

Chapter 6: Security and the new medium

According to [American] federal law enforcement estimates, online thieves steal more than \$10 billion worth of data in the United States annually. Law enforcement officials admit that, so far, they're fighting a losing battle. -Information Week for Business and Technology Managers'

The only safe computer is a dead computer. – S. M. Lieu. Netsurfer Digest

The stakes associated with security are high. Security is now recognised as one of the Internet's true weaknesses. Advances in protecting systems, software and various transactions are rapid and new products can alleviate most security concerns. However, copyright and privacy are rooted in security.²

Treating these subjects in one chapter illustrates how the medium of the Internet creates the environment for rethinking or redefining some fundamental issues. In the seamless, global, fast-cycle network, generally accepted views of protection, ownership and secure or private information are undergoing change. In many instances, traditional mechanisms of protection can no longer function. The flows of data are too great and the resources needed for effective monitoring are not available. Nevertheless, people with assets want protection or at least assurance that compensation will be paid and privacy assured. The reality of the networked world is not congruent with these expectations.

The fears about security are real. America Online and Microsoft have both been plagued by a new type of security threat. The powerful macro capability of word processing programs has allowed hackers to create a Microsoft Word-delivered 'mail bomb'. A simple electronic message, when opened on the recipient's computer, can destroy the contents of the hard drive or play other mischief. Victims of the America Online mail bomb must open a message and then execute an attached file with the name aolgold.exe or install.exe. The outcome is loss of data. With mail on the most used electronic information services expected to exceed one trillion messages a year by 1997, these security threats are an unpleasant development in the network medium. If one is not safe reading one's own mail, how safe are

-
- [1] Clinton Wilder and Bob Violino, 'Data Security: Online Theft', *Information Week for Business and Technology Managers*, 28 August 1995, p. 30.
- [2] Inevitably, the treatment of copyright and privacy will need to be somewhat abbreviated in these present pages. Readers are referred, however, to the second edition of Charles Oppenheim's excellent companion Briefing to this one, where the two subjects are explored in greater depth and detail: *The Legal and Regulatory Environment for Electronic Information*, second edition, Infonortics, Calne, 1995.

financial transactions, personal medical reports and other types of sensitive information?

No discussion of security techniques and the closely-related topics of copyright and privacy can do justice to the immense financial, social, political and personal issues tethered to the concept *security*. Techniques do exist to provide excellent protection. However, no system can provide protection if an individual with knowledge of the system uses that expertise for nefarious purposes. The same vulnerability exists with any medium or any person-to-person interaction. What is dramatically different is the scale of the consequences. The Internet medium raises the stakes all along the information chain. In short, the stakes have never been higher and the risks never greater.

Secure systems: an oxymoron?

Security, particularly security lapses, makes headlines. A Russian cracker stole \$400,000 from a major American bank, a French student broke Netscape Communications' security system for credit card transactions.

Even Hollywood exploits the Internet. The film *Hackers* exploits the vulnerability of networks to experts with a chip on their shoulder and in their laptop. What did

How secure are different media: a pragmatic view of piracy and theft		
<i>Medium</i>	<i>Security techniques</i>	<i>Effectiveness and vulnerability</i>
Audio CDs	Embedded data	Not effective Low-cost optical drives and software permit duplication
Facsimile publication	Timeliness Limited distribution Fax-on-demand can record telephone number of caller	Effective if published frequently
Live broadcasts or performances	Prohibitions against unauthorised copying	Not effective Can be copied
Motion pictures	Limit access via distribution channel	Not effective Duplication systems permit high-speed copying of protected tapes
Online data	Layers of security	Effective if human factor can be managed Vulnerable to an expert attack
Printed matter	Frequency of updates Number of pages to be copied Language of the original Limiting distribution in some way	Effective if page count is high Vulnerable to digital copying
Software	Copy protection schemes Large program size CD-ROM distribution discs 'Dongles' or hardware locks	Not effective Software is available to defeat external hardware devices

not appear in the advertisements for the film was the revelation that hackers vandalised the Web site set up to promote the film. Talking about security appears to be less useful than achieving it.

Furthermore, like any computer discipline, security has its own peculiar jargon: *public key encryption, authentication, digital signatures, firewalls* and *proxies*, to name a handful. These words describe a range of techniques that can be used to keep information from unwanted alteration, ensure that copyright is enforced, collect money for transactions and keep private information safe from unauthorised eyes.

If the present trend toward digitisation and Internet linkage to data repositories continues, security will remain a pivotal subject for many years. A type of feedback somewhat akin to a tornado is taking place. More users, more information, more knowledge of systems and more payoff for high value information, equal security risks. As soon as one new feature is implemented, a bright programmer will defeat it. The rationale is similar to mountain climbing. The higher the barrier, the greater the climbing thrill, the more significant the achievement. Not surprisingly, the most hacked sites in the world are those which purport to have the best security. In many cases, good security is a result of positioning and planning, not technology. The table on page 94 provides a list of useful sites that contain security information.

With other media, risks were ever present, but they have been controllable. Printing presses and photocopy machines can be locked up or destroyed. Broadcast licences can be revoked or transmitters disabled. The Internet medium does not lend itself to local or even national controls. The individual and a global, interactive communications system interact in distinctive, new and remarkable ways.

The technology and the social fabric of the global Internet community are in the process of redefining such concepts as copyright, privacy-and even security itself.

Security breaches

Security is a nebulous concept. Systems that can be used in authorised ways also permit a wide spectrum of unauthorised behaviour:

- High school students built a dozen bombs and stashed them in their school lockers. School officials said: "It's very disturbing that this type of information is out there" [on the Internet.]
- Employees simply steal information and then sell it; a telephone company employee admitted selling more than 50,000 credit and telephone card numbers. Stolen credit card numbers are available on 'black' USENET discussion groups and private bulletin board systems.' Using anonymous mailers, data thieves advertise stolen or black data on USENET groups.

[1] A *black* discussion group is one concerned with illegal or illicit information.

