترتیبی اشتباه (بترتیب 0 و 1) به گیرنده رسیده است. گذارهای 10 و 11 نیز بترتیب حاکی از رسیدن صحیح و سالم فریمهای 0 و 1 به گیرنده، و تحویل آنها به لایه شبکه هستند.



شکل ۳-۲۳. مدل شبکه پتری پروتکل ۳.

با شبکه پتری نیز می توان مانند ماشین حالت محدود مشکلات یک پروتکل را تشخیص داد. برای مثال، اگر یک توالی آتش وجود داشته باشد که در آن دو گذار 10 (بدون یک گذار 11 بین آنها) رخ دهد، پروتکل ما مشکل دارد. مفهوم بن بست در شبکه پتری نیز کاملاً شبیه همین مفهوم در ماشین حالت محدود است.

شبکه پتری را بصورت جبری نیز می توان نوشت: هر گذار یک جمله جبری است، که در یک دستور (که مکانهای ورودی و خروجی گذار را مشخص می کند) بکار می رود. از آنجائیکه شکل ۳-۲۳ دارای ۱۱ گذار است، برای نمایش جبری آن به ۱۱ دستور (که آنها را متناظر با گذارها شماره گذاری می کنیم) نیاز داریم. معادل جبری شبکه پتری شکل ۳-۲۳ چنین است:

- 1: $BD \rightarrow AC$
- 2: $A \rightarrow A$
- 3: $AD \rightarrow BE$
- 4: $B \rightarrow B$
- 5: $C \rightarrow$
- 6: $D \rightarrow$
- 7: $E \rightarrow$
- 8: $CF \rightarrow DF$
- 9: $EG \rightarrow DG$
- 10: $CG \rightarrow DF$
- 11: $EF \rightarrow DG$

همانطور که می بینید، یک پروتکل نسبتاً پیچیده به ۱۱ جمله جبری ساده تبدیل شده که به آسانی می توان آنرا با کامپیوتر تحلیل کرد.

حالت فعلی شبکه پتری با مجموعه نامنظمی از مکانها (که هر مکان به تعداد نشانه هایی که دارد، ظاهر می شود) نشان داده می شود. در هر دستور، مکانهایی که سمت چپ جمله قرار دارند می توانند آتش کرده، و (بعد از حذف خود از حالت فعلی) خروجی خود را به حالت سیستم اضافه کنند. شکل ۳-۲۳ را می توان با علامت ACG نشان داد (یعنی، مکانهای A ، C و G هر کدام یک نشانه دارند). در نتیجه، دستورات ۲، ۵ و ۱۰ فعال هستند، و هر کدام از آنها می توانند اجرا شوند و حالت سیستم را عوض کنند (که البته حالت قبلی هم جزء حالتها می باشد). توجه داشته باشید که، برای مثال، در این لحظه دستور ۳ ($AD \rightarrow BE$) نمی تواند اجرا شود، چون D در علامت سیستم وجود ندارد.

۶-۳ چند نمونه از پروتکل های لینک داده

در این قسمت چند نمونه از پروتکل های لینک داده را که کاربرد وسیعی دارند، مورد بررسی قرار خواهیم داد. اولین آنها، HDLC، یک پروتکل بیت-گرا (bit-oriented) است که سالهاست در برنامه های بسیاری از ویرایشهای مختلف آن استفاده می شود. دوم، PPP، یک پروتکل لینک داده است که برای اتصال کامپیوترهای خانگی به اینترنت بکار می رود.

