NETWORKS WITHOUT USER OBSERVABILITY

Andreas Pfitzmann, Michael Waidner

Institut für Informatik IV, Universität Karlsruhe, Postfach 6980,

D 7500 Karlsruhe 1, West Germany

ABSTRACT

In present-day communication networks, the network operator or an intruder could easily observe when, how much and with whom the users communicate (traffic analysis), even if the users employ end-to-end encryption. With the increasing use of ISDNs, this becomes a severe threat.

Therefore, we summarize basic concepts to keep the recipient and sender or at least their relationship unobservable, consider some possible implementations and necessary hierarchical extensions, and propose some suitable performance and reliability enhancements.

Keywords
traffic analysis, user observability, anonymity, fault-tolerance, ISDN, MIX-network, DC-network, RING-network, switched/broadcast network

A previous version entiteled "Networks without User Observability -- Design Options" was presented at EUROCRYPT 85.

0 Motivation

Public and private communication networks have a growing importance for our daily life. We use them for telephony, telegraphy, television, videotex, radio and in the near future we will use them for video telephony, electronic mail, ordering and receiving of newspapers, home banking, etc.

All these services will be integrated in a so-called Integrated Services Digital Network (ISDN). If such a network is built as planned e.g. by the german PTT and operated on a "transmission on demand basis" even for the classical broadcast services TV and radio, major parts of any user's life might easily be observed by the PTT or by an intruder.

Eavesdropping can be foiled by link-by-link encryption [2], but this does not foil attackers at the stations (e.g. via Trojan Horses).

There are some well known measures allowing users themselves to decrease their observability. The content of a message can be sufficiently hidden by end-to-end encryption. However, an attacker can still observe who sends how many messages to whom and at what time (traffic analysis). To hide this information, too, they can use public network stations (e.g. telephone boxes) instead of private ones. This will prevent observation but is very uncomfortable for the users (e.g. who would watch TV in a video telephone box?). If they use private network stations, they can only try to hide their behaviour by making their network stations do more things than necessary at other times than necessary. For example a user can order a whole newspaper or several newspapers instead of a single article, and he can do so at any time before he wants to read them. This is an easy but expensive measure and not suitable for services like telephony.

So the only way to decrease user observability in a comfortable and cheap fashion seems to be to design a network for anonymity and not to try to realize anonymity afterwards.

Of course, the standard requirements for an ISDN, i.e. high performance and reliability, have to be met, too.

This paper describes the existing proposals for anonymous networks in a systematic way and some options allowing them to be adapted to meet the stringent requirements on performance and reliability.

1 Basic concepts for anonymous networks

1.1 A closer look at anonymity

What we would like to realize is absolute anonymity against every possible attacker. But an attacker can control all network stations, all lines, and even the communication partner and so absolute anonymity is theoretically impossible. Therefore we need reasonable models of possible attackers.

There are several possible attackers: the administration, foreign states, companies, one's neighbours and communication partners. During the design of an anonymous network these possible attackers have to be translated into terms of stations and lines. A station is always under control of its owner and might be under control of everybody who has had access to it so far, e.g. its manufacturer, because he might have installed a Trojan Horse [24, 29]. Trojan Horses are a serious problem in stations with high complexity, e.g. switching centers. In simple user stations they can be detected (if tried) more easily. Lines are assumed to be owned by the PTT. Normally they can easily be observed by the PTT or an eavesdropper, but by physical measures such an attack can be made much more difficult.

Given a model of the attacker we have to define what we want to keep hidden from him. A strong possibility is to keep the sender and the recipient of a message secret. A weaker possibility is to keep only their relationship secret, i.e. sending and receiving of physical messages is observable, but it is infeasable for an attacker to link the physical message sent by the sender and the physical message received by the recipient.

1.2 Recipient anonymity

Receiving a message can be made completely anonymous to the network by delivering the message to all stations (broadcast). If the message has an intended recipient, a so called addressee, it has to contain an attribute by which he and nobody else can recognize it as addressed to him [12]. This attribute is called an implicit address in contrast to an explicit address, which describes a place in the network.