Sites to monitor for security-related information		
<i>Site</i>	<i>Address</i>	<i>Comment</i>
Authentication Information	http://www.w3.org	Details of a multi-user, multilevel authentication process
CERN	http://www.w3.org/hypertext/www	Security information about CERN's Hypertext Transfer Protocol (HTTP) server
Direct Access Authentication Scheme	http://www.spyglass.com	Authentication process proposal
General Security Information	http://www-ns.rutgers.edu	Information and links to security sites
Internet Security Background	http://www.home.mcom.com	Description of Secure Sockets Layers protocol
Mosaic User Authentication	http://hoohoo.ncsa.uiuc.edu	Security set up for the National Centre for Supercomputer Architecture server

The messages are encrypted with state-of-the-art public keys. Only those in the inner circle know the innocuous message's real content. The basic tool is the encryption algorithm known as Pretty Good Privacy.

- Public domain programs allow hackers to 'spoof' security systems.' Firewall systems usually look at packet headers to determine if the information stream comes from an authorised or unauthorised source. A spoof feeds a legitimate header to the firewall, thus allowing an unauthorised user access to the system. Agents, hidden in the Byzantine structures of UNIX servers, capture log-in names and passwords of legitimate users. After snagging passwords, often hundreds, the agent mails the log-in data to the hacker and erases itself.
- When thwarted, hackers can bombard a site with 'pings', or requests for the server to identify itself with its name and location. Ping bombing overloads the targeted server causing it to fail. Automatic remailer tools, agent software such as tcl (tool control language) and original C programs make short work of the targeted server. Public domain software with names such as *spook*, *hacktick* and *penet* performs these functions.
- Team attacks, often from public access Internet sites in libraries or university computer laboratories, are used to breach highly secure sites. Targets of these attacks are seldom willing to announce that they have been the focal point of concerted hacking. Information about these breaches is circulated by word-of-mouth at Internet security meetings and on hacker news groups.

[1] Spoofing is routinely used to allow staff to access the Internet via proxy servers maintained by their employers. The techniques can be used to breach security systems.

Internet security checklist

- Are security features integrated with the Internet server and any other network?
- Is public key cryptographic technology from RSA Data Security or comparable supplier used?
- Is enhanced user authorisation, including HTTP 1 .0 or higher access authorisation supported?
- Are Internet packet and domain name server access control, local access control, user-controlled passwords and named group capabilities implemented by the security system software?
- Is support for multiple Internet addresses (permits a server to support more than one domain name) provided to authorised users?
- Is there security support for UNIX and Windows NT servers?
- Are user tracking and log analysis tools with real-time monitoring functionality in place and operational?
- Are multiple-user public information directories with variable paths supported?
- Are layers of security supported with a combination of tokens, dongles and authentication?
- Is a firewall in place and maintained by a systems professional and monitored by an oversight service like that offered by Bolt, Beranek and Newman?
- Are such meta-security processes as SSL, authentication, etc. supported?
- Do passwords become invalid on an established cycle?

- A decent, law-abiding person may send to a friend, via the Internet, a commercial software program. The friend might then post the software on a software server. The original sender acted out of friendship. The recipient wanted to share the useful program. Within minutes, the commercial program is replicated and copied worldwide to servers specialising in software. The developer of the popular game Doom said in a radio interview: “We found two Internet sites with the unreleased version of our game on them. Every stolen game is food from my child’s mouth.”

Does this exhaust the tricks, risks and deceptions? No. It does not scratch the surface of what is taking place in the new medium. Can law enforcement and legal institutions take effective action? No. When it comes to technology, these professions are far from the bleeding edge of technology. They are lucky to be able to log on and gather evidence about the most overt abusers. The subtle, clever thieves are ghosts.

The Internet requires a rethinking of certain assumptions about information. Data can be instantly duplicated and sent anywhere in the world in seconds. The most effective hackers are not clever teen-agers. Significant risks are posed by profes-

[1] Interview on National Public Radio, 14 November 1995

sionals who use their knowledge for financial or personal gain and by average computer users who find nothing morally objectionable about copying a program or information for a friend or colleague.

Those without knowledge of the new medium's security weaknesses are digital sheep waiting for their turn on the shearing block

Security basics

The first line of defence for a secure system is to make a decision about what to protect and from whom, A security audit can identify information priorities. The individuals within the organisation who have access to this information often provide the key to unlock the most elaborate systems. Security is only as good as its weakest link. If that link is a person within an organisation, steps must be taken to ensure the right people are on the job.'

The options available range from common sense to elaborate high-technology set-ups worthy of a James Bond film. The common sense steps are the first ones to consider and, truth be told, for many situations they will provide the first line of defence. An organisation or person must define minimum standards for a 'security envelope' and then implement procedures and processes to deliver that level of security.

Because security systems can be implemented at almost any cost and scale, the right tools for the specific task at hand must be selected. The most readily available and widely used tools are:

- Firewalls
- Encryption
- Physical devices (also known as 'dongles') that plug into a user's PC.

First, a firewall is either software or a software/hardware combination that lets authorised users into or out of a system and keeps out all others. Most firewalls require that the organisation define two parameters for each user who is to obtain access via the firewall. The software firewall limits access from the public connection to the protected information on the server. The entire server can be protected by the firewall, or certain files or services can be protected. To accomplish this, the firewall software must know who can gain access, what type of verification is required, and the information on the server to which a person with specific authorisation can gain access.

A typical hardware/software system designed for Internet use is the Sun Microsystems Netra product that retails for about \$6,000. When delivered, the Netra is ready to run, and watches the data packets to make sure that they belong to an

[1] The resource for security information is the National Computer Security Association, 10 South Courthouse Avenue, Carlisle, PA 17013, e-mail firewall@ncsa.com.

authorised user. Once the system is set up, it provides a reasonable degree of security. Most breaches occur because customers lack the resources or the will to maintain the passwords and manage the security systems.

Variations on the firewall theme abound. The commerce servers from Netscape Communications and Open Market include firewall functions along with state monitoring and password functions to protect the data on the server from unauthorised access.

For the majority of cases, properly administered password protection is sufficient. Intrusions or incidents occur mainly because of human error. Many people tape their passwords to their computer monitor, or give them to another person who wants to use their system. Typical systems can be breached when a person copies log-on scripts from PCs or workstations left running, even by a technician who performs routine maintenance.