۶-۳-۱ HDLC - کنترل سطح بالای لینک داده

در این قسمت گروهی از پروتکل های نزدیک به هم را بررسی می کنیم که با وجود قدیمی بودن، همچنان کاربرد گسترده ای دارند. همه این پروتکلها از اولین پروتکل لینک داده که برای کامپیوترهای بزرگ IBM توسعه داده شد، مشتق شده اند: پروتکل SDLC (کنترل لینک داده سنکرون - Synchronous Data Link Control). بعد از توسعه این پروتکل، IBM آنرا برای پذیرش بعنوان استاندارد آمریکایی و بین المللی به ANSI و ISO فرستاد. ANSI و ISO هر کدام تغییراتی در این پروتکل دادند، و بترتیب پروتکل های ADCCP (روش پیشرفته کنترل مخابرات داده - Advanced Data Communication Control Procedure) و HDLC (کنترل سطح بالای لینک داده - High-Level Data Link Control) را از آن مشتق کردند. بعد از مدتی، CCITT تغییراتی در پروتکل HDLC داده، و پروتکل جدید را که LAP (روش دسترسی لینک - Link Access Procedure) نام گرفت، بعنوان قسمتی از استاندارد شبکه های X.25 معرفی کرد (این پروتکل بعدها برای سازگاری بهتر با ویرایشهای جدید HDLC باز هم اصلاح و LAPB نامیده شد). خوبی استانداردها همین تنوع زیاد آنهاست، و بالاخره می توانید یکی را مطابق سلیقه تان پیدا کنید؛ اگر هم پیدا نکردید، زیاد جای نگرانی نیست: سال آینده

مدلهای جدیدتری به بازار خواهد آمد!

تمام این پروتکلها مبنای واحدی دارند: همه آنها بیت-گرا هستند، و از تکنیکهای لاگذاری بیت (bit stuffing) برای افزونگی داده‌ها استفاده می‌کنند. اختلاف آنها کوچک (ولی بهر حال، ناراحت‌کننده) است. برای اطلاع از مشخصات دقیق هر پروتکل می‌توانید به منابع مربوطه مراجعه کنید.

تمام پروتکلهای بیت-گرا از فریمهایی با ساختار شکل ۳-۲۴ استفاده می‌کنند. هر فریم با یک توالی پرچم (01111110) شروع می‌شود. فیلد آدرس (Address) در خطوطی اهمیت می‌یابد که ترمینالهای متعددی دارند، و از این فیلد برای مشخص کردن ترمینال مقصد استفاده می‌شود. در خطوط نقطه-به-نقطه (point-to-point) گاهی از این فیلد برای تشخیص فرمان (command) از پاسخ (response) استفاده می‌شود.

فیلد کنترل (Control) برای شماره ترتیبی فریم، تصدیق دریافت، و مقاصد دیگر بکار می‌رود (در ادامه این فیلد را بیشتر توضیح خواهیم داد).

Bits بیت‌ها	8	8	8	≥ 0	16	8
	0 1 1 1 1 1 1 0	Address (آدرس)	Control (کنترل)	Data (داده)	Checksum (جمع تطبیقی)	0 1 1 1 1 1 1 0

شکل ۳-۲۴. فرمت فریم در پروتکلهای بیت-گرا.

فیلد داده (Data) محتوی اطلاعاتیست که فریم باید منتقل کند. طول این فیلد می‌تواند هر اندازه‌ای باشد، اگر چه با زیاد شدن آن کارایی تکنیک جمع تطبیقی (بدلیل بالا رفتن احتمال بروز خطاهای فورانی) کاهش خواهد یافت.

فیلد جمع تطبیقی (Checksum) یک کُد افزونگی چرخه‌ای (cyclic redundancy) است، که در قسمت ۲-۲-۳ توضیح دادیم.

در انتها، فریم به یک توالی پرچم دیگر (01111110) ختم می‌شود. وقتی یک خط نقطه-به-نقطه بیکار است، بطور پیوسته توالیهای پرچم را ارسال می‌کند. هر فریم باید حداقل سه فیلد (مجموعاً ۳۲ بیت) داشته باشد (البته منهای پرچمهای ابتدا و انتها).

