

AD709366

SECURITY OF COMPUTERIZED INFORMATION SYSTEMS

Rein Turn
H. E. Petersen

July 1970

Best Available Copy

U.S. GOVERNMENT
PRINTING OFFICE
1970

Not approved

for release

P-4405

SECURITY OF COMPUTERIZED INFORMATION SYSTEMS*

Rein Turn
H. E. Petersen

The RAND Corporation, Santa Monica, California

*Any views expressed in this Paper are those of the authors. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

This was a Paper presented at the 1970 Carnahan Conference on Electronic Crime Countermeasures, Lexington, Kentucky, April 16-18, 1970.

SECURITY OF COMPUTERIZED INFORMATION SYSTEMS

by

Rein Turn

and

Harold E. Petersen

The RAND Corporation
Santa Monica, Calif. 90406

Abstract. This paper addresses the vulnerabilities of remotely accessible computerized information systems to electronic crime -- penetration of the information system for illicit copying, altering or destruction of selected information. A condensed survey of the probable threats and applicable countermeasures is presented. A particular emphasis is placed on the use of cryptographic techniques for protecting information in the communication lines and in computer files.

Introduction

The spectacular advances in computer technology in the last decade have set the stage in the 70's for a large increase of computerized information systems in all areas of endeavor: government, law enforcement, business, industry and banking. Already over 50,000 computer systems are installed in the United States and by 1975 this number is expected to pass the 150,000 mark.

Many of the present computerized information systems (and a much larger percentage in the future) operate in the on-line, time-shared manner -- an individual user types his information request and receives the requested data on a terminal that is connected to the central data processor and its information storage units through a data transmission system. The central processor cyclically allocates slices of time to each user, thus effectively processing all information requests simultaneously. In a large information system hundreds of terminals may be serviced simultaneously and they may be located hundreds of miles from the central facility.

Representative among the large computerized information systems are banks with terminals located at the branch offices, and commercial time-sharing firms that install terminals on the premises of their customers. It has been estimated that presently there are over 70,000 terminals connected to computers and this number may reach 400,000 by 1973.

The categories of information stored in remotely accessible time-shared computer systems range from personal data on tens of millions of individuals to proprietary industrial data, trade secrets, bank accounts and stock market transactions. In the hands of criminals or unscrupulous business competitors almost any category

of this information could be turned into a business advantage or a direct payoff in cash.

Storage of information in computerized form allows rapid retrieval and updating of files and drastically reduces the required storage space. However, information previously in the form of printed documents in locked file cabinets is now replaced by magnetization patterns on tapes and disks -- they can be anonymously read, altered or erased without a trace of evidence that this has occurred. Hence, anyone that has gained access to the information system could, in principle, manipulate any information in the files -- perhaps plant damaging information on a competitor, change bank accounts or copy trade secrets.

The increasingly large numbers of on-line information systems and associated terminals provide increased access opportunities and may make penetration of these systems appear profitable to a wider class of technically sophisticated but larcenously inclined individuals. Indeed, the "electronically perpetrated crime" appears to be characterized by a low physical risk, small probability of detection, anonymity, lack of evidence and a lack of applicable laws. Further, the level of expertise previously required for successful embezzlement has been reduced by simplification of business procedures for computerized operation. On the other side of the ledger we find, however, that the resources, both in equipment and know-how, required for successful penetration are considerably higher than those necessary for conventional burglaries or holdups.

The question of providing increased protection to data files in remotely accessible information systems has recently found considerable attention in the computer and industrial security fields (1,2,3,4,5). This paper presents a condensed survey of more likely threats and discusses several

classes of countermeasures. A particular emphasis is placed on the use of cryptographic techniques for protecting information in communication lines and computer files.

Threats to Data Security

The operation of an on-line time-shared information system is controlled by a set of master programs -- the operating system. The basic tasks of the operating system are to:

1. Receive access requests from terminals and verify that the user is authorized for access -- possesses a valid account number and/or password.
2. Control all communications with the terminals.
3. Schedule time slices to user programs or information requests.
4. Provide protections to users' programs and data (and to the operating system itself) against inadvertent destruction by other users.