Serious experts in thievery such as Kevin D. Mitnick, captured in early 1995 by other computer wizards, use sophisticated techniques to breach a system. Programs with sniffing and spoofing functions intercept passwords and responses to prompts. Mr. Mitnick's 'crime' was to penetrate computer systems accessible to authorised persons. Mr. Mitnick's targets ranged from Digital Equipment to the San Diego Supercomputer Centre system (SDSC). Mr. Mitnick copied without paying – that is, stole – the DEC VMS operating system. He compromised the California motor vehicles information system, snagged more than 15,000 valid credit card numbers from an online system's accounting database, and bragged about his accomplishments on online fora. (Mr. Mitnick was captured by a team of American government security experts and Tsutomu Shimomura, then a security expert from the University of California-San Diego Supercomputer Centre).

Second, encryption or 'payload security' offers an additional level of protection. In its simplest form readable text is converted into a jumble of letters that makes no sense. When the jumble is processed with a software key, the original message is displayed. Encryption technology is available in the public domain for little or no charge. One excellent program is PGP (Pretty Good Privacy), written by Philip Zimmerman. Commercial encryption tools are available from numerous vendors. Most of the more robust implementations are based upon technology developed at Stanford and Harvard Universities and commercialised by RSA Data Security.

Encryption technology comes in different degrees of security. The widely-publicised cracking of the Netscape Navigator encryption technology by a student in Europe was a result of a less rigorous encryption technique. A shorter key was used to secure the message. Longer keys, 128 bits or more, are used within the United States, but America's export laws prohibit the sale of the technology outside its borders. Data encrypted with shorter keys can easily be cracked. The longer the key, the less likely the message will be unscrambled.

Encryption is becoming more widely deployed to protect general business information. IBM has introduced the crypto-envelope. The idea is to combine encryption with verification that the sender and the recipient are who they say they are. The

verification process permits a publisher, for example, to send a document to a customer. The information, if intercepted, cannot be read without the proper decryption keys.

Similarly, the major vendors of Internet commerce servers have developed systems that implement encryption and various other checks on the authenticity of the sender and receiver of a message. Netscape has developed the Secure Sockets Layer protocol. It provides the ability to keep the connection between buyer and server private. Netscape has worked to make its SSL protocol the standard by providing the details of the SSL to the Internet Engineering Task Force as an Internet Request for Comment.

CompuServe, Bell Laboratories and others are developing solutions to security issues. If and when a standard emerges, it becomes a powerful competitive weapon. Not surprisingly, Microsoft and one of its partners (Visa International) have developed an alternative Secure Transaction Technology (SST). It has been designed to handle secure financial transactions over insecure transport media such as the Internet. SST incorporates RSA encryption technology. Microsoft has also developed a Private Communication Technology protocol that fits between the network transport and high-level applications such as telnet, ftp and http. Microsoft's approach also permits compatibility with the other standards and supports different keys for message authentication and message encryption. With Microsoft's embracing of the Internet, each of the major vendors of secure servers will support transactions from other secure systems in order to minimise data incompatibilities and customer frustration.

Encryption systems are available from such companies as Cylink Corporation in Silicon Valley, Isolation Systems in Toronto, Raptor Systems in Waltham, Massachusetts and Technical Communications Corporation near Boston.

The outlook for encryption as a security tactic is robust. What is emerging is layers or wrappers of encryptions. Once a hacker bores through one layer, one or more other encryption layers must then be decoded,

The third major category of security techniques is devices. These are cards, plug, cables or combinations of physical devices that are attached to a computer. When the computer is equipped with such a device, encryption, access information and log-on procedures are transmitted to a specific server. When a computer operates without the device, access is not permitted to the secure system.

Each of these broad categories of security techniques rests upon a solid technical foundation. Among the technologies underlying these widely used protection schemes are cryptographic techniques, applied-number theory, authentication and integrity.

The security planning process

One can never be too thin, too rich or too secure. A point of diminishing returns is reached when the cost of providing security outweighs the value of the information or the effort required to get an Internet server operational. When the stakes are high,

can any effort be spared to provide a suitable level of security? What is needed? How much must one spend to have a secure publishing site on the network? Tough questions, indeed.

There are three keys to Internet security. The first is knowledge. The second is appropriate support in terms of money, staff and hardware and software infrastructure. The third, usually the most difficult to control, is people.

What are the elements of a secure system? The table on page 95 provides an expanded checklist. Obviously the type of security and specific mechanisms for ensuring an appropriate level of security depend on situations. The security system can be tailored to a situation when questions such as these are answered:

- What is the purpose of the security system? Is this server designed to permit public access to the data on the server while protecting the server from intentional or accidental harm?
- What is the user's or customer's expectation of security? Do those visiting the site have a tolerance for security precautions? Will users provide passwords, or will users forget passwords and expect some type of on-demand customer support to assist them?
- What particular information or parts of the system must be the most secure? What parts of the system can be comparatively open to outside access?
- What are the physical safeguards that must be taken to provide a suitably secure environment for the customers, employees and visitors?
- What person or organisation will be responsible for the security precautions that are implemented? Is a member of the staff able to handle this responsibility? Should the security monitoring and operation be delegated to a third-party contractor?

Answering these questions provides a baseline of information upon which to build an appropriate security system.

The other dimension of knowledge pertains to the options themselves. Security can be visualised as a ladder. In order to reach the highest rung or the most secure information, more steps must be taken. How does one differentiate between a security system based upon passwords and authentication, versus a system based upon passwords that change on a weekly cycle and require authentication as well as a digital signature? Making distinctions is important because the cost differentials can be significant – and in some instances, sobering.

The best guarantees for secure online services are:

- A system that is tailored to the specific needs of the user or the customer. Complex log-on procedures are often resisted as too cumbersome or complex for the pace of work.