فریمها بر سه نوعند: اطلاعاتی (Information)، سرپرستی (Supervisory)، و بدون شماره (Unnumbered). در شکل ۳-۲۵ فیلد کنترل هر یک از این سه نوع فریم را ملاحظه می‌کنید. این پروتکل از تکنیک پنجره لغزنده، با شماره‌های ترتیبی ۳-بیتی، استفاده می‌کند. در هر لحظه تا هفت فریم تصدیق نشده می‌تواند در بافر فرستنده وجود داشته باشد. فیلد Seq در شکل ۳-۲۵ (الف) شماره ترتیبی فریم را نشان می‌دهد. فیلد Next نیز فیلد سواری مجانی برای تصدیق دریافت است. با این حال، در تمام انواع پروتکلهای HDLC مرسوم است که بجای سوار کردن شماره آخرین فریم دریافت شده، شماره اولین فریمی که هنوز دریافت نشده (فریمی که گیرنده منتظر آن است) برگردانده شود. این دو روش هیچ مزیتی بر یکدیگر ندارند، و انتخاب یکی از آنها به طراح پروتکل بستگی دارد؛ البته مشروط باینکه همواره از یک روش استفاده کند.

فیلد P/F مخفف Poll/Final (سرکشی/پایان) است. این فیلد در کامپیوترها یا مودمهایی بکار می‌رود که به چندین ترمینال سرکشی (polling) می‌کنند. اگر این فیلد محتوی P باشد، کامپیوتر (یا مودم) ترمینال را دعوت به ارسال داده می‌کند. در تمام فریمهایی که ترمینال می‌فرستد، فیلد P/F مقدار P دارد (بجز در آخرین فریم، که مقدار آن F است).

در برخی از پروتکلها بیت P/F باعث می‌شود تا ماشین طرف مقابل بلافاصله فریم سرپرستی را بفرستد، و منتظر سواری مجانی نشود. در ارتباطاتی که از فریمهای بدون شماره سود می‌برند، نیز این بیت کاربرد دارد.

Bits	1	3	1	3
(الف)	0	Seq (توالی)	P/F	Next (بعدی)
(ب)	1	0	Type (نوع)	P/F
(ج)	1	1	Type (نوع)	Modifier (تغییر دهنده)

شکل ۳-۲۵. فیلد کنترل در یک فریم اطلاعاتی، (ب) فریم سرپرستی، (ج) فریم بدون شماره.

انواع مختلف فریمهای سرپرستی با فیلد *Type* مشخص می شود. نوع 0 فریم تصدیق دریافت (که رسماً RECEIVE READY نامیده می شود) است، و مشخص می کند که گیرنده آماده دریافت فریم بعدیست. این فریم وقتی ارسال می شود که ترافیک جهت برگشت برای اجرای تکنیک سواری مجانی وجود نداشته باشد. نوع 1 فریم تصدیق دریافت منفی (که رسماً REJECT نامیده می شود) است، از آن برای تصدیق دریافت فریم با خطا استفاده می شود. در این حالت فیلد *Next* حاوی اولین فریمیست که درست دریافت نشده است (یعنی فریمی که باید دوباره ارسال شود). در اینجا فرستنده باید تمام فریمهای بعد از فریم *Next* را مجدداً بفرستد، و از این نظر شبیه پروتکل 5 است تا پروتکل 6.

نوع 2 فریم RECEIVE NOT READY است؛ این فریم اعلام می کند که تمام فریمهای قبل از *Next* درست دریافت شده اند، ولی خود *Next* خیر (و از این نظر شبیه RECEIVE READY است، با این تفاوت که جلوی ادامه ارسال را می گیرد). فریم RECEIVE NOT READY در واقع نوعی اعلام مشکل است از سوی گیرنده، مثلاً مشکل پُر شدن بافرها. بعد از برطرف شدن این وضعیت، گیرنده یکی از فریمهای کنترلی دیگر (مانند RECEIVE READY یا REJECT) را می فرستد.

نوع 3 فریم SELECTIVE REJECT است. این فریم فقط زمانی فرستاده می شود که گیرنده خواستار ارسال مجدد یک فریم خاص باشد؛ HDLC از این نظر شبیه پروتکل 6 است تا پروتکل 5، و برای مواقعی مفید است که اندازه پنجره ارسال نصف شماره ترتیبی باشد. بنابراین اگر گیرنده بخواهد فریمهای نامنظم را بافر کرده و فقط ارسال مجدد برخی از فریمهای معیوب را طلب کند، می تواند از SELECTIVE REJECT استفاده کند. این فریم کنترلی فقط در HDLC و ADCCP وجود دارد، ولی در SDLC و LAPB تعریف نشده است.