These functions of the operating system are the "bare bones" protective shield that must be provided in any time-shared system. The protection can be augmented by establishing additional access controls at certain data files (e.g., the company payroll) and imposing operating restrictions against certain set of users (e.g., allowing reading but not altering data in certain files). Each increase in the protective features of the operating system is accompanied by a decrease in the efficiency of the information system as more and more computer time is diverted to these nonproductive tasks.

A successful penetration of an information system depends on the ability of the penetrator to cope with the existing access controls and protective measures -- he must gain physical access to a terminal or communication line of the system and then defeat the operating system. Several probable penetration tactics are enumerated below.

Deception

A user who is normally allowed to use the information system may have learned the account number or passwords that permit access to restricted files. He can now masquerade as an authorized user and see all files or perform all operations that were authorized for the latter.

Deception may also be attempted from outside of the system by connecting a compatible privately owned terminal into the system. If terminals are normally connected by dialing a telephone number, the illicit terminal could be connected in the same manner. If private lines are employed

it may be possible to physically tap the terminal into one of the lines. The necessary account numbers and passwords could be obtained from discarded printouts, by theft, wire tapping or bribery.

Wire Tapping

Passive eavesdropping by wire tapping is a low cost approach to copying all information communicated over the line. It is necessary to gain physical access to the communication lines and sort out the correct wires. A pickup device, tape recorder and a terminal (or equipment that can emulate the terminal) are required for recording and uncovering the information. Wire tapping must be considered as a very probable starting point for any determined external attack against the system.

Circumvention

Circumvention of the normal protective features may be possible through existing imperfections or oversights in the operating system. A penetrator could discover these by detailed studies of the operating system programs or experimentally. However, an experimental search for weaknesses in an operating system requires access to the system for prolonged periods of time and the persistent rejection of the penetrator's access attempts are likely to alarm the system management. If a circumvention scheme is discovered the payoff is great -- information can be copied, altered and destroyed with impunity.

Tampering

Deliberate insertion of circumvention schemes in the operating system implies the cooperation of persons who gain normal access to the operating system and to the computer hardware -- systems programmers and maintenance engineers. If appropriate controls are in force a conspiracy may be difficult to establish and a lone wolf approach ineffective.

Physical Penetration

Physical penetration of the computer facility by an outsider for the purposes of stealing access information or selected information files is so what outside the category of electronic crime although it may directly achieve objectives that may be exceedingly costly or time consuming by electronic means.

A different threat to information systems arises from persons whose objectives on breaking and entering are to cause disruption and destruction. Witness the recent damage to computers at the Stanford Research Institute and at the Williams University in Montreal. Many of these individuals are fanatics not deterred by physical risk, detection or apprehension.

Case Histories

To date the published accounts of electronic crime against information systems deal exclusively with tampering of the operating system or application programs by programmers themselves for financial gain. There was a programmer in Minneapolis whose checking account overdrafts were ignored by the computer and a conspiracy at a New York stock brokerage firm who mailed checks to their homes under fictitious names (6). It is interesting to note that a claim of "computer error" was used by the embezzlers whenever their activities caused suspicion. The management readily accepted this explanation! No penetrations of information systems from remote locations have been reported but it is known that tapping of data lines has been attempted (5).

Countermeasures

The normal protective features of a remotely accessible information system are not designed to resist sophisticated penetration attempts. For increased security they must be augmented by additional programmed procedures or electronic devices. The objective is not absolute security -- this can never be achieved, but rather an increase of the cost of penetration, the "work factor", to a level where the expected payoff becomes relatively small. At the same time, a balance must be maintained between the cost of countermeasures and the value of the protected information.

Improved Operating System

The key to successful application of any programmed protection procedure is the integrity of the operating system. It must be designed to be free of any vulnerabilities and imperfections -- every sequence of input statements by the user, however illogical, must be properly handled. The design of such a system requires a high degree of programming competence (8,9), a thorough initial checkout and periodic tests to verify its integrity.