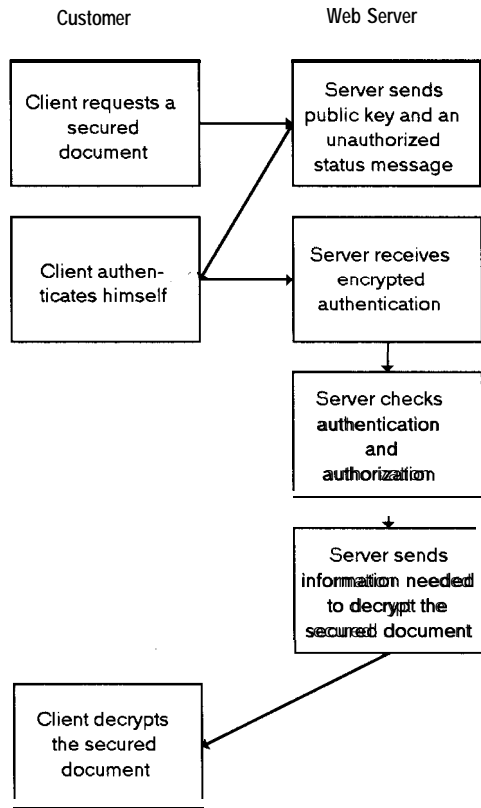
- Anonymity can be a powerful security step. The presence of the server is communicated only to a specific customer group.
- Content is a security feature. Certain types of information cannot easily be copied either because its volume is too great (eg, digital image libraries) or too volatile (frequent updates make it more efficient for users to get the most recent material automatically).
- Monitoring. A system, regardless of levels and types of security imposed, must be watched. When an intrusion occurs, action can be taken ranging from shutting down the server to monitoring the packets of the intruder. The data in the packets allow the intruder's location (in most instances) to be determined.
- Designing the system so that different types of information or levels of service are wrapped in different security layers. Marketing information might reside on a public server with little or no formal security beyond a form requesting visitors to provide their name, address and electronic mail address. More secure information can require a password and an authentication process that is changed regularly.
- Making use of Web servers that are separate from mission critical networks. The outbound Internet link is through a separate communications mechanism; for example, ISDN through the modem pool which supports transient connections. The public Web server is updated with a temporary connection to the organisation's internal network only when specific functions such as updates are processed.
- Using firewall and Web server software that support encryption and secure transactions. Supplement these steps with site monitoring processes.
- Having security as a priority. Staff and a budget are required to maintain a secure environment. The most vulnerable sites are those with security procedures that are not maintained rigorously.

An overview of the building blocks of a secure system appears on page 114.

The future of security

The basic authentication scheme used by the original HyperText Transfer Protocol does not provide a secure method of user authentication. Consequently anyone can masquerade as another person. The recipe for trouble compounds HTTP with such common ingredients as the 'open' approach of UNIX, the ethos of the Internet and the intelligence embedded in the packets of data. The development of standards is difficult.

HTTP does not provide end-to-end protection across the net. The body of an HTTP message is transmitted as clear text across the physical network which is used as the carrier. Some proposed standards for allowing secure transport are Shen, S-HTTP and SSL. Although the technical details of each approach vary, they all make use to some degree of a combination of techniques.



A simplified view of the Identification and Authentication process for an Internet secure Web service

The future for security is 'I&A', or identification and authorisation. I&A operates at different levels of stringency depending upon the specific implementation. A single-level identification requires that specific information be provided to the system. In some implementations, the required information changes daily.

The more robust applications require the use of an external device (dongle) and specific personal identification numbers. A token is generated by the device. Access is possible only when the I&A system generates the code required to establish access. Systems using encryption or a time-based code use public-key and private-key mechanisms to control access to the host. The Virtual Open Network Environment Corp. (V-ONE) approach uses smart cards, precoded data and a PIN (Personal Identification Number) in its system.'

[1] Information about devices used for I&A security applications may be located at Security Dynamics, Inc. <http://www.sdti.com>; Digital Pathways Inc. <http://dp.com>; Cylink Corporation <http://cylink.com>; and V-ONE <http://www.v-one.com>.

A 'challenge-response' mechanism has been developed by Bell Laboratories and is being explored by Mosaic licensee Spyglass Technologies. The client and server agree upon a value derived from a password or other token. It is then the value that is transmitted on the network, not the password itself.

Authorisation technology that is now widely available from server software vendors offers two basic techniques to protect a site:

- Username/password-level access authorisation. The security can be assigned as a single user per password, multiple users per password, or group access per password.
- Rejection or acceptance of connections based on the Internet address of the client with network protection based upon matching packet addresses to a list of authorised points of origin.

There are several levels at which authentication can work. The building blocks of a secure system make use of one or more of these components:

- Rule file that defines which directory trees are public or protected.
- Protection set-up file that spells out the authentication scheme; that is, the process to determine that a packet has not been tampered with.
- Access control specifies the users, groups or domain names that have access rights in a specific directory.
- Password file that contains user names and passwords. This file can be automatically maintained by a tool that is provided with the server software.
- Group file that lists which user names and Internet addresses belong to which groups.

In summary, it is possible to engineer a secure system. Many organisations do not address security with a proper degree of enthusiasm until a break occurs.

Nibbling at privacy: monitoring Internet usage

The proliferation of advertising supported Internet sites has created a demand for monitoring and tracking usage at sites. The majority of Internet sites are stateless; that is, no record is kept of who sends the data that flow in and out of a site.

It is possible to keep track of usage using UNIX utilities tweaked for the site operator's particular interest. A system administrator with fluency in `perl` or `C` can write programs that perform a number of system monitoring functions. The most common monitoring programs track the number of accesses a site has, the amount of data flowing into and out of the system, and various performance parameters.

One of the most misleading statistics collected using public domain site monitoring tools is the number of 'hits' a particular Web site receives in a 24 hour period. The majority of the usage reports count mouse clicks on a hyperlink as a hit. The system is not able to distinguish between one user with an overactive finger, and ten users with one mouse click each.

Until recently, most server software that would track the number of different users accessing a site, the specific Web pages and hyperlinks they followed and the amount of time each user spent within a particular Web page, was unaudited. Sites, such as <http://www.persimmon.com> that touted their ability to track the 'states' – that is, actual number of users and what they did within a particular site – were not audited. But now, monitoring utilities offer site operators the ability to accumulate a wealth of information not previously available.

A representative example of these utilities is the OMI WebReporter. This software, created in Cambridge, Massachusetts, tracks the demographics of the users of a particular Web site. OMI's tools can work at the single server level. A broader deployment of the tools allows one to monitor traffic, pricing and information activities across a group of servers. The canny implementer can even monitor servers that are operated by others. The report generator provides a wide range of tabular summaries of usage, traffic and time.

Nielsen Media Research, a unit of Dun & Bradstreet, introduced its web audit product. The I/COUNT and I/AUDIT software, developed by a third party, give Web site operators an independent, standardised report of usage. The data captured by the Nielsen product include numbers of unique users, specific Web pages accessed, the number of Web pages accessed, and ordering or messaging functions executed. The software generates standardised reports that can be used by the site operator to sell advertising or demonstrate the traffic on a particular site.

