

- [3] ISO. *IS 10116. Modes of operation for an n-bit block cipher algorithm*. ISO/IEC JTC1/SC 27, 1991.
- [4] ISO. *IS 10181-2. Security Frameworks in Open Systems – Part 2: Authentication Framework*. ISO/IEC JTC1/SC 21, 1991.
- [5] ISO/IEC. *IS 10167. Guidelines for the application of Estelle, LOTOS and SDL*. ISO/IEC, 1991.
- [6] ITU-T. *Recommendation Q.24, Multifrequency Push-Button Signal Reception*. International Telecommunication Union, Telecommunication Standardization Sector, 1988,1993.
- [7] Xuejia Lai. Markov ciphers and differential cryptanalysis. In *Proceedings of EUROCRYPT'91*, 1991.
- [8] National Bureau of Standards (U.S.). *DES modes of operation*. Federal Information Processing Standards, Publication 81, National Technical Information Service, Springfield, VA, Dec. 1980.
- [9] W. Tuchman. Hellman presents no shortcut solutions to DES. *IEEE Spectrum*, July 1979.

5.4. Key Distribution

As already mentioned, the key distribution can be made simple and effective with smart cards which also contain the DES algorithm. The cards are customised by the operator and sent to each individual subscriber. The key stored in the card is uniquely related to the subscribers logical directory number, which again is uniquely related to the physical line. A smart card can thus only be used by the subscriber it is intended for. However, the keys must be protected so that an intruder can not misuse it to fabricate a counterfeited smart card. Counterfeits will, if successful, allow an intruder to connect to the line and masquerade as the real subscriber in the same way as shown in Figure 1.

5.5. Compatibility and Feature Interaction

The AD is best suited to be used together with DTMF terminal equipment, but it will still be possible to use it with rotary dial sets or other older equipment using register signalling. It would then be necessary to define the state “Wait for Challenge” from Figure 8 to also be transparent so that conversation can take place. This is because rotary sets use alternated high and low impedance register signalling, which would cause the AD to jump between the first two states and then finally end up in “Wait for Challenge” when the last digit has been dialled. In addition, the local exchange must be able to receive both rotary and DTMF signalling during the same call.

Most telephone exchanges allow DTMF equipment to use the “Recall” function which in fact is a calibrated period of high impedance. The Recall function can be used for many purposes. Sometimes a Recall alone can activate a service, and in other cases the subscriber receives a special dialtone as invitation to dial a new directory number or a special service code. Because the AD will be in “Wait for Challenge” after a Recall, it is now possible to challenge the AD for each new service or new directory number called, which would make it impossible for an intruder to take over a call which is already established, and use it to make further calls.

6. Conclusion

We have analysed the problem of theft (misuse or resources) in analogue telephone access networks, and described how to solve the problem by introducing authentication of the legitimate subscriber by an authentication module. By considering system constraints and security requirements, the adequate cryptographic mechanisms could be determined, and a practical implementation of the scheme has been described in some detail. This study is the solution to a concrete and identical problem in many countries. If operators cooperate, manufacturers will be able to deliver similar equipment to all of them, and thereby be able to offer an affordable solution to the operators.

7. References

- [1] ANSI. *American National Standard for Data Encryption (DEA), or Data Encryption Standard (DES)*. American National Standards Institute, New York, 1981. ANSI X3-92-1981.
- [2] D Coppersmith. The data encryption standard (DES) and its strength against attacks. Technical report, T.J.Watson Research Center, December 1992. IBM Research Report RC 18613 (81421).

with the signals it can send and receive are shown in Figure 1.