Real-Time Monitoring

A capability to continuously monitor the system activity -- access requests by users, granting or refusing of access, status of the lists of current users and terminals -- provides a further increase in the system's security. Attempts at deception can be detected when two users claim the same identity or when two identically labeled terminals are connected. Unusual activity in a file or abnormally large numbers of access rejections may indicate attempts to penetrate the system.

The cost of real-time monitoring arises essentially in the storage space required for maintaining the logs and in the computer time for their updating. In large systems the cost may become sufficiently high to restrict monitoring only to selected sets of users and files.

Positive Identification

The normal access control procedures identify a user by his account number and may require additional passwords to authenticate his identity or gain access to restricted files. In principle, the protection afforded by the password approach may be increased to any level by requiring longer sequences of passwords. Protection against wire tapping is obtained by use of "onetime" passwords that are discarded after a single use.

A high degree of confidence in identification of an individual can be obtained by using fingerprints, signatures or voice characteristics as passwords (10,11,12,13). These techniques require installation of appropriate sensors, maintenance of a library of measurements that characterize the particular identifying feature for a set of known individuals, and a processor for comparison of the given sample against the master file.

A natural implementation of the fingerprint, signature or voice print identification devices to controlling access to information systems locates the sensors at terminals and performs the processing in the central computer. Note, however, that once again a threat arises from wire tapping -- the scan patterns from the sensor can be recorded and played back into the system during penetration. The cost depends on the complexity of the sensors and the amount of required processing.

Protected Communication Lines

Telephone lines connecting a terminal and the central computer are extremely vulnerable to wire tapping. By this means all communications on the line, including the passwords, can be recorded at a minimal cost and low risk. It is clear that a prerequisite for a high level of security is either an adequate physical protection of the communication lines or concealment of the transmitted information by cryptographic techniques.

Placing of wire taps on telephone lines, terminal boards in manholes or directly inside a telephone or data modem has become a sophisticated art (14). Detection of a tap on the external wires is extremely difficult by other than visual inspection.

The only effective protection against the placement of wire taps is to prevent access to the communication lines by using buried and armored cables, locked and alarmed terminal boards and manhole covers, and physical protection of the terminal facilities at the remote locations and at the central computer.

The concealment of information by cryptographic techniques is based on using a "key" -- a particular sequence of operations on the message -- to systematically transform the message at the transmitting

terminal into an unintelligible and apparently random sequence of characters. The original information is recovered at the receiving terminal by applying the same key in an inverse order. A properly designed cryptographic transformation provides a high level of protection by making the detection of the key by analysis of the intercepted messages an extremely difficult task. The keys themselves must receive the highest level of protection and be handled only by the most trustworthy employees of the information system.

Historically, the cryptographic equipment for communication lines has been very expensive and essentially inaccessible to commercial users. Only recently have a few preliminary models of data privacy devices, as they are called, become available for private business.

The use of cryptographic transformations for protecting computer data is discussed in more detail in a subsequent section of this paper.

Physical Protection of Premises

A large selection of electronic intrusion detection devices and associated alarms are available for designing physical protection systems for terminal and central computer facilities (15). Among the major categories of these devices are the following:

1. Electrical burglar alarms to detect opening of doors and windows.
2. Electromagnetic, optical or acoustical barriers for protecting an area. An intruder is detected when he approaches or crosses the barrier.
3. Electronic motion detectors register even the slowest motions.
4. Vibration detectors can be used to protect specific objects.

The output signals from anti-intrusion devices are normally connected to local alarms and also used to activate automatic telephone dialing equipment to notify appropriate security forces.

Cryptographic Countermeasures

Traditionally the applications of cryptography for concealment of information and the solution of resulting cryptograms (16) have been treated in open literature as a form of entertainment similar to crossword puzzles or anagrams. Except for Shannon's early work (7) the application of cryptography to the protection of commercial data has found wider attention only in the recent few years (1,17,18,19).