Implicit addresses can be distinguished according to their visibility, i.e. whether they can be tested for equality or not. An implicit address is called invisible, if it is only visible to its addressee and is called visible otherwise [30].

Invisible implicit addresses can be realized with a public key cryptosystem. A message is addressed by encrypting it (or a part of it) with a public key of the addressee. Each station decrypts all messages with each of its private keys and uses the message redundancy to decide which messages are addressed to it.

Conversely, any invisible addressing scheme can be used for public key distribution: If A wants to communicate an n bit key to B, A chooses n messages randomly, and addresses them to B if the corresponding key bit is 1, and addresses them not to B otherwise. A sends these n messages in one explicitly addressed message to B. This addressee-anonymity-based public key distribution protocol corresponds to a sender-anonymity-based one proposed by Alpern and Schneider [1].

If a secret key of a faster conventional cryptosystem has already been exchanged, that key can also be used for invisible implicit addressing [17 pp. 111..112]. If this is done to save decipherment cost, each packet should start with a bit telling which cryptosystem is used for invisible implicit addressing.

Visible implicit addresses can be realized much easier: Users choose arbitrary names for themselves, which can then be prefixed to messages.

Another criterion to distinguish implicit addresses is their distribution. An implicit address is called public, if it is known to every user (like telephone numbers today) and private if the sender received it secretly from the addressee either outside the network or as a return address or by a generating algorithm the sender and the addressee agreed upon [12, 17].

Public addresses should not be realized by visible implicit addresses to avoid the linkability of the visible public address of a message and the addressed user.

Private addresses can be realized by visible addresses but then each of them should be used only once.

All possible combinations of addressing modes and address distribution are shown in figure 1.

1.3 Unlinkability of sender and recipient

This form of anonymity can be realized by a special network station, a so called MIX, which collects a number of messages of equal length from the senders, discards repeats, changes their encodings, and forwards the messages to the recipients in a different order (figure 2).

This measure hides the relation between sender and recipient of a message from everybody but the MIX and the sender of the message. By using more than one MIX to forward a message from the sender to the recipient, the relation is hidden from every attacker in the network who does not control all MIXes which the message passed, nor has the cooperation of the sender [5].

Care must be taken that an attacker cannot submit all messages which are mixed at the same time except for one. Otherwise he might link the message he did not send, too.

1.4 Sender anonymity

A powerful scheme for sender anonymity is superposed sending which is published in [8, 10] and is called DC-network (dining cryptographers network) there.

Each user station generates at least one key bit for each message bit and sends each key bit to exactly one other user station over a secure channel. To send one bit every user station adds modulo 2 (superposes) all generated and received key bits and its message bit if there is one. The sums are sent over the network and added up modulo 2. The result is distributed to all user stations. The result is the sum of all sent message bits, because every key bit was added twice (figure 3). Therefore the scheme realizes a multi-access channel with collisions. For its efficient use a medium access protocol [28] preserving anonymity is needed. Two of them are mentioned in [8].

If an attacker controls all lines and some of the user stations, he gets no information about the sender of a message among the other users, as long as their key graph, i.e. the graph with the users as nodes and the keys as edges, is connected.

Superposed sending requires the exchange of a tremendous amount of randomly chosen keys. To reduce costs, pseudorandomly generated keys can be used instead, reducing information-theoretic [26] to complexity-theoretic security.

The expensive generation, distribution and superposing of keys (and messages) of the concept of superposed sending can be avoided if the network is designed for preventing attackers from physically observing all lines connecting a user with the rest of the world.

A simple and efficient way to do so is to connect the user stations by RINGs, which are in wide use for local area networks. If an anonymous medium access protocol is used, a user station is only observable if its two neighbour stations collude or the lines are tapped. The latter attack can be prevented by running the cable in an appropriate way [20, 21]. Then, it is approximately as difficult to observe the sending of a user station as to observe its owners behavior at home directly by hidden microphones, laser-based bugs or EMI emissions of the user station [15, 16, 11].