سومین نوع از فریمها، فریمهای بدون شماره (Unnumbered) است. این نوع گاهی کاربردهای کنترلی دارد، ولی در سرویسهای غیر قابل اعتماد غیر متصل نیز می توان از آن برای انتقال داده استفاده کرد. برخلاف دو نوع دیگر، این نوع فریم عملکردهای متفاوتی در انواع پروتکلهای بیت-گرا است. در این حالت پنج بیت تعیین کارکرد فریم وجود دارد، ولی تمام ۳۲ حالت آن استفاده نمی شود.

در تمام پروتکلها فرمانی وجود دارد بنام DISC (قطع ارتباط) که اجازه می دهد یک ماشین خاموش شدن خود را به ماشینهای دیگر اعلام کند. فرمان دیگری وجود دارد که به یک ماشین اجازه می دهد تا بازگشت خود را اعلام کرده، و تمام شماره های ترتیبی را به 0 برگرداند؛ این فرمان SNRM (ست شدن حالت پاسخ عادی - Set Normal Response Mode) نام دارد. متأسفانه، «حالت پاسخ عادی» هر چیزی هست جز عادی. این حالت عبارتست از یک رابطه نامتقارر، که در آن یک سر خط «اریاب» و سر دیگر خط «رعیت» است. این فرمان از زمانی

به جای مانده که یک کامپیوتر بزرگ مرکزی وجود داشت و تعداد زیادی ترمینال به آن متصل می‌شد، که در این حالت مسلماً رابطه نامتقارن اریاب و رعیتی «عادی» محسوب می‌شود. برای آن که این پروتکلها بتوانند پاسخگوی نیازهای جدید (رابطه متقارن) باشند، در HDLC و LAPB فرمان دیگری وجود دارد بنام SABM (ست شدن حالت آسنکرون متعادل - Set Asynchronous Balanced Mode)، که خط را ریست کرده و دو سر آنرا به حالت هم‌ارز و متعادل درمی‌آورد. در این پروتکلها دو فرمان دیگر وجود دارد بنامهای SABME و SNRME که بر ترتیب با SABM و SNRM مناظر هستند، و فقط در آنها شمارهٔ ترتیبی فریمها (بجای ۳-بیتی) ۷-بیتی است. فرمان سومی که در تمام پروتکلها وجود دارد، FRMR (FRaMe Reject) است که نشان می‌دهد جمع تطبیقی فریم صحیح است ولی از نظر شکلی درست نیست. از میان چنین شکلهای نادرستی می‌توان به فریم سرپرستی نوع 3 در LAPB، فریمی که کوتاهتر از ۳۲ بیت است، و تصدیق دریافت فریمی که خارج از پنجرهٔ دریافت است، اشاره کرد. در فریم FRMR یک فیلد دادهٔ ۲۴ بیتی وجود دارد که علت خطا را توضیح می‌دهد. اطلاعاتی که در این فیلد مخابره می‌شود، عبارتند از: فیلد کنترل فریم معیوب، پارامترهای پنجرهٔ دریافت، و اطلاعاتی دربارهٔ طبیعت خطا.

احتمال خراب یا گم شدن فریمهای تصدیق دریافت نیز مانند سایر فریمها وجود دارد، بنابراین برای آنها نیز باید نوعی تصدیق دریافت پس فرستاد. برای این منظور فریم کنترلی خاصی بنام UA (تصدیق دریافت بدون شماره - Unnumbered Acknowledgement) در نظر گرفته شده است. از آنجائیکه همیشه فقط یک فریم تصدیق دریافت تصدیق نشده در گیرنده وجود دارد، هیچ ابهامی وجود ندارد که منظور از یک فریم UA کدام فریم تصدیق دریافت است.