Transformations

Two basic classes of cryptographic transformations have been used since ancient times:

1. Substitution of the characters in the message with other characters or groups of characters. The replacement characters may come from an alphabet different from that used for the message.
2. Transposition of the sequencing of the characters in the message.

The set of parameters or rules that specify a particular substitution or transposition scheme is the key of the cryptographic transformation. The set of all possible keys is the key space. Even for very simple transformations the key space may be very large. For example, there are 10^{26} possible one-to-one substitutions of the 26 characters of the English alphabet.

The substitution transformations may be increased in complexity by cyclically applying N different substitutions -- a polyalphabetic substitution is used. The length of the key is now increased by N and the size of the key space by $N!$ (the number of permutations of the N sets of substitution rules).

The key space of the transposition transformations depends on the length of the group of characters that is being permuted. For a group of length M there are $M!$ possible transpositions. Practical aspects of implementing substitution and transposition transformations in a digital computer are discussed in Reference 19.

The best protection is obtained by making the keys very long and correspondingly, the key spaces very large. In fact, Shannon (7) has proven that the necessary and sufficient condition for a cryptographic transformation to be perfect (totally unbreakable) is that the key is random, the same length as the message, and used only once. A perfect cryptographic system is known as the Vernam system.

Although implementation of the Vernam system is not practical in commercial information systems, it is still possible to generate keys with very long periods by utilizing feedback shift-register circuits (20). It should be pointed out, however, that very large key spaces, for example the 10^{26} possible monoalphabetic substitutions mentioned above, do not necessarily imply large amounts of security. Indeed, Shannon (7) has shown that less than 100 properly chosen tests may be sufficient to determine the key in this case.

It is clear that unless the Vernam system is implemented, a large key space is only a necessary but by no means a sufficient prerequisite for obtaining increased security through cryptographic techniques.

Basic Elements

Essential in any application of cryptographic techniques to information in computer files or data communication links are the following:

1. A hardware device or a set of computer programs for automatic encrypting of the data at the transmitting terminal or when storing in the files.
2. A similar device or program for performing the decryption operation at the receiving terminal or upon reading the data from the files.
3. A mechanism for inserting the key into the device or the programs.

It is prudent to assume that neither the selected cryptographic system nor the construction details of the hardware or programs can be kept secret indefinitely. The entire burden of providing security rests, therefore, on the keys -- the penetrator who intercepts an encrypted message must not know which of the thousands of equally probable keys was used. Thus, it is imperative to provide the keys with the highest attainable protection. Further, they should be changed as frequently as operationally feasible and every communication link and file should use a different key.

Vulnerabilities

Intercepted encrypted messages are not solved by trying all possible keys. Instead, the solution is attempted and often obtained with amazingly little effort by utilizing the following known or postulated characteristics of the message language:

1. Average frequencies of occurrences of single characters or pairs of characters (digrams).
2. The size of the vocabulary and the frequencies of occurrences of words -- the "probable word" approach.
3. Redundancy in the words -- the number of characters that may be deleted from a word and still allow its unique determination.
4. The grammatical structure.

If a sufficient amount of encrypted message is intercepted then the results of these analyses on parts of the message can be used to form hypotheses concerning the key and tested on other parts of the message. If partial clear text emerges the hypotheses are modified and tested until the key is determined.

Historically, all but a very few non-perfect cryptographic systems have eventually been "broken." The principal factors that contributed to their weaknesses were use of keys with short periods, use of highly formatted messages with limited vocabularies (as in military communications), and encrypting many messages with the same key.

Although the wouldbe penetrators of an information system are not expected to have infinite resources and time available for breaking of keys, it is still necessary to analyze the computer languages used in the system and the data in the files to determine whether their characteristics may seriously increase the vulnerability of the cryptographic system under consideration.