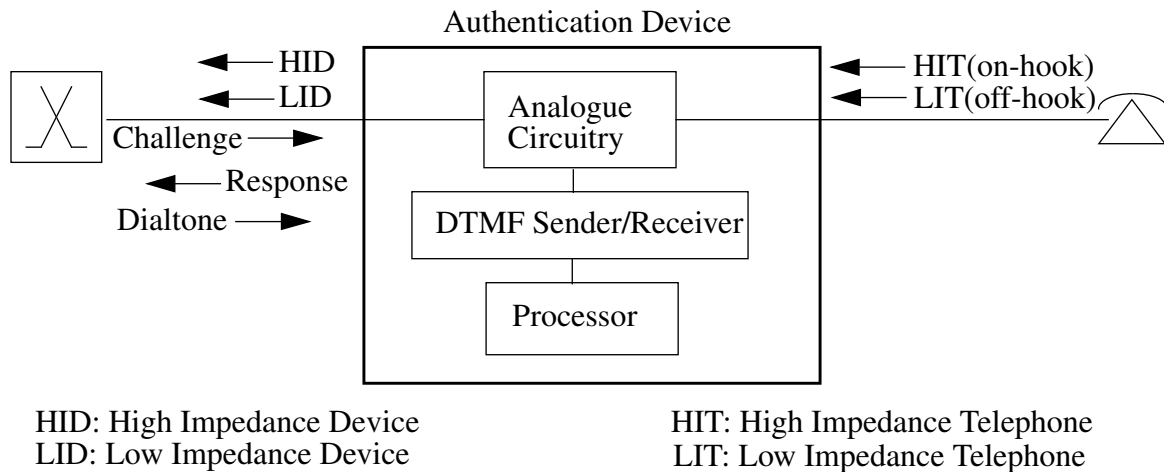


Figure 7. Authentication Device functional blocks

The processor can be a smart card so that the AD itself does not contain any processor, but only a smart card reader. The advantage of using a smart card as processor are twofold: Standard smart cards with the DES are available at a very low price, and secondly, it facilitates key distribution. The functionality of the DTMF Sender/Receiver and of the Analogue Circuitry is found in mass consumption telephony devices such as automatic answering machines and number diallers, so the building blocks of the AD are already being produced today. Taking into consideration that the AD itself will be manufactured in large numbers, it should be possible to reach an acceptably low production cost.

The state diagram of the AD is shown in Figure 8 on the right, where the description language used is SDL [5]. The AD is considered as a “black box” so that only its external behaviour is observed. At off-hook, the telephone set changes from high to low impedance. This is detected by the AD which also changes from high to low-impedance. When the telephone exchange detects the low-impedance from the AD, it initialises the challenge-response protocol. When the AD detects the challenge, it computes the response which is returned to the local exchange. The AD then enters Transparent Mode, i.e. it does not interfere with the analogue signals passing between the telephone and the exchange. It only exits the transparent mode again if high-impedance, caused by on-hook or a recall at the telephone set, is detected.

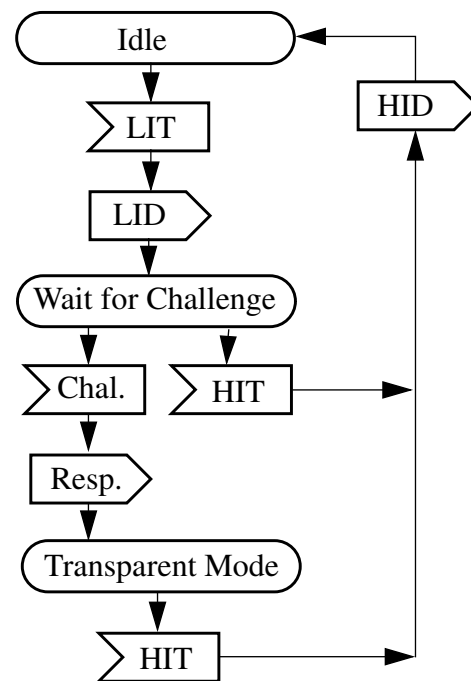


Figure 8. AD State Diagram

5.1. Cryptographic Mode of Operation

The simplest possible implementation is obtained if initialisation and synchronisation of the cryptographic algorithm can be avoided. ECB mode [3,8] has this property, and we will describe in detail how it can be adapted for our purpose. In the previous section, we found $m = 2$ byte and $k = 14$ byte. Since DES has a block length of 8 byte, and our message length is only 2 byte, the block must somehow be padded. As the simplest solution we propose to let the authentication device repeat 4 times the received challenge of 2 byte, in order to obtain a block of 8 byte. DES also produces a response of 8 byte, but in the previous section, we concluded that only a 2 byte response is needed. We can get the 2 byte by picking 16 bit from predetermined positions, e.g. the two first byte in direct order. In the previous section, we have already foreseen the risk of an intruder randomly guessing the right response out of 2^{16} , so this does not add to the vulnerability. The cryptographic scheme is shown in Figure 1.

