INTERNET DRAFT

Neil Haller, Bellcore
Randall Atkinson, NRL
5 October 1993

Internet Authentication Guidelines

STATUS OF THIS MEMO

This document is an Internet Draft. Internet Drafts are working documents of the Internettt Engineering Task Force (IETF), it's Areas and Working Groups. Note that other groups may also distribute working documents as Internet Drafts.

Internet Drafts are draft documents valid for a maximum of six months. Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet Drafts as reference material or to cite them other than as a "working draft" or "work in progress."

To learn the current status of any Internet Draft, please check the 1id-abstracts.txt listing contained in the Internet-Drafts Shadow Directories on ds.internic.net, nic.nordu.net, ftp.nisc.sri.com, or munnari.oz.au.

The distribution of this Internet Draft is unlimited. It is filed as <draft-haller-auth-requirements-01.txt>, and it replaces an earlier Internet Draft "Internet Authentication Requirements", August 1993. It expires on April 5, 1994.

ABSTRACT

The authentication requirements of computing systems and network protocols vary greatly with their intended use, accessibility, and their network connectivity. This document describes a spectrum of authentication technologies and provides guidance to protocol developers on what kinds of authentication might be suitable for what kinds of protocols and applications used in the Internet.

DEFINITION OF TERMS

This section briefly defines some of the terms used in this paper to aid the reader in understanding the draft.

Active Attack: An attempt to gain authentication or authorization by inserting false packets into the data stream or by modifying packets transiting the data stream. (See passive attacks and replay attacks.)

Asymmetric Cryptography: An encryption system that uses different keys, for encryption and decryption. Also called Public Key Cryptography. (See Symmetric Cryptography)

Authentication: The verification of the identity of the source of information, possibly including verification that the information has not been tampered with since being sent.

Authorization: The granting of access rights based on an authenticated identity.

Confidentiality: The protection of information so that someone not authorized to access the information cannot read the information even though the unauthorized person might see the information's container (e.g. computer file or network packet).

Encryption: A mechanism often used to provide confidentiality.

Integrity: The protection of information from unauthorized modification.

Key Certificate: TBD

Passive Attack: An attack on an authentication system that inserts no data into the stream, but instead relies on being able to passively monitor information being sent between other parties. This information could be used a later time in what appears to be a valid session. (See active attack and replay attack)

Plain-text: Unencrypted text.

Replay Attack: An attack on an authentication system by recording and replaying previously sent valid messages (or parts of messages). Any constant authentication information, such as a password or electronically transmitted biometric data, can be recorded and used later to forge messages that appear to be authentic.

Symmetric Cryptography: An encryption system that uses the same key for encryption and decryption. Sometimes referred to as Secret Key Cryptography.

AUTHENTICATION TECHNOLOGIES

There are a number of different classes of authentication, ranging from no authentication to very strong authentication. Different authentication mechanisms are appropriate for addressing different kinds of authentication problems, so this is not a strict hierarchical ordering.

No Authentication

For completeness, the simplest authentication system is not to have any. A non-networked PC in a private location or a stand-alone public workstation containing no sensitive data need not authenticate potential users.

Disclosing Passwords

The simple password check is by far the most common form of authentication. Password checks come in many forms: the key may be a password memorized by the user, it may be a physical or electronic item possessed by the user, or it may be a unique biological feature. Simple password systems are said to use disclosing passwords because if the password is transmitted over a network it is disclosed to eavesdroppers. Access keys may be stored on the target system, in which case a single breach in system security may gain access to all passwords. Alternatively, as on most systems, the data stored on the system can be enough to verify passwords but not to generate them.

Non-disclosing Passwords

Non-disclosing password systems have been designed to prevent replay attacks. Several systems have been invented to generate non-disclosing passwords. For example, the SecurID Card from Security Dynamics uses synchronized clocks for authentication information. The card generates a visual display and thus must be in the possession of the person seeking authentication. The S/KEY authentication system developed at Bellcore generates multiple single use passwords from a single secret key. [SKEY] It does not use a physical token, so it is also suitable for machine-machine authentication. In addition there are challenge-response systems in which a device or computer program is used to generate a verifiable response from a non-repeating challenge. These systems vary in the sensitivity of the information stored in the authenticating host, and thus vary in the security requirements that must be placed on that host.

Stronger Authentication Systems

The growing use of networked computing environments has led to the need for stronger authentication. In open networks, many users can gain access to any information flowing over the network, and with additional effort, a user can send information that appears to come from another user.