سایر فریمهای کنترلی کارهایی مانند آماده‌سازی (initialization)، سرکشی (polling) و گزارش وضعیت (status report) انجام می‌دهند. فریم کنترلی دیگری نیز وجود دارد بنام UI (Unnumbered Information)، که می‌توان با آن اطلاعات دلخواه ارسال کرد؛ این اطلاعات برای لایهٔ پیوند داده هستند، و به لایهٔ شبکه داده نمی‌شوند.

پروتکل HDLC علیرغم کاربرد وسیع آن، بهیچوجه کامل نیست. در (Fiorini et al., 1994) می‌توانید بحثی دربارهٔ مشکلات این پروتکل را ببینید.

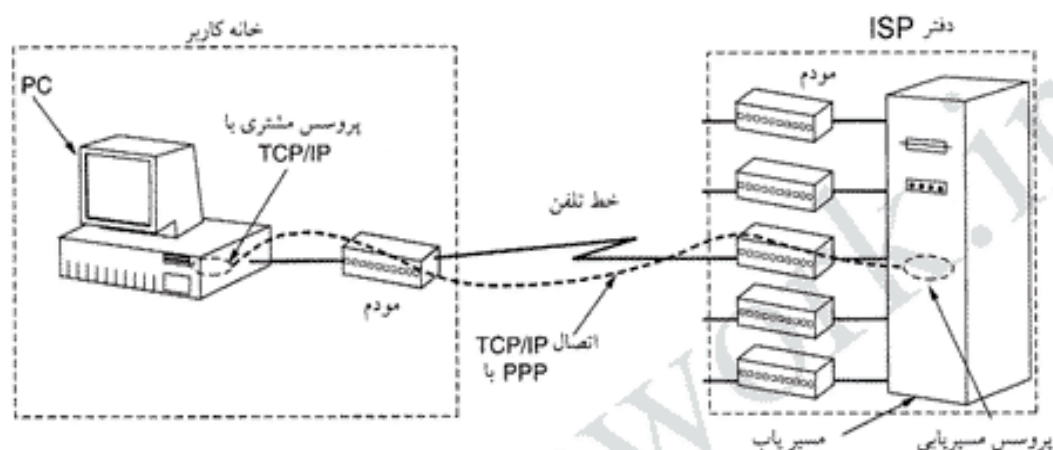
۲-۶-۳ لایهٔ پیوند داده در اینترنت

اینترنت شامل ماشینهای متعددی (میزبان و مسیریاب) است، که توسط یک ستون فقرات به یکدیگر متصل می‌شوند. در یک ساختمان کوچک می‌توان از تکنیکهای LAN برای ارتباط استفاده کرد، ولی در اینترنت اغلب ارتباطات از نوع نقطه-به-نقطه (point-to-point) است. در فصل ۴ دربارهٔ لایهٔ پیوند داده در LAN صحبت خواهیم کرد؛ در این قسمت به لایهٔ پیوند داده در خطوط نقطه-به-نقطه می‌پردازیم.

در عمل، ارتباط نقطه-به-نقطه در دو حالت بکار برده می‌شود. اول، هزاران شرکت و مؤسسه دارای شبکه‌های محلی با تعداد زیادی ماشین میزبان (کامپیوترهای رومیزی، ایستگاههای کاری، کامپیوترهای سرور و غیره) و یک مسیریاب (یا یک پل - bridge) که در عمل همان وظیفه را انجام می‌دهد) هستند، و این مسیریابها در یک شبکهٔ بزرگتر بیکدیگر متصل شده‌اند. معمولاً این مسیریابها بوسیلهٔ ارتباط نقطه-به-نقطه و از طریق خطوط اجاره‌ای (leased line) به یک (یا دو) مسیریاب دیگر متصل می‌شوند. همین مسیریابها و خطوط اجاره‌ای هستند که زیرشبکهٔ اینترنت را می‌سازند.