The key to the newer monitoring technologies is analysis of the header information on each data packet and requiring that users register and use a password for subsequent visits. Traditional online systems, and an increasing number of Web sites such as IBM's much-publicised Information Marketplace server, require that the user initially provide complete address information, name, password and verification key – typically mother's maiden name, name of a child, or spouse's maiden name. Each subsequent visit to the site requires use of the name and password. With such a system in place, monitoring software can track the actions of each registered user.

However, for sites that permit public access, monitoring software must use the information in the packet header to determine each log-on. The Nielsen software has the capability to inspect the header packet identification information and capture the point of origin for the pack at the sender's electronic mail address. The monitoring software then tracks the activities of this particular user and maintains a log of user activity. Because header identification information is unique, the Nielsen approach provides a reliable way to track actual site usage. Hits can be defined precisely. With robust tracking software in place, sites that report hundreds of thousands of hits find that they have only thousands of users. Even sites that have

been characterised as receiving millions of hits per day are finding that traffic in the 400,000 to 500,000 unique log-ons a day more realistic. The number of absolute log-ons is increasing at a rapid rate, but the numbers are an order of magnitude or more below what had been reported before the advent of standardised site monitoring software.

The benefits of site monitoring are considerable. Site operators have more information about what information services are attracting users. Network planning benefits from data about when data are accessed and what data are generating the greatest system load. Advertisers can have greater confidence in the data upon which rates are based at sites that sell advertising space.

The downside of monitoring is the encroachment of monitoring usage upon privacy. The information contained in the data packet header provides enough information for a site operator to build a mailing list of visitors. The data captured by voluntary log-on provides a goldmine of customer information. These data can be used for communication with interested customers. Other site operators may elect to provide the data to third parties. Resale of subscriber names is a common practice in the magazine and direct mail communities. It is likely that Internet usage data will spawn a similar business.

The American Direct Mail Marketing Association members give magazine subscribers an option to have their name used or not used in lists for resale. No similar provision is in place for the Internet, nor is there likely to be in the foreseeable future. The Internet will become a rich source of marketing information about particular e-mail users.

Thus, the success of the various directory services that have appeared from large companies such as NYNEX to small start-up operations such as McKinley Data in Berkeley, California, underlines the need for pointers or standard index services. The transactional services such as those introduced by the United Nations and dozens of others have a potential vulnerability. Consider a user who sends an electronic mail message to several vendors of a similar product. The vendors respond with product information, a request for more information and possibly a specific price or price range for the product.

The vendors assume that the person making the request is sincere. However, what if the sender of the message were a vendor trying to determine what his competitors' prices are? What if the sender were a consumer advocate preparing a commercial report on the pricing practices of vendors of a similar product in a specific region?

The example illustrates that unless considerable thought is given to the type of information provided in the Internet environment and the vetting procedures required for what appear to be harmless inquiries, the new medium's environment can be used in surprising ways.

Now consider the publishing company wanting to create an interactive directory. Advertisers can post their own electronic brochures or point to pre-existing Web sites. The user consults the directory and browses listings and added-value informational attachments. Well and good until one finds that a hot link from an

innocuous listing drops the user into a site that is perceived as unsavoury by the user. The user can either exit the site or assume that other 'problem' locations are hosted by the well-intentioned service.

But such troubles are too obvious and miss the key point. Once the person has visited the site and provided information directly, or the site operator has made use of site monitoring tools, the visitor can begin receiving messages or solicitations from companies making use of this customer or visitor 'mailing list'. The rapid advances in site monitoring technology and the characteristic of headers to contain substantial amounts of information about the origin of a message allow technologically-adept site operators to build lists of prospects.

Site monitoring advances have now broadened the privacy issue, therefore, beyond that of a hacker who looks at financial or legal records. Without the knowledge of the Internet site visitor, usage data are compiled. Those data have commercial value as a list of qualified prospects. Many sites, including the Point Communications site profiled in the Pacesetters chapter of this Briefing, provide upon request automatic electronic mail notification of new sites that are likely to be of interest to the customer. At Point, the customer identifies sites of interest. Other services can monitor usage and send notifications automatically without the user making any overt request.

When the site monitoring matches the customer's interests, there is little or no objection in most instances. When the messages are about sites that the user visited unintentionally or confidentially, difficulties arise. For a public utility or a family-oriented publisher, providing links to certain types of information raises the stakes considerably. Will an offended parent sue a telephone company or a respectable publishing company for sending her daughter to a questionable site and thus triggering a flow of salacious electronic messages?

The ubiquity of electronic information systems provides numerous opportunities for systems to gather and disseminate information with or without the consent of the users. The practice of states such as Indiana and California to create online files of public records raises concern among some about the use and possible abuse of computerised information about people. Medical records exist in computerised form. The majority of this information resides in proprietary systems. However, efforts in Japan, the United States and elsewhere to standardise medical information creates the possibility that these data can migrate into more public fora.

With software that can assemble, collate and analyse data on a range of criteria becoming more widely available, the debate about privacy seems warranted. Governments in their well-intentioned efforts for cost control may be providing database-savvy entrepreneurs with a digital goldmine. Cross tabulation and analysis of seemingly innocuous data can provide some startling insights into individuals and groups of people. Without delving too deeply into this subject, one can visualise the existence of data about usage of certain online information services, the user's finances and criminal or employment record.

Copyright eroded by digital flows

A small Massachusetts company, Star Burst Communications, released, to little fanfare, a product named Burst Multicast. The software automates a nettlesome Internet task: that of sending the same information to a number of people at one time. It provides a point-and-click interface and a useful tool for identifying the particular recipients of a file transfer. A few mouse drags and keystrokes allow a single user to distribute a single file or the contents of an entire Web server to any number of sites connected to the Internet at the speed of the connection. Armed with StarBurst and a T-1 line, megabytes of data can be moved to dozens or hundreds of servers simultaneously. Each site or recipient receives a mirror or perfect copy of a file or another Internet service.

The mechanics of transfer with a tool such as StarBurst, which is just one of many Internet utilities becoming available for the new medium, are so simple that they promise to make the transfer of large amounts of information trivial. The more widely such tools become available, the greater the information flows.