More powerful authentication systems make use of the computation capability of the two authenticating parties. Authentication may be unidirectional such as most time sharing systems, or it may be mutual in which case the entity logging in is assured of the identity of the host. Authentication systems use cryptographic techniques and establish as a part of the authentication process a shared secret (session key) that can be used for further exchanges. One example is the passing of a ticket that can be use to obtain other services without further authentication. These authentication systems can also provide confidentiality (using encryption) over insecure networks when required.

Symmetric Cryptography

Symmetric Cryptography includes all systems that use the same key for encryption and decryption. This means that knowledge of the key by an undesired third party fully compromises the confidentiality of the system. Therefore, the keys used need to be distributed securely, either by courier or perhaps by use of a key distribution protocol, of which the best known is perhaps that proposed by Needham and Schroeder. [NS78, NS87] The widely used Data Encryption Standard (DES) algorithm, which has been standardized for use to protect unclassified civilian US Government information, is perhaps the best known symmetric encryption algorithm. [NBS77]

A well known system that addresses insecure open networks as a part of a computing environment is the Kerberos Authentication Service that was developed as part of Project Athena at MIT. [SNS88, BM91] Kerberos is based on Data Encryption Standard (DES) symmetric key encryption and uses a trusted (third party) host that knows the secret keys of all users and services, and thus can generate credentials that can be used by users and servers to prove their identities to other systems. As the Kerberos server knows all secret keys, it must be physically secure. Kerberos session keys can be used to provide confidentiality between any entities that trust the key server.

Asymmetric Cryptography

In the recent past, a major breakthrough in cryptology has led to the availability of Asymmetric Cryptography. This is different from Symmetric Cryptography because different keys are used for encryption and decryption, which greatly simplifies the key distribution problem. The best known asymmetric system is based on work by Rivest, Shamir, and Adleman and is often referred to as "RSA" after the authors' initials. [RSA78]

SPX is an experimental system that overcomes the limitations of the trusted key distribution center of Kerberos by using RSA Public Key Cryptography. [TA91] SPX assumes a global hierarchy of certifying authorities at least one of which is trusted by each party. It uses digital signatures that consist of a token encrypted in the private key of the signing entity and that are validated using the appropriate public key. The public keys are known to be correct as they are obtained under the signature of the trusted certification authority. Critical parts of the authentication exchange are encrypted in the public keys of the receivers, thus preventing a replay attack.

Digital Signatures

Digital signatures are a comparatively recent addition to the tools available to protocol designers. A digital signature performs a function analogous to written signatures. It serves to authenticate a piece of data as to the sender and possibly as to the integrity of the data. A digital signature using asymmetric technology (public key) can also be useful in proving that data in fact originated with a party even if the party denies having sent it, a property called non-repudiation. A digital signature provides authentication without confidentiality and without incurring some of the difficulties in full encryption. For example, Secure SNMP and SNMPv2 both calculate a MD5 cryptographic checksum over a shared secret item of data and the information to be authenticated. [Rivest92, GM93] This serves as a digital signature and it is believed to be very difficult to forge such a digital signature or to invert it to recover the shared secret data. Digital signatures can be used to provide relatively strong authentication and are particularly useful in host-to-host communications.

USER TO HOST AUTHENTICATION

There are a number of different approaches to authenticating users to remote or networked hosts. Two hazards are created by remote or networked access: First an intruder can eavesdrop on the network and obtain user ids and passwords for a later replay attack. This is called a passive attack. Second, an intruder can "take over" a connection after authentication; this is an example of an "active attack".

Currently, most systems use plain-text disclosing passwords sent over the network (typically using telnet or rlogin) from the user to the remote host. [Anderson84 , Kantor91] This system does not provide adequate protection from reply attacks where an eavesdropper gains remote user ids and remote passwords.

Failure to use at least a non-disclosing password system means that unlimited access is unintentionally granted to anyone with physical access to the network. For example, anyone with physical access to the Ethernet cable can impersonate any user on that portion of the network. Thus, when one has plain-text disclosing passwords on an Ethernet, the primary security system is the guard at the door (if any exist). The same problem exists in other LAN technologies such as Token-Ring or FDDI. In some small internal Local Area Networks (LANs) this may be acceptable to take this risk, but it is an unacceptable risk in an Internet.

The minimal defense against eavesdropping is to use a non-disclosing password system. Such a system can be run from a dumb terminal or a simple communications program (e.g. CTRM or PROCOMM) that emulates a dumb terminal on a PC class computer. Using a stronger authentication system would certainly defend against passive attacks against remotely accessed systems, but at the cost of not being able to use simple terminals. It is reasonable to expect that the vendors of communications programs and non user-programmable terminals (such as X-Terminals) would build in non-disclosing password or stronger authentication systems if they were standardized or if a large market were offered.