Anonymous medium access protocols are slotted ring with sender remove and token ring, both with exhaustive service [13, 14]. This is illustrated in figure 4.

2 Performance

The two main performance characteristics of networks are throughput and transfer delay. Their importance depends on the services the network is supposed to offer. Throughput and transfer delay are less critical for services like electronic mail, only throughput is critical for services like file transfer, only transfer delay for services like telephony and both are critical for video telephony.

2.1 Some remarks on the basic concepts for anonymous networks

Analyzing the performance of the concepts of section 1 must go along with considering how they would be implemented physically.

In local areas with a few hundred stations the performance of a RING-network implemented as a physical ring is roughly as good as or even better than that of an equally expensive usual star or bus network [3, 4, 18]. However, performance and reliability of RING-networks with more than 10000 stations become unacceptable.

In [8], David Chaum suggests implementing superposed sending on a physical ring network. Each message bit requires two circulations around the ring: in the first round, the user bits are successively superposed by the users, in the second round, the resulting bit is broadcasted.

This implementation seems quite efficient because under the assumption of uniformly distributed traffic it increases the average expenditure of transmission only by a factor of four compared with a traditional ring access protocol in which the recipient removes the message from the ring, whereas on a star or tree network the factor is the number of stations. But the amount of transmission on each line, i.e. the required bandwidth, is the same for all implementations, so implementations on stars or trees might still be better if their transfer delay is shorter. The nodes of such networks can be less complex than normal switching centers and may be constructed in a way that the overall transfer delay in the network is only proportional to the logarithm of the number of stations, whereas in ring networks it is always proportional to the number of stations [23].

As throughput and reliability of any network based on superposed sending cannot be superior than that of a RING-network, these networks cannot be built with more than 10000 stations either.

In the MIX-network, several factors are to be considered: How many and which stations act as MIXes and how many MIXes are used per message?

The message length grows proportionally with the number of MIXes chosen: to disable an attacker to reproduce the change of encoding of the message, it must include about 100 random bits for each MIX, which the deciphering MIX discards. Therefore, the expenditure of transmission of a message grows quadratically with the number of MIXes chosen for it. So this number must not be too large. Especially, not all stations can be chosen as MIXes for all messages.

To guarantee short transfer delay for time critical services, the throughput of a station that acts as a MIX must be very high because it always has to have enough messages to mix. These messages must be decrypted, rearranged, and forwarded. Thus, a MIX should be extremely powerful and complex, and therefore will not be cheap. Consequently, only a limited number of MIXes can be afforded in the network.

If the MIX-network is implemented using some user stations of an existing physical network as MIXes, each message must pass through the physical network several times, which adds additional delay to that occuring in the MIXes. But using the switching centers of the physical network as MIXes can not be recommended either, because the probability that they collude is too great (and the assumption that they are independent becomes altogether absurd in countries with a telecommunication monopoly like the FRG).

2.2 Hierarchical networks

As mentioned above, networks which provide sender and recipient anonymity cannot be built for the number of stations an ISDN would have. To achieve high performance, it seems reasonable to divide the network stations statically or dynamically into groups which perform one of the schemes of paragraph 1.4 and to support the possible groupings by a physical structure.

The simplest form of such a structure is the switched/broadcast network (SBNS) which has two levels, broadcast networks based on RINGs or superposed sending at the lower level and an arbitrary switched network as backbone (figure 5) [19, 20, 21, 22, 23]. If the scheme of superposed sending is used, the SBNS can easily be generalized to a tree network. The partitioning into local broadcast networks can then be made variable by changing the depth of the backbone network [23].

2.3 Channel switching

So far only networks based on slotted rings with exhaustive service are suitable for services that require a continuous stream of information with short transfer delay (channel switching) because once a station is allowed to use a slot, it can use this slot again and again as a channel.

The MIX-network is inappropriate for such services because of the delay during the transport of each message. The networks based on the concept of superposed sending are of limited use because the basic medium access protocols do not guarantee synchronous service (called isochronous in [27 p. 819]).

New possibilities of increasing the performance of these network are obtained by dropping one requirement for anonymity that seems unreasonable for channel switching services anyway: the requirement that the relationship between different messages of the same connection is hidden [23].