Tools such as StarBurst amplify the networked environment. From the vantage point of the site administrator, routine tasks can be accomplished more efficiently. From the user's point of view, access to needed information becomes more convenient. A less sanguine response can be expected from the 'author' or 'owner' of the information transferred at the click of the mouse. Unauthorised dissemination of information, regardless of format, is a serious breach of expectations, laws, mores, social conventions, etiquette and undoubtedly several dozen other concepts dear to the heart of the lawful and ethical person.

In the era of print, which has now evolved into a newer, more malleable medium, fast copying and perfect reproductions required several centuries to realise. In the leisurely innovation environment of the pre-computer era, copyright gained a place in most cultures' canons of law. Aside from technology, copyright flourished among those who valued scholarship and business, although one cannot be sure of the relative importance of the two factors. Unauthorised copying took time and was, when compared with the new medium's environment, relatively easy to trace.

Not so with digital copies: perfect copies are the outcome of processes that move at ultra high speeds. Information or software in Location One appear in Locations Two, Three and beyond as if by magic. A clever new medium craftsman can render the duplication anonymous. The data simply are everywhere without clues as to origin, copier, other locations and pedigree of the perpetrator.'

-
- [1] IBM and other companies have introduced technology that attaches a digital counter to an information object. When the counter reaches a pre-set limit, the article or other 'object' can no longer be examined. At this time the technology shows promise and is actively being investigated by the US Copyright Clearance Center. The approach uses a variation of encryption and requires special processing of the information object before it is distributed via an Internet or intranet.

Quite understandably, many commercial organisations and individuals are concerned about the unauthorised dissemination, use, re-use and re-publishing in other media of high-value information. Publishers worry that information posted on their Web sites will find its way into the hands of competitors. Authors wonder if they will lose revenues from unauthorised versions of their software residing in the bellies of pirate servers located in Hong Kong or Cleveland.

The perfect copy is a consequence of the new medium. Copyright does not stretch comfortably to cover the nuances of Internet publishing. Over time, systems and tools will be developed to provide appropriate controls for certain types of materials. In the interim, the perception of the Internet as an insecure, fast-and-loose environment where anything can happen and most certainly will, is largely correct for a certain category of information.

At risk are postings in various USENET discussion groups. Electronic mail transferred without encryption can be intercepted by the motivated expert. Software can be and is copied from user to user, often as an act of charity to fellow inventors of the new medium. Other software and information companies provide high-value information as part of their marketing activities. Users can be (or pretend to be) confused about what is free and what is for-fee. In short, the technology has outstripped many of the cultural restraints to unauthorised copying and re-use. There is little chance of a short-term resolution to the problem. There are several reasons for concern:

- The rapid growth of the Internet and the volume of information flowing on it provides ample proof of the pent-up demand for electronic publishing by millions of people and organisations. Some percentage of users knowingly or unknowingly will take actions that invade the rights of others. Furthermore, the behaviour of a small percentage of users will become increasingly noticeable as the user community grows.
- Systems to meter, limit, charge and protect certain information objects are becoming more robust and more widely available. Nevertheless, the technologies to control terabytes of information flow are immature. As Prodigy and CompuServe, two commercial services exercising control over who gains access to their for-fee services, have found, implementing appropriate safeguards is no easy task even in a private, password-protected online environment. In the digital Wild West, monitoring of misappropriation is just about impossible; control almost unimaginable.
- The status of information as 'free' versus 'for-fee', or software as 'freeware', 'shareware', or 'demonstration-ware' is often difficult to determine. Jim Button's Buttonware began as freeware and evolved into a commercial product. Large amounts of commercial software circulate on the Internet for legitimate reasons. Equally large and possibly larger

amounts circulate for illegitimate reasons. The transgressors today can plead ignorance. Proving guilt is difficult and requires considerable resources. ¹

Copy-right and author-might

The relationship between publishers and authors is a tenuous one. The publishers work hard to keep the authors happy. But some authors are growing wary of the publishers' extension of rights to all media, not just print. Several large American firms, for example, found themselves under the spotlight when 'militant' writers wanted a royalty from the electronic re-use of their materials. The transgressors included *The New York Times*, Lexis-Nexis and University Microfilms International.

Different rules of varying clarity apply in different countries. Publishers who want to obtain rights to materials created by an author in Canada, Australia or Britain must navigate through murky waters and then figure out how to re-use these materials in a CD-ROM or a syndicated column without running foul of copyright somewhere along the line.

To solve this problem, at least in part, the British publisher EMAP Metro revised its standard agreement in mid-1995. In addition to print rights, EMAP clearly stated in the agreement circulated to authors that it wanted rights to all subsequent re-use. The umbrella language embraced online, CD-ROM and virtually any electronic medium. Similar language has been added to agreements issued from Reed-Elsevier and many other publishing firms as well.

Copyright in America and Western Europe is crystal clear compared with that of many countries in other parts of the world. Windows 95 software programs, complete with holograms, were available in Asia two weeks before the release of the products in the United States. One of the senior managers of Engineering Information was fond of offering two volumes of Compendex citations for inspection. After allowing a visitor to inspect each, he would ask: "Which is the one we published?" Invariably, the visitor would point to the volume with the higher quality paper and precisely bound and trimmed pages. "Wrong", he replied. "A colleague in the Pacific Rim gave me this duplicate as an example of the regard in which our data are held."

[1] Prodigy has established a precedent of monitoring information on its system. For its trouble, Prodigy's diligence has been rewarded with legal action for alleged libel and with derision from interest groups whose messages were deleted by a vigilant Prodigy. CompuServe avows a hands-off policy on postings, but continues to use moderators of special interest groups (SIGs). To minimise software piracy, file size limits have been imposed on binary transfers and binary files in the personal file areas cannot be directly sent to another CIS user via CompuServe mail. Neither approach has proven 100% effective. The debate rages in various quarters of the online community about the responsibility of an online operator who unwittingly creates an environment where copyrighted or objectionable material is improperly posted.

What is the impact of a world hyper-extending copyright? Most importantly, creators (authors, photographers, designers, musicians) are beginning to move rapidly up the learning curve about the value of their information objects in the digital world.

Publishing has for centuries been defined within the boundaries of relatively well known slow-to-develop media. Even the breathtaking deployment of television technology unfolded over decades. Mechanisms driven by unions and professional association worked to develop formulae that would return some money to the owner of rights. The formulae were peculiar, arcane and largely disconnected from the realities of business life in the real world, but they were workable.