Properties of Computer Languages

Computer languages are designed for certain specific purposes such as writing computer programs or querying computerized information files. Unlike the natural languages, they are designed to be precise and logical. They are characterized by the following:

1. Limited vocabulary -- a relatively small number of words or sentences are used.
2. Rigid format and grammatical structure -- spelling and rules for constructing sentences must be followed to the last detail.
3. Predictable statistics -- in order to accommodate the users, the vocabulary and sentence structure is usually designed to closely resemble the natural language. Hence the character frequencies may approximate those of the natural language.
4. Enriched character set -- computer languages generally employ a large set of special characters and symbols. They are often used in a predictable manner (e.g., every query statement may have to be terminated by a dollar sign).

All of the enumerated characteristics appear to make solution of the key easier by reducing the uncertainty regarding character or word usage.

In addition to statements in computer language, the encrypted messages will also contain large amounts of data from the computer files. Depending on the particular application area these data may consist largely of numbers, names, addresses and various sets of abbreviations. These can be expected to exhibit markedly different statistical features -- all possible N-digit numbers may be equally likely, names are not constrained to be words of the natural language, and all possible combinations of letters may be used as abbreviations or codes. The reduction of key uncertainty using statistical tests is now much more difficult.

Computer programs, however, can be expected to exhibit statistical properties somewhere between those of the highly structured query languages and the numerical data -- certain fixed statements must be used but the programmer is also free to invent his own names for variables and use numerical values for constants. To illustrate this,

Figure 1 displays a plot of the relative frequencies of character usage as found in the English language and in a typical computer program written in FORTRAN. Note the high usage of special characters (commas, parentheses, and the equal sign) in FORTRAN and the generally flatter shape of the usage curve.

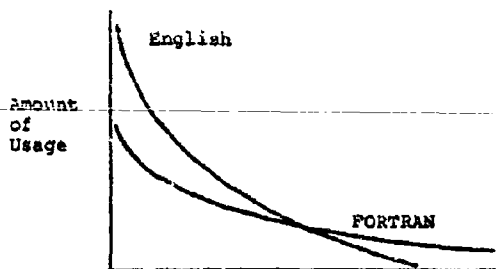


Figure 1 Letter Usage in English Text and in a FORTRAN Computer Program

Work Factor

It was suggested previously that the cost of breaking a key could be estimated in terms of a work factor for this task. This could then be used as relative measure of security of the cryptographic systems that are being considered.

Breaking of the key of an intercepted encrypted message is largely a problem of mathematics and logic. In general it is possible to derive mathematical equations and formulate statistical procedures that can be iteratively applied to evaluate hypotheses concerning the unknown key. If these procedures can be converted into computer programs, the work factor can be estimated in terms of the average number of arithmetic operations required to obtain the solution.

A dollar value of the work factor can now be estimated for a particular computer that is expected to represent the resources of the penetrator by determining the cost of computer time required to perform the necessary arithmetic operations.

The estimated work factor represents the cost of breaking one key. Its adequacy as a deterrent depends on the amount of information and the degree of penetration that could be attained on this basis. If the system uses many keys and they are frequently changed then even a work factor of just a few hours per key may be more than the penetrator's resources could handle.

Practical Aspects

There are a number of practical problems associated with the implementation of cryptographic transformations in a real-time communication network:

1. **Synchronization.** In all but the simplest monoalphabetic substitution it is necessary to maintain the encryption and decryption devices at both ends of the communication link in synchronism.
2. **Reserved codes.** A communication system usually employs a set of code words used for the control of the network. Inadvertent generation of these codes by the encryption process may interfere with the normal operation of the network. Special circuitry may be needed for their detection and handling.

The lack of computing capability at the remote terminals suggests use of electronic hardware rather than programmed procedures for key generation and application. Circuits based on the feedback shift-register action appear especially suitable (20) and can be miniaturized by using the integrated circuit technology. For example, it is possible to package a 23-stage shift-register and associated circuits in a space less than one cubic inch in volume. This generator can produce a key sequence of more than one million characters and has a key space of 356,000 different keys. Modern computing technology may permit even more complex procedures at modest cost (19).

Concluding Remarks

The threat to computerized information systems by electronic means -- the electronic crime -- is becoming a reality. The necessary technology exists, risk is low and pay-off may be great. Unlike in other forms of crime, however, the technology used for penetration can be applied even more effectively to provide protection.