In a network based on superposed sending, channels can then be switched as in normal broadcast networks.

In a MIX-network in its pure form, the delay results essentially from the fact that every MIX has to await all bits of a long packet before it can decrypt it and send the first bit to the next MIX. This can be avoided if a single message is used for setting up a connection and giving each MIX a key of a fast private key system used as a stream cipher. These private keys are used to encrypt the following messages of the initiated connection just like the public keys in the normal MIX-network [23].

In a hierarchical network, channels are switched by concatenating channels of the different levels of the hierarchy.

3 Fault tolerance

So far, all networks are serial systems in the sense of reliability: all MIXes of a chosen sequence of MIXes, all stations of a RING, and all stations taking part in superposed sending have to work correctly. To fulfil the high reliability requirements on an ISDN, each scheme must be extended to include some fault-tolerance mechanisms. These mechanisms may work end-to-end, i.e. the sender retransmits a message if it does not receive an acknowledgement after a certain period of time. Even if the sender chooses a different encoding of the message for each retransmission, the retransmitted messages may enable statistical attacks in some networks. Moreover, the performance of such mechanisms in terms of average transfer delay, variance of transfer delay, or usable throughput may be unsatisfactory. Therefore, it seems worthwhile to use mechanisms which avoid end-to-end retransmission wherever possible.

3.1 MIX-network

There are two possibilities to avoid end-to-end retransmission in the MIX-network: MIXes may choose backups for itself or the user of the network may enable the MIXes to bypass MIXes which broke down.

Both possibilities are especially useful in the case of return addresses where the end-to-end timeout and retransmit approach does not work satisfactorily for all services (e. g. electronic mail).

To realize the first possibility, the private key of a MIX is distributed to its backups. If a certain amount of time can be spent on establishing the backup service, a threshold scheme [25] may be used. MIXes sharing a key must be coordinated to mix each message at most once [23].

If every MIX in a sequence of chosen MIXes can bypass the next MIX, a failure of one MIX (or more, as long as no two consecutive MIXes break down) is tolerable. As in the first possibility, a coordination protocol must guarantee that each message is mixed at most once [23].

To bypass one MIX, its predecessor must not only get the message part for it but also for its successor (figure 6). If it receives both message parts and this is done for every MIX, the length of the whole message grows exponentially. To avoid this, the sender of a message chooses a different key (e.g. of a fast private key system) for each MIX. Together with its message part, each MIX has to get its key, that of its successor, and the addresses of the next two MIXes, all together encrypted with its own public key.

Let A1,...,An be the sequence of addresses and e1,...,en the sequence of public keys of the chosen MIXes MIX1,...,MIXn; An+1 the address of the addressee (called MIXn+1 for convenience), and en+1 his public key; k1,...,kn the chosen sequence of keys, and Mi the message that MIXi shall receive. The messages Mi are formed according to the following scheme, starting from the message content M that MIXn+1 shall receive:

Mn+1 = en+1(M)
Mn = en(kn,An+1),kn(Mn+1)
Mi = ei(ki,Ai+1,ki+1,Ai+2),ki(Mi+1) i=1,...,n-1

So MIXi can compute Mi+1 and Mi+2 out of Mi, but as long as at least two consecutive MIXes are not controlled by the attacker, the scheme is as secure as the original one [23]. It is not necessary that both consecutive uncontrolled MIXes are really passed by the message. One may be down or intentionally left out to increase performance: Whenever MIXi+2 is up, MIXi sends Mi+2 to it. Only if MIXi+2 is down, MIXi sends Mi+1 to MIXi+1. Compared with the obvious solution where MIXi sends Mi+1 to MIXi+1 first, this saves transmission cost, increases throughput and decreases transfer delay in the average (figure 7).

The scheme can easily be modified to tolerate the failure of d consecutive MIXes instead of one for every fixed number d.

3.2 Other networks

The RING-network can be made fault tolerant by using a braided ring (figure 8) and special protocols [18] (a quantitative examination of the reliability improvement is given there).