Perimeter defenses are becoming more common. In these systems, the user first authenticates to an entity on an externally accessable portion of the network, possibly a "firewall" host on the Internet, using a non-disclosing password system. The user then uses a second system to authenticate to each host, or group of hosts, from which service is desired. This decouples the problem into two more easily handled situations.

There are several disadvantages to the perimeter defense, so it should be thought of as a short term solution. The gateway is not transparent at the IP level, so it must treat every service independently. The use of double authentication is, in general, difficult or impossible for computer-computer communication. End to end protocols, which are common on the connectionless Internet, could easily break. The perimeter defense must be tight and complete, because if it is broken, the inner defenses tend to be too weak to stop a potential intruder. For example, if disclosing passwords are used internally, these passwords can be learned by an external intruder (eavesdropping). If that intruder is able to penetrate the perimeter, the internal system is completely exposed. Finally, a perimeter defense may be open to compromise by internal users looking for shortcuts.

A frequent form of perimeter defense is the application relay. As these relays are protocol specific, the IP connectivity of the hosts inside the perimeter with the outside world is broken and part of the power of the Internet is lost.

An administrative advantage of the perimeter defense is that the number of machines that are on the perimeter and thus vulnerable to attack is small. These machines may be carefully checked for security hazards, but it is difficult (or impossible) to guarantee that the perimeter is leak-proof. The security of a perimeter defense is complicated as the gateway machines must pass some types of traffic such as electronic mail. Other network services such as the Network Time Protocol (NTP) and the File Transfer Protocol (FTP) may also be desirable. [Mills92, PR85] Furthermore the perimeter gateway system must be able to pass without bottleneck the entire traffic load for its security domain.

In the foreseeable future, the use of stronger techniques will be required to protect against active attacks. Many corporate networks based on broadcast technology such as Ethernet probably need such techniques. To defend against an active attack, or to provide privacy, it is necessary to use a protocol with session encryption, for example Kerberos, or use an authentication mechanism that protects against replay attacks, perhaps using time stamps. In Kerberos, users obtain credentials from the Kerberos server and use them for authentication to obtain services from other computers on the network. The computing power of the local workstation is used to decrypt the credentials (using a key derived from the user-provided password) and store them until needed. If the security protocol relies on synchronized clocks, then NTPv3 will be useful because it distributes time amongst a large number of computers and is one of the few existing Internet protocols that includes solid authentication mechanisms.

Another approach to remotely accessible networks of computers is for all externally accessable machines to share a secret with the Kerberos KDC. In a sense, this makes these machines "servers" instead of general use workstations. This shared secret can then be used encrypt all communication between the two machines enabling the accessable workstation to relay authentication information to the KDC in a secure way.

Finally, workstations that are remotely accessible could use asymmetric cryptographic technology to encrypt communications. The workstation’s public key would be published and well known to all clients. A user could use the public key to encrypt a simple password and the remote system can decrypt the password to authenticate the user without risking disclosure of the password while it is in transit.

KEY DISTRIBUTION & MANAGEMENT

The discussion thus far has periodically mentioned keys, either for encryption or for authentication (e.g. as input to a digital signature function). Key management is perhaps the hardest problem faced when seeking to provide authentication in large internetworks. Hence this section provides a very brief overview of key management technology that might be used.

The Needham & Schroeder protocol, which is used by Kerberos, relies on a central key server. In a large internetwork, there would need to be significant numbers of these key servers, at least one key server per administrative domain. There would also need to be mechanisms for separately administered key servers to cooperate in generating a session key for parties in different administrative domains. These are not impossible problems, but this approach clearly involves significant infrastructure changes.

Most public-key encryption algorithms are computationally expensive and so are not ideal for encrypting packets in a network. However, the asymmetric property makes them very useful for setup and exchange of symmetric session keys. In practice, the commercial sector probably uses asymmetric algorithms primarily for key exchange and not for encryption. Both RSA and the Diffie-Hellman techniques can be used for this. One advantage of using asymmetric techniques is that the central key server can be eliminated. The difference in key management techniques is perhaps the primary difference between Kerberos and SPX. Privacy Enhanced Mail uses asymmetric digital signatures for trusted key authorities to sign the public keys of users. The result of this use is key certificates which contain the public key of some party and authentication that the public key in fact belongs to that party. Key certificates can be distributed in many ways. One way might be to extend the existing Domain Name System by adding a new DNS record that would hold the key certificate for a host.