دومین حالتی که ارتباط نقطه-به-نقطه بکار برده می‌شود، میلیونها کاربر اینترنتی هستند که از منزل و با یک مودم (از طریق خط تلفن) به اینترنت متصل می‌شوند. اتفاقی که معمولاً می‌افتد اینست که کامپیوتر کاربر به

مسیریاب سرویس دهنده اینترنت (ISP) زنگ می‌زند، و از آن طریق (درست مثل یک میزبان معمولی) به اینترنت متصل می‌شود. در این روش فرقی نمی‌کند که خط تلفن معمولی است یا اجاره‌ای، فقط بعد از اینکه کاربر دیگر نیازی به آن نداشت، ارتباط قطع می‌شود. در شکل ۳-۲۶ این نوع ارتباط نقطه-به-نقطه را ملاحظه می‌کنید. مودمی که در این شکل نشان داده‌ایم یک مودم خارجی است، اما مودمهای داخلی نیز دقیقاً همان کار را انجام می‌دهند.



شکل ۳-۲۶. یک کامپیوتر خانگی می‌تواند نقش میزبان اینترنت را بازی کند.

در هر دو حالت (ارتباط مسیریاب-مسیریاب یا میزبان-مسیریاب) به یک پروتکل لینک داده نقطه-به-نقطه نیاز داریم تا وظایفی از قبیل فریم‌بندی، کنترل خطا و مانند آن را انجام دهد. پروتکلی که در اینترنت از آن استفاده می‌شود (و در این قسمت آنرا بررسی خواهیم کرد)، PPP نام دارد.

PPP - پروتکل نقطه-به-نقطه

اینترنت در موارد مختلفی، از قبیل ترافیک مسیریاب-به-مسیریاب یا ترافیک کاربر-به-ISP، به پروتکل نقطه-به-نقطه نیاز دارد. این پروتکل PPP (پروتکل نقطه-به-نقطه - Point-to-Point Protocol) نام دارد، که در RFC 1661 تعریف شده و در چند RFC دیگر (از قبیل RFC 1662 و RFC 1663) مشخصات آن بهبود یافته است. PPP ویژگیهای کنترل خطا دارد، از پروتکل‌های مختلف پشتیبانی می‌کند، اجازه می‌دهد تا آدرس IP در زمان اتصال به طرف مقابل درخواست شود، تعیین هویت (authentication) انجام می‌دهد، و دهها ویژگی دیگر. مهمترین بخشهای PPP عبارتند از:

۱. یک روش فریم‌بندی، که ابتدا و انتهای فریمها را بوضوح مشخص می‌کند. فرمت فریم در PPP تشخیص خطا را نیز انجام می‌دهد.
۲. یک پروتکل کنترل لینک برای برقراری ارتباط، تست آن، مذاکره برای سایر گزینه‌ها، و در پایان قطع ارتباط بصورتی آبرومندانه. این پروتکل LCP (پروتکل کنترل لینک - Link Control Protocol) نام دارد. LCP از هر دو ارتباط سنکرون و آسنکرون، و بیت-گرا یا بایت-گرا پشتیبانی می‌کند.
۳. روشی برای ارتباط و مذاکره درباره گزینه‌های لایه شبکه، که از طرف کار این لایه مستقل است. در این روش برای هر نوع لایه شبکه یک NCP (پروتکل کنترل شبکه - Network Control Protocol) وجود دارد.

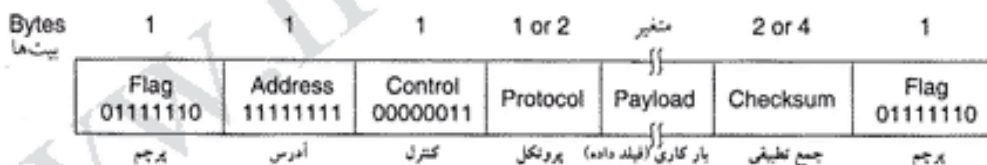
برای اینکه ببینید این قطعات چگونه با یکدیگر جفت می‌شوند، سناریوی ارتباط کاربر-به-ISP را در نظر می‌گیریم. ابتدا PC کاربر از طریق مودم خود به مسیریاب ISP زنگ می‌زند. بعد از اینکه مودم مسیریاب گوشی را

برداشت و ارتباط برقرار شد، PC از طریق فیلد داده یک یا چند فریم PPP چند بسته LCP به مسیریاب می فرستد. پارامترهای PPP از طریق همین بسته ها و پاسخ آنها انتخاب می شود.