The cost of protective devices and techniques is greatly reduced if they are included in the terminals and communication links in the design phase and not as a retrofit. We feel that is a high time for the electronic crime measures researchers and industry to turn their attention to providing effective low cost devices to counter the emerging threat of electronic crime now while it still is in its infancy.

References

1. Petersen, H. E. and Rein Turn, "Systems Implications of Information Privacy," AFIPS Conference Proceedings Vol. 30, 1967 Spring Joint Computer Conference, Thompson Book Co., New York, 1967, pp. 291-300.
2. Hoffman, L. J., "Computers and Privacy: A Survey," Computing Surveys, Vol. 1, No. 2, June 1969, pp. 86-103.
3. "The Considerations of Data Security in a Computer Environment," International Business Machines Corporation, New York, New York, 1969.

4. "Problems and Potential Solution in Computer Control," *Industrial Security*, Vol. 13, No. 2, April 1969.
5. "Computer Security," *Industrial Security*, Vol. 13, No. 6, December 1969.
6. "Crooked Operators Use Computers to Embezzle Money from Companies," *The Wall Street Journal*, April 5, 1968.
7. Shannon, C. E., "Communications Theory of Secrecy Systems," *Bell System Technical Journal*, Vol. 28, October 1969, pp. 656-715.
8. Peters, Bernard, "Security Considerations in a Multi-Programmed Computer System," *AFIPS Conference Proceedings*, Vol. 30, 1967 Spring Joint Computer Conference, Thompson Book Co., New York, 1967, pp. 283-286.
9. Weissman, Clark, "Security Controls in the ADEPT-50 Time-Sharing System," *AFIPS Conference Proceedings*, Volume 35, 1969 Fall Joint Computer Conference, The AFIPS Press, Montvale, New Jersey, 1969, pp. 119-133.
10. Busch, G. E., "Applications of Electro-Optical Fingerprint Correlators," *Proceedings, 1969 Carnahan Conference on Electronic Crime Countermeasures*, University of Kentucky, Lexington, Ky., pp. 90-97.
11. Dyche, J. W., "Positive Personnel Authentication by Handwriting," *Proceedings, 1969 Carnahan Conference on Electronic Crime Countermeasures*, University of Kentucky, Lexington, Ky., pp. 114-126.
12. Kersta, L. G., "Voice Pattern Identification of Speakers," *Proceedings, 1969 Carnahan Conference on Electronic Crime Countermeasures*, University of Kentucky, Lexington, Ky., pp. 127-136.
13. Luck, J. E., "Description of a Real Time Completely Automatic Speaker Verification System," *Proceedings, 1969 Carnahan Conference on Electronic Crime Countermeasures*, University of Kentucky, Lexington, Ky., pp. 98-113.
14. Purgslove, S. D., "The Eavesdroppers: 'Fallout' from R&D," *Electronic Design*, Vol. 14, No. 15, June 21, 1966, pp. 35-43.
15. Cantor, Lon, "Electronic Intrusion Alarms," *Electronics World*, Vol. 80, No. 3, September 1968, pp. 44-46.
16. Gaines, H. F., "Cryptanalysis," Dover Publications Inc., New York, 1956.
17. Van Tassel, Dannie, "Cryptographic Techniques for Computers," *AFIPS Proceedings*, Vol. 34, 1969 Spring Joint Computer Conference, The AFIPS Press, Montvale, New Jersey, 1969, pp. 367-372.
18. Kahn, David, "The Code Breakers," The McMillan Company, New York, 1967.
19. Skatrud, R. O., "A Consideration of the Application of Cryptographic Techniques to Data Processing," *AFIPS Conference Proceedings*, Vol. 35, 1969 Fall Joint Computer Conference, The AFIPS Press, Montvale, New Jersey, 1969, pp. 111-117.
20. Reed, I. S. and Rein Turn, "A Generalization of Shift-Register Sequence Generators," *Journal of the ACM*, Vol. 16, No. 3, July 1969, pp. 461-473.