Some remarks on the DC-net and the hierarchical anonymous networks can be found in [23].

4 Concluding remarks

This paper dealt with the design of a network with high performance and reliability which allows its users to send and to receive anonymously. As mentioned in section 0, the content of a message can be hidden by using end-to-end encryption.

If using the network is not for free, the charges must either be paid anonymously with each use of the network (e.g. by anonymous numbered accounts [20, 21] or digital banknotes [9, 10]), which seems rather troublesome, or measured anonymously (e.g. by safeguarded counters at user stations [20, 21]), or paid by flat rates.

The initially mentioned services like electronic mail, ordering of newspapers or home banking can be implemented by higher protocols on top of such a network.

If identification is required instead of anonymity, the well known authentication schemes can be used. Otherwise it is necessary to implement the services in a way which preserves the anonymity of the network. This must be proved in addition to proofs that the implementation fulfils its normal specification, e.g. security against fraud [31].

It should be mentioned that many communication services where users nowadays have to identify themselves can be used in an anonymous way in the future if there is a protocol that allows people to act under several pseudonyms and to transform documents that carry one of these pseudonyms into documents carrying another of their own pseudonyms, in a secure and anonymous way [6, 7, 10].

Acknowledgements

We are grateful to David Chaum for sending us his drafts and to him, Klaus Echtle, and Birgit Pfitzmann for a lot of useful comments and stimulating discussions. Dr. Klaus Dittrich read and polished this text very carefully.

Literature

[1] Bowen Alpern, Fred B. Schneider: Key exchange Using 'Keyless Cryptography'; Information Processing Letters Vol. 16, 26 February 1983, pp. 79..81

[2] P. Baran: On Distributed Communications: IX. Security, Secrecy, and Tamper-Free Considerations; Memorandum RM-3765-PR, Aug. 1964, The Rand Corporation, Santa Monica, California

[3] G. Bürle: Leistungsvergleich von Sternnetz und Schieberegister-Ringnetz; Studienarbeit, Univ. Karlsruhe, 1984

[4] G. Bürle: Leistungsbewertung von Vermittlungs-/Verteilnetzen; Diplomarbeit, Univ. Karlsruhe, Mai 1985

[5] D. Chaum: Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms; CACM Vol. 24, Nu. 2, Feb. 1981, pp. 84..88

[6] D. Chaum: A New Paradigm for Individuals in the Information Age; Proc. of the 1984 Symp. on Security and Privacy, IEEE, Apr. 1984, Oakland, California, pp. 99..103

[7] D. Chaum: Showing Credentials Without Identification. Signatures Transferred Between Unconditionally Unlinkable Pseudonyms; Eurocrypt 85, Draft, received May 13, 1985;

[8] D. Chaum: The Dining Cryptographers Problem. Unconditional Sender Anonymity; Draft, received May 13, 1985;

[9] D. Chaum: Privacy Protected Payments. Unconditional Payer and/or Payee Anonymity; Draft, received May 13, 1985;

[10] David Chaum: Security Without Identification: Transaction Systems to Make Big Brother Obsolete; CACM Vol. 28, Nu. 10, Oct. 1985, pp. 1030..1044

[11] Wim van Eck: Electromagnetic Radiation from Video Display Units: An Eavesdropping Risk? Computers & Security Vol. 4, Nu. 4, December 1985, pp. 269..286

[12] D. J. Farber, K. C. Larson: Network Security Via Dynamic Process Renaming; Fourth Data Communications Symp., Oct. 1975, Quebec City, Canada, pp. 8-13..8-18

[13] G. Höckel: Untersuchung der Datenschutzeigenschaften von Ringzugriffsmechanismen; Diplomarbeit, Univ. Karlsruhe, Aug. 1985

[14] G. Höckel, A. Pfitzmann: Untersuchung der Datenschutzeigenschaften von Ringzugriffsmechanismen; 1. GI-Fachtagung "Datenschutz und Datensicherung", Okt. 1985, München, IFB Band 113, Springer-Verlag, Heidelberg, pp. 113..127