بعد از آنکه بر سر این پارامترها توافق حاصل شد، یک سری بسته NCP برای پیکربندی لایه شبکه رد و بدل می شود. معمولاً PC هایی که به اینترنت متصل می شوند از TCP/IP استفاده می کنند، پس PC ماه به یک آدرس IP نیاز دارد. تعداد آدرس های IP آنقدر زیاد نیست که بتوان به همه آدرس ثابت اختصاص داد، به همین دلیل ISP تعدادی آدرس IP را بصورت دینامیک به PC هایی که به آن وصل می شوند، اختصاص می دهد. اگر یک ISP دارای n آدرس IP باشد، می تواند در هر لحظه (حداکثر) n کاربر متصل به اینترنت داشته باشد (البته تعداد کل مشترکان آن می تواند خیلی بیشتر باشد). یکی از بسته های NCP که درخواست IP می کند، این آدرس را به PC اختصاص می دهد.

از این لحظه به بعد PC ما درست مثل یک میزبان معمولی اینترنت است، و می تواند بسته های IP رد و بدل کند. وقتی کاربر ارتباط را قطع کند، NCP نیز ارتباط لایه شبکه را قطع کرده و آدرس IP را آزاد می کند. پس از آن LCP ارتباط لایه پیوند داده را قطع می کند، و کامپیوتر نیز به مودم دستور می دهد که گوشی را بگذارد و به ارتباط فیزیکی پایان دهد.

فرمت فریمهای PPP بسیار شبیه HDLC است (همیشه که نباید چرخ را از نو اختراع کرد). تفاوت اصلی PPP و HDLC اینست که، PPP کاراکتر-گرا (character-oriented) است نه بیت-گرا. بویژه، PPP از تکنیک لاگذاری بایت (byte stuffing) استفاده می کند، بنابراین تعداد بایتهای یک فریم همیشه عددی صحیح است. در PPP (برخلاف HDLC) نمی توان فریمی فرستاد که مثلاً 30.25 بایت داشته باشد. البته PPP می تواند (علاوه بر خطوط تلفن معمولی) روی SONET یا خطوط بیت-گرای HDLC نیز کار کند. فرمت فریم PPP را در شکل ۳-۲۷ ملاحظه می کنید.



شکل ۳-۲۷. فرمت فریم کامل PPP برای حالت بدون شماره.

تمام فریمهای PPP با بایت پرچم استاندارد HDLC (یعنی 01111110)، شروع می شوند، که اگر این بایت در داخل داده ها وجود داشته باشد از تکنیک لاگذاری بایت برای متمایز کردن آن استفاده می شود. بعد از آن فیلد Address می آید، که همیشه 11111111 است و مشخص می کند که تمام گیرنده ها باید این فریم را قبول کنند. با استفاده از این آدرس مشکل تخصیص آدرس های لینک داده نیز هم حل می شود.

بدنبال آدرس فیلد Control قرار دارد، که مقدار پیش فرض آن 00000011 است، و نشان می دهد که فریمها بدون شماره (unnumbered) هستند. بعبارت دیگر، در PPP فریمها شماره ترتیبی ندارند، و از فریم تصدیق دریافت نیز خبری نیست. البته در محیط های پرنویز (مانند لینکهای بیسیم) می توان از فریمهای شماره دار استفاده کرد. جزئیات دقیق این حالت در استاندارد RFC 1663 مشخص شده، ولی در عمل بندرت از آن استفاده می شود.

از آنجائیکه در پیکربندی پیش فرض فیلدهای Address و Control همواره ثابت هستند، LCP مکانیزم خاصی را بین دو سر خط پیاده می کند که این دو بایت بکلی حذف و فریمها کوتاهتر شوند.