The digital revolution has moved more rapidly. The new medium of the Internet and its sister networks has exploded into the mainstream in a matter of months. The pace of innovation is fed by the flow of information itself. Like a nuclear breeder reactor, the more information stuffed into the medium, the more explosive the growth, the pace of innovation and the speed of data transfer. Why go to a record shop when the music can be downloaded directly to a sound-equipped PC? Why purchase non-digital data, when the electronic form will meet many needs nicely and let the money go directly to the author or digital intermediary?

Regardless of one's view of the creative act, the new medium hinges on individuals who invent content, images, sounds and animations. Packagers and distributors still retain some importance, but the creators are the raw material of the new medium and they now have the technology to go it alone if necessary, or to forge a virtual global community of like-minded creators. Either of these two actions poses significant challenges up and down the publishing or intermediary line, since publishers' power in the past rested mainly on their ability to offer the expensive packaging and global distribution skills and facilities that authors and creators lacked. Many publishing companies have not digested the significance of the new medium for their futures.

Rights are now the keys to the information kingdom. Creators have more power than at any time in recent history. The new medium is giving the creator the same type of influence enjoyed by the literati in Athens or Rome when the media were less malleable and copyable.

Other challenges warrant comment as well:

- *Control.* Technology can provide only some mechanisms. The nature of the new medium itself makes it almost impossible to prevent duplication of certain information objects. Proprietary file formats such as Adobe's Portable Document Format and digital signatures provide some safeguards. The fact remains that once data are displayed on a monitor, they can be captured and re-used.
- *Tracking.* One can know who accessed a particular information object when the access system matches users to sessions and keeps track of who

viewed or copied what objects. But there are unfortunately many ways to fool even the most sophisticated systems.

- *Re-use.* Once text, images, sound or any other information object resides in digital form on a workstation, the data can be manipulated. The proliferation of powerful computers throughout the world provides an innovation environment that has never before existed. When bright people make use of information resources, it becomes difficult to know if a particular bit of information has a legitimate provenance. When a single bit is altered, is the 'new' object original, or is it an unauthorised copy?

To be clear: copyright in the new medium will be difficult to protect. It is unlikely that legislation alone will be the solution. Contracts and litigation may be practical approaches in some instances.

The unbreakable linkage: security, copyright and privacy

Most analysts of the Internet do not see the linkages that exist between the three issues of security, copyright and privacy. The new medium requires a different way of thinking about each of these hitherto separate concepts.'

Security is more a matter of system design, planning and skilful technological implementation than an absolute. Publicly-available networks are difficult to bend to the strictures of one-to-one confidential interaction. Regardless of the steps one takes, it is inevitable that someone will breach the system. Thus, the best remedies are planning, monitoring and appropriate staffing and resources.

Copyright laws in their present form are not up to the task of controlling the digital environment. Thus, effective protection of information can rely upon technology to an extent. The use of proprietary file formats, digital signatures, tokens and elaborate security procedures are effective – to a point. Then they break down because virtually any type of digital information can be instantly copied, modified and moved from point to point in a matter of seconds.

Privacy is becoming increasingly difficult to protect in the new medium. One solution is the use of what CompuServe used to refer to as 'handles' and avatars on the Internet. These are artificial personae that mask the identity of the user. However, if the look-up table with the real name and the false identity are breached, the concept of anonymity breaks down. Data about personal and private matters are a subset of security.

[1] A new software niche has been created to block sites that parents or employers find objectionable. The most interesting of these is called Software Nanny. It gives the parent-oriented adult the ability to limit what the computer user can do by blocking Uniform Resource Locators and certain key words. CompuServe has removed 200 Internet sites from its services as a result of pressure applied by Germany and Spain.

What is the outlook for these three issues? Are these problems substantive, or are they the fall-out from the supernova of the Internet's explosion into popular consciousness?

First, these issues are real and vital. The technology of the new medium has created an environment filled with paradoxes and contradictions. The very openness that has given the new medium its vitality is vulnerable because it is difficult to provide certain safeguards. After centuries of print, it should be possible to protect one's intellectual effort from inappropriate or unlawful use. Common sense might argue that such safeguards must be put in place; the reality of the new medium is that reasonable assurances are indeed possible. Absolute guarantees will be a long time coming, if indeed they ever come at all.

Second, creators and organisations have difficulty seeing that the blend of security, copyright and privacy are not separate issues. The result of this mix is the key issue of the new medium. The long-term success of the new medium as more than a public relations and marketing tactic hinges upon how we come to terms with:

- Providing any access to the right person at the appropriate time to public, private, proprietary or confidential information.
- Implementing systems that provide appropriate safeguards that are neither onerous, costly nor ineffective.
- Managing the information about users of systems in some appropriate way so that abuses can be minimised.
- Building systems and processes that operate globally.

At this time, none of these four points has been satisfactorily resolved at other than on a limited level.

Third, solutions will not emerge from outside the new medium. The medium itself will have to generate solutions from within. Copyright is a consequence of the experiences of publishers, artists and other creators over centuries. The body of law, despite its flaws, represents cumulative historical experience. The new medium has a short history. Furthermore, its dynamics make historical accumulation of experience almost laughable. Law and real-time quantum changes in technical functionality have different time scales. Monitoring of new media activity is technically impossible. Regulation for most of the new medium's short history has come from the users themselves. Will the user community be able to generate solutions to the new issue of security-copyright-privacy?

A solution or cluster of solutions will emerge over time. In the interim, the best safeguards are those driven by the careful design, implementation and monitoring of the information constructs created for the new medium.

Outlook: new medium and new rules?

Consider this: how easy is it for a dishonest person to work as a waiter for a day or two, copy the credit card number and expiry date of the well-heeled patron and use

those numbers? Credit card companies struggle to contain such routine credit card theft. Cyberspace is another class of problem entirely:

- The notion of cyberspace is new and not well understood by the majority of the world's population. Only two or three percent of the population of the United States uses online services.
- Security hinges upon technology. The expertise required to design, set up, maintain and defeat various safeguards is outside the mainstream of most computer specialists.
- Security and threats to security engage in a form of war game. Each advance challenges those who would crack the system to renew their efforts to defeat the new barrier. Each breach leads to new safeguards. The cycle of measure and countermeasure is locked in a high-stakes game of one-upmanship.
- Standards have not emerged because the free market promises huge financial rewards and significant competitive advantage to the organisation able to impose its solution upon the Wild West of the datasphere.
- National security agencies do not want security to prevent government monitoring of certain activities. Thus, secure systems are, in truth, not too secure. If they were, powerful entities of major governments would be blinded; that is, unable to 'see' or monitor potentially threatening activities.