[15] John Horgan: Thwarting the information thieves; IEEE Spectrum Vol. 22, Nu. 7, July 1985, pp. 30..41

[16] John Horgan: Inventor seeks to warn Government of threat from laser-based bug; The Institute, IEEE, October 1985, p. 8

[17] P. A. Karger: Non-Discretionary Access Control for Decentralized Computing Systems; Master Thesis, MIT, Laboratory for Computer Science, May 1977, Report MIT/LCS/TR-179

[18] A. Mann: Fehlertoleranz und Datenschutz in Ringnetzen; Diplomarbeit, Univ. Karlsruhe, Okt. 1985

[19] A. Pfitzmann: Ein Vermittlungs-/Verteilnetz zur Erhöhung des Datenschutzes in Bildschirmtext-ähnlichen Neuen Medien; 13. Jahrestagung der GI, Okt. 1983, Univ. Hamburg, IFB Band 73, Springer-Verlag Heidelberg, pp. 411..418

[20] A. Pfitzmann: Ein dienstintegriertes digitales Vermittlungs-/Verteilnetz zur Erhöhung des Datenschutzes; Fak. f. Inform., Univ. Karlsruhe, Interner Bericht 18/83, Dez. 1983

[21] A. Pfitzmann: A switched/broadcast ISDN to decrease user observability; 1984 Intern. Zurich Seminar on Digital Communications, March 1984, Zurich, Switzerland, Swiss Federal Inst. of Tech., Proc. IEEE Cat. No. 84CH1998-4 pp. 183..190

[22] A. Pfitzmann: Technischer Datenschutz in diensteintegrierenden Digitalnetzen - Problemanalyse, Lösungsansätze und eine angepaß Systemstruktur; 1. GI-Fachtagung "Datenschutz und Datensicherung", Okt. 1985, München, IFB Band 113, Springer-Verlag, Heidelberg, pp. 96..112

[23] A. Pfitzmann: How to implement ISDNs without user observability - Some remarks; Fak. f. Inform., Univ. Karlsruhe, Interner Bericht 14/85, 1985

[24] G. J. Popek, C. S. Kline: Issues in Kernel Design; Operating Systems, An Advanced Course, Ed. by R. Bayer et. al.; LNCS 60, 1978; Springer-Verlag, Heidelberg, pp. 209..227

[25] Adi Shamir: How to Share a Secret; CACM Vol. 22, Nu. 11, November 1979, pp. 612..613

[26] C. E. Shannon: Communication Theory of Secrecy Systems; Bell Syst. Tech. J., Vol. 28, No. 4, Oct. 1949, pp. 656..715

[27] Daniel T. W. Sze: A Metropolitan Area Network; IEEE Journal on Selected Areas in Communications Vol. SAC-3, No. 6, November 1985, pp. 815..824

[28] A. S. Tanenbaum: Computer Networks; Prentice-Hall, Englewood Cliffs, N. J., 1981

[29] K. Thompson: Reflections on Trusting Trust; CACM, Vol. 27, No. 8, Aug. 1984, pp. 761..763

[30] M. Waidner: Datenschutz und Betrugssicherheit garantierende Kommunikationsnetze. Systematisierung der Datenschutzmaß und Ansätze zur Verifikation der Betrugssicherheit; Diplomarbeit, Fak. f. Inform., Univ. Karlsruhe, Interner Bericht 19/85, Aug. 1985

[31] M. Waidner, A. Pfitzmann: Betrugssicherheit trotz Anonymität. Abrechnung und Geldtransfer in Netzen; 1. GI-Fachtagung "Datenschutz und Datensicherung", Okt. 1985, München, IFB Band 113, Springer-Verlag, Heidelberg, pp. 128..141; Revised version appeared in DuD, "Datenschutz und Datensicherung, Informationsrecht, Kommunikationssysteme", Vieweg Verlag, Wiesbaden, Heft 1, Februar 1986, pp. 16..22

und Datensicheru•“x´s¸pÛk kh sc ˜` ü[-X US ‹P ÚK î • î î • î î • î î • î î • î î î