For multicast sessions, key management is harder because the widely used key management techniques have their number of operations proportional to the number of participating parties. In the future, more scalable multicast key management techniques are desirable.

Finally, key management mechanisms described in the public literature have a long history of subtle flaws. There is ample evidence of this, even for well-known techniques such as the Needham & Schroeder protocol [NS78, NS87]. In some cases, subtle flaws have only become known after formal methods techniques were used in an attempt to verify the protocol. Hence, it is highly desirable that key management mechanisms be kept separate from authentication or encryption mechanisms as much as is possible.

AUTHENTICATION OF NETWORK SERVICES

In addition to needing to authenticate users and hosts to each other, many network services need or could benefit from authentication. This section describes some approaches to authentication in protocols that are primarily host to host in orientation. As in the user to host authentication case, there are several techniques that might be considered.

The most common case at present is to not have any authentication support in the protocol. Bellovin and others have documented a number of cases where existing protocols can be used to attack a remote machine because there is no authentication in the protocols. [Bellovin89]

Some protocols provide for disclosing passwords to be passed along with the protocol information. The original SNMP protocols used this method and a number of the routing protocols continue to use this method. [Moy91, LR91, CFSD88] This method is useful as a transitional aid to slightly increase security and might be appropriate when there is little risk in having a completely insecure protocol.

There are many protocols that need to support stronger authentication mechanisms. For example, there was widespread concern that SNMP needed stronger authentication than it originally had. This led to the publication of the Secure SNMP protocols which support optional authentication, using a digital signature mechanism, and optional confidentiality, using DES encryption. The digital signatures used in Secure SNMP are based on appending a cryptographic checksum to the SNMP information. The cryptographic checksum is computed using the MD5 algorithm and a secret shared between the communicating parties so is believed to be difficult to forge or invert.

Digital signature technology has evolved in recent years and should be considered for applications requiring authentication but not confidentiality. Digital signatures may use a single secret shared among two or more communicating parties or it might be based on asymmetric encryption technology. The former case would require the use of predetermined keys or the use of a secure key distribution protocol, such as that devised by Needham and Schroeder. In the latter case, the public keys would need to be distributed in an authenticated manner. If a general key distribution mechanism were available, support for optional digital signatures could be added to most protocols with little additional expense. Each protocol could address the key exchange and setup problem, but that might make adding support for digital signatures more complicated and effectively discourage protocol designers from adding digital signature support.

For cases where both authentication and confidentiality are required on a host-to-host basis, session encryption could be employed using symmetric cryptography, asymmetric cryptography, or a combination of both. Use of the asymmetric cryptography simplifies key management. Each host would encrypt the information and within the host, the existing operating system mechanisms would provide protection.

In some cases, possibly including electronic mail, it might be desirable to provide the security properties within the application itself in a manner that was truly user-to-user rather than being host-to-host. The Privacy Enhanced Mail (PEM) work is employing this approach. [Linn93, Kent93, Balenson93, Kaliski93]

FUTURE DIRECTIONS

Systems are moving towards the cryptographically stronger authentication protocols described in the first paragraph. This move has two implications for future systems. We can expect to see the introduction and eventually the widespread use of public key crypto-systems. Session authentication, integrity, and privacy issues are growing in importance. As computer-to-computer communication becomes more important, protocols that provide simple human interfaces will become less important. This is not to say that human interfaces are unimportant; they are very important. It means that these interfaces are the responsibility of the applications, not the underlying protocol. Human interface design is beyond the scope of this memo.

The use of public key crypto-systems for user-to-host authentication solve many security issues, but unlike simple passwords, a public key cannot be memorized. As of this writing, public key sizes of at least 500 bits are commonly used in the commercial world. It is likely that larger key sizes will be used in the future. Thus, users might have to carry their private keys in some electrically readable form. The use of read-only storage, such as a floppy disk or a magnetic stripe card provides such storage, but it might require the user to trust their private keys to the reading device. Use of a smart card, a portable device containing both storage and program might be preferable. These devices have the potential to perform the authenticating operations without divulging the private key they contain. They can also interact with the user requiring a simpler form of authentication to "unlock" the card.

The use of public key crypto-systems for host-to-host authentication appears not to have the same key memorization problem as the user-to-host case does. A multiuser host can store its key(s) in space protected from users and obviate that problem. Single user inherently insecure systems, such as PCs and Macintoshes, remain difficult to handle but the smart card approach should also work for them.