Unlike print or television, people in cyberspace are coalescing into a global market of sorts. Furthermore, there remains some vestiges of the climate of mutual support and co-operation that characterised the pre-commercialisation of the Internet. Crime in cyberspace is abstract, one that involves bits and bytes, not life and limb. An injured party might not know of an intrusion for years, or until an unexpected charge appears on a Visa statement.

But the tangible consequences of security breaches can be serious. A stolen credit card or telephone calling card number can be disseminated globally in seconds. Vital medical information might be corrupted with potentially life-threatening consequences. Companies can lose their credit or credibility and fail.

However, special care must be taken to set up interactive services that walk a fine line between reasonable security and procedures that alienate honest customers. The nature of network environments is that breaches of security are usually the act of one person or a small group of individuals working as a team.

But as the financial stakes go up, the payoff for wrongdoers will escalate. Security must respond, or the momentum of network publishing will be lost. The reality is that security will be one of the major network publishing business opportunities in the next 12 to 36 months. The important trends that are gaining momentum include:

- *Embedded, standard security.* Major operating systems such as those from Sun Microsystems and Novell, and Microsoft's Windows NT include more robust tools.
- *New hardware devices.* The range of devices available for online interactivity is remarkable. Consider modems: laptop users must place calls through modems that provide the proper response.
- *Firmware* that provides user-defined security features. The AT&T Bolt, Beranek and Newman Internet monitoring service incorporates proprietary software and devices to monitor, track, follow and neutralise unauthorised users. The devices can be installed at the client site and monitored from anywhere in the world.
- *Network security* features will become more robust. The value-added network is able to enforce end-to-end security built upon various encryption technologies. Users will be able to authenticate their documents with digital signatures. Documents will have digital date, time and permission 'stamps' included in them.

Before probing into the links between security, copyright and privacy, it is unlikely that network publishing will enjoy the type of safeguards associated with more traditional media. The principal differences are highlighted in the table on the following page.

What a glance at the table reveals is that network publishing requires a different way of viewing how and why to protect intellectual property. A document available via network publishing technology can be copied perfectly an infinite number of times, distributed globally in seconds and manipulated easily and often automatically into a new revenue-generating network-published document almost invisibly. Digital trails are difficult to trace.

The potential of network publishing is not likely to be realised unless creators (author, publishers, aggregators) can be assured that:

- Suitable protection exists for their products and services.
- The money they collect can be credited to their account and used like hard currency, traditional bank transfers, or plastic credit cards.
- Their business and personal activities can be conducted without fear of eavesdropping or compromise.

There can be no clear line drawn between what have traditionally been thought of as the separate domains of security (in the sense of protection of systems and information from unauthorised intrusion or alteration), copyright (in the sense of appropriate protection for unauthorised or improper use of another's intellectual product) and privacy (in the sense of protection of information about the private affairs on an individual). This list raises difficult issues for different governments and cultures. In the world without borders of cyberspace, legal, political and legislative resolution of these issues is a daunting task.

Security snapshot: the building blocks for protection of information use and re-use

<i>Technique</i>	<i>How it works</i>	<i>Yardstick for use</i>	<i>Qualitative security rating</i>
Authentication	A process that reads bits and exchanges information with another server to determine that the sender has sent a message, not a person posing as a sender.	Promising technology for financial transactions. Technology is not widely known outside of security specialists and, therefore, seems to offer promise for financial transactions.	Mid level, possibly very high. Will require specialised software on the sender's and receiver's computer. <i>Weakness:</i> may be subject to software that fools the authenticating system.
Encryption	Messages are encoded in a form that cannot be read in their native form unless a key is used to unlock the message. Fees for encryption technology range from zero for public domain tools like Pretty Good Privacy to five figure license fees for commercial-grade algorithms and support tools.	All public network traffic intended for a single reader should be encrypted.	Mid to high level of security. Depends upon the type of encryption technology employed and the individuals who use that technology. The Netscape encryption technology for non-American use was compromised by a student in France in late 1995. <i>Weakness:</i> decryption keys are cracked by hackers or leaked by a trusted user.
Firewalls	Software looks at each packet to make certain that the sender has authorisation to use certain server-facilitated information or services.	All servers connected to the Internet should use firewall technology between the server and the Internet. If the server is connected to another network, another firewall is required.	Mid level security. Proven technology. Many sources for both hardware, software and firmware implementations. <i>Weakness:</i> system administration faulty.
Password	Authorised users must enter a string of alphanumeric characters to gain access.	Password control is provided by most server software packages. Specialised password software is available to supplement the built-in functionality.	Low to mid level of security. Depends upon the user community, the frequency with which passwords change, and other security procedures in place.
Private network	A value-added network reseller like General Electric Information Services Co. Fees vary from a low of five figures to six figures or higher. Average: \$500,000 for a worldwide system with 200 access sites.	Technique used by major banks and certain government agencies. Traffic runs on a proprietary network architecture or TCP/IP. Cost is not a barrier.	Highest possible level of security because it can implement all of the other techniques. Security can be compromised by a trusted individual. <i>Weakness:</i> a person can compromise the network.
Unidentified or 'hidden' server	Users do not know the address for the server.	Early public systems available via dial up from value-added network providers offered little security other than keeping the 'address' secret. Once the server is located, it can be hacked using various tools unless other security measures are taken.	Low to mid level of security. Depends upon the user community, the content, and the other security devices in place.
Tokens	Special sequences of bits are embedded in the encrypted message. Special software adds data to a message or a file that provides information about origin, copies that can be made, etc.	Promising technology to meter the re-use of certain types of information. Technology is not widely known outside of security specialists and, therefore, seems to offer promise for publishers.	Mid level, possibly very high if the embedded bits can resist tampering or being defeated by some type of spoofing technique. <i>Weakness:</i> not subjected to widespread public use. Vulnerability not known.