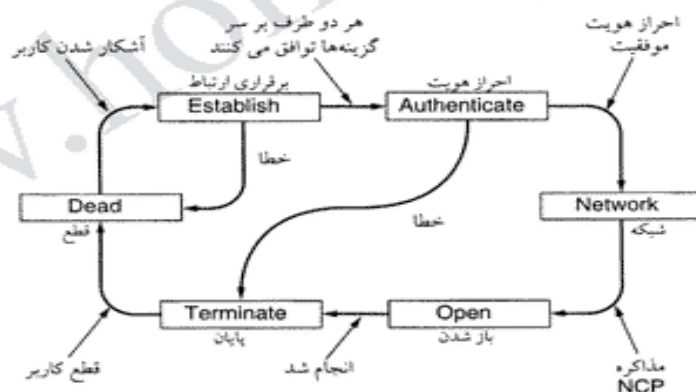
فیلد چهارم PPP فیلد Protocol است، و مشخص می کند که داده موجود در فریم (قسمت Payload) از چه

نوعیست. برای پروتکل‌های مختلف از قبیل LCP، NCP، IP، IPX، Apple Talk و پروتکل‌های دیگر کدهایی تعریف شده است. پروتکل‌هایی که با 0 شروع می‌شوند، پروتکل‌های لایه شبکه (مانند IP، IPX، OSI CLNP و XNS) هستند؛ آنهایی که با 1 شروع می‌شوند، برای مذاکره در باره پروتکل‌های دیگر بکار می‌روند (از جمله LCP، و یک NCP خاص برای هر یک از انواع لایه‌های شبکه). اندازه پیش فرض این فیلد 2 بایت است، ولی دو طرف می‌توانند از طریق LCP مذاکره کرده و اندازه آنرا به 1 بایت تقلیل دهند.

طول فیلد Payload متغیر است، و حداکثر مقدار آن در مذاکره اولیه مشخص می‌شود. اگر دو طرف در مذاکره اولیه (هنگام برقراری ارتباط) بر سر این عدد به توافق نرسند، از عدد پیش فرض 1500 بایت استفاده خواهد شد. اگر مقدار داده ارسالی به این حد نرسد، لایه پیوند داده بقیه را با کاراکتر خاصی پر خواهد کرد. بدنبال Payload فیلد Checksum می‌آید، که مقدار آن معمولاً 2 بایت است، ولی دو طرف می‌توانند بر سر جمع تطبیقی 2 بایتی هم توافق کنند.

بطور خلاصه، PPP یک مکانیزم فریم‌بندی چندپروتکلی است، که می‌توان از آن روی مودم، خطوط بیت-گرای HDLC، SONET و سایر لایه‌های فیزیکی استفاده کرد. این پروتکل از کشف خطا، مذاکره برای گزینه‌های مطلوب، فشرده‌سازی سرآیند (header compression) (و در صورت نیاز، از ارتباط قابل اعتماد یا فرمت HDLC) پشتیبانی می‌کند.

اکنون اجازه دهید ببینیم در PPP برقراری خط و قطع ارتباط چگونه انجام می‌شود. در شکل ۳-۲۸ مراحل ساده شده برقراری خط، بکارگیری، و قطع آن نشان داده شده است. این مراحل برای ارتباط مودمی و یا مسیریاب به مسیریاب هر دو صادق است.



شکل ۳-۲۸. مراحل ساده شده برقراری و قطع خط در پروتکل PPP.

در شروع کار پروتکل خط در وضعیت DEAD است، و این به معنای آنست که هیچگونه کاربر یا ارتباطی در لایه فیزیکی وجود ندارد. بعد از برقراری ارتباط در لایه فیزیکی، خط به وضعیت ESTABLISH می‌رود. در این لحظه مذاکره بر سر گزینه‌های LCP شروع می‌شود، که اگر موفقیت‌آمیز باشد، به وضعیت AUTHENTICATE منجر می‌شود. حال دو طرف می‌توانند در صورت تعادل هویت یکدیگر را چک کنند. بعد از ورود به مرحله NETWORK، لایه شبکه با اجرای پروتکل NCP مناسب پیکربندی می‌شود. اگر این پیکربندی موفقیت‌آمیز