The implications of this taxonomy are clear. Strong cryptographic authentication is needed in the near future for many protocols. Public key technology should be used when there is benefit to do so and when it is practical and cost-effective. In the short term, the use of disclosing password systems should be phased out in favor of non-disclosing systems and digital signatures. Additional research is needed to develop improved key management technology and scalable multicast security mechanisms.

SECURITY CONSIDERATIONS

The entire Internet Draft discusses Security Considerations in that it discusses authentication technologies and needs. There are no security issues regarding the public release of this draft.

REFERENCES

[GM93] J. Galvin & K. McCloghrie, Security Protocols for Version 2 of the Simple Network Management Protocol (SNMPv2), RFC-1446, DDN Network Information Center, April 1993.

[SNS88] J.G. Steiner, C. Neuman, & J.I. Schiller, Kerberos: "An Authentication Service for Open Network Systems", *USENIX Conference Proceedings* , Dallas, Texas, February 1988.

[BM91] Steven M. Bellovin & Michael Merritt, "Limitations of the Kerberos Authentication System", ACM Computer Communications Review, October 1990.

[Bellovin89] Steven M. Bellovin, "Security Problems in the TCP/IP Protocol Suite", ACM Computer Communications Review, Vol. 19, No. 2, March 1989.

[NBS77] National Bureau of Standards, "Data Encryption Standard", *Federal Information Processing Standards Publication 46*, Government Printing Office, Washington, DC, 1977.

[LR91] K. Lougheed & Y. Rekhter, "A Border Gateway protocol 3 (BGP-3)", RFC-1267, DDN Network Information Center, October 1991.

[SKEY] TBD

[TA91] J. Tardo & K. Alagappan, "SPX: Global Authentication Using Public Key Certificates", Proceedings of the 1991 Symposium on Research in Security & Privacy, IEEE Computer Society, Los Amitos, California, 1991. pp.232-244.

[Kaliski93] B. Kaliski, Privacy Enhancement for Internet Electronic Mail: Part IV: Key Certification and Related Services, RFC-1424, DDN Network Information Center, February 1993.

[Balenson93] D. Balenson, Privacy Enhancement for Internet Electronic Mail: Part III: Algorithms, Modes, and Identifiers, RFC-1423, DDN Network Information Center, February 1993.

[Kent93] S. Kent, Privacy Enhancement for Internet Electronic Mail: Part II: Certificate-Based Key Management, RFC-1422, DDN Network Information Center, February 1993.

[Linn93] J. Linn, Privacy Enhancement for Internet Electronic Mail: Part I: Message Encryption and Authentication Procedures, RFC-1421, DDN Network Information Center, February 1993.

[NS78] R.M. Needham & M.D. Schroeder, "Using Encryption for Authentication in Large Networks of Computers", Communications of the ACM, Vol. 21, No. 12, December 1978.

[NS87] R.M. Needham & M.D. Schroeder, "Authentication Revisited", ACM Operating Systems Review, Vol. 21, No. 1, 1987.

[RSA78] R.L. Rivest, A. Shamir, & L. Adleman, "A Method for Obtaining Digital Signatures and Public Key Crypto-systems", Communications of the ACM, Vol. 21, No. 2, February 1978.

[Mills92] D. Mills, Network Time Protocol (Version 3) Specification, Implementation, and Analysis, RFC-1305, DDN Network Information Center, March 1992.

[PR85] J. Postel & J. Reynolds, File Transfer Protocol, RFC-959, DDN Network Information Center, October 1985.

[Kantor91] B. Kantor, BSD Rlogin, RFC-1258, DDN Network Information Center, September 1991.

[Anderson84] B. Anderson, TACACS User Identification Telnet Option, RFC-927, DDN Network Information Center, December 1984.

[CFSD88] J. Case, M. Fedor, M. Schoffstall, & J. Davin, "Simple Network Management Protocol", RFC-1067, DDN Network Information Center, August 1988.

[Moy91] J. Moy, "OSPF Routing Protocol, Version 2", RFC-1247, DDN Network Information Center, July 1991.

EXPIRATION

This Internet Draft expires on April 5, 1994.

AUTHORS' ADDRESSES

Randall Atkinson <atkinson@itd.nrl.navy.mil>
Information Technology Division
Naval Research Laboratory
Washington, DC 20375-5320

Neil Haller <nmh@thumper.bellcore.com>
Bell Communications Research
445 South Street -- MRE 2Q-280
Morristown, NJ 07962-1910