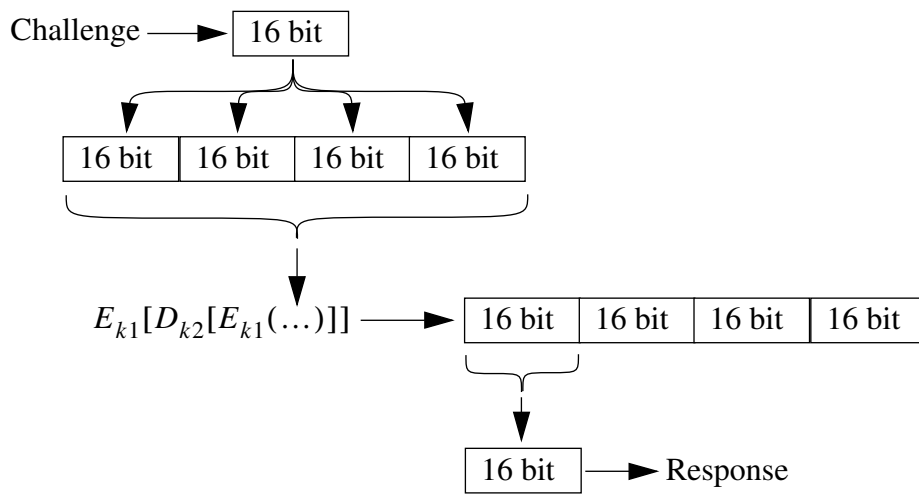


Figure 6. Cryptographic Mode of Operation

5.2. Authentication Software Module

At the central office, the functionality for subscriber authentication can be provided by software alone, since every modern telephone exchange is equipped with hardware for DTMF-signal sending and reception. The computation of the expected response can be done during challenge transmission and response reception so that no extra delay is added to the call setup time. However, DES computation in software may be a bottleneck in case of high processor load. If high processor load is a problem, we propose two alternative solutions. Firstly, the challenge-response pairs can be precomputed off-line (e.g. during otherwise idle periods), so that no extra processor capacity is needed during call setup, but the disadvantage is the extra storage capacity needed for the precalculated challenge-response pairs. The second alternative is to provide specialised hardware for DES encryption so that it can be done in real time during call setup without extra load to the processor. The disadvantage is that new hardware must be installed in every telephone exchange.

5.3. The Authentication Device

For the design of the AD, great care must be taken to limit the production cost. Each device may not be very expensive, but when multiplied by the number of subscribers using it, the total sum quickly becomes very high. The functional blocks of the authentication device together

4. Compute the maximum message space usage

Assume that only half the challenge-response pairs shall be used during the lifetime of a single key. Hence, from (4) we have another constraint on the minimal message length

$$m = \log_2\left(\frac{r}{U_m}\right) = \log_2\left(\frac{36500}{0.5}\right) = 16.1 \text{ bit}$$

5. Check solution for consistency

In item (1) we found the maximum message length to be 16 bit while 16.1 bit is required as from item (4). However, when we estimated the required message space usage, we assumed that the subscriber made 10 calls per hour, every hour, every day for a total of 10 years. This is a very conservative estimate. Hence, we accept a message length of 16 bits.

6. Determine the required key length

We have assumed DES or IDEA as the best choices of algorithm by considering the requirements of low price, high speed, availability and security. Due to U.S. export restrictions, IDEA may be the better choice for global use. However, the following discussion assumes DES as the features of DES are well known.

Exhaustive key search on simple DES is today within reach of commercial general purpose computing equipment in a few weeks, if not, within a few days time. Double DES with two different keys is vulnerable to meet-in-the-middle attacks, and although such an attack is probably not practically feasible, Triple DES using two different keys[9] is considered to be much stronger. Currently, there are no practical cryptanalytic attack on Triple DES. By assuming that Single DES can be broken in 1 day, and that Triple DES increases the search time by a factor in the order of 2^{56} , compared to Single DES[2], the time needed for exhaustive search would be approximately 10^{14} years. This provides more than enough strength, and we therefore propose to use a Triple DES with a double key of length $k_{min} = 112$ bit, as a conservative choice.

7. Determine maximal key length from storage constraints

A key of 112 bit = 14 byte is easily stored in the authentication device, and the only difficulty may be to store the key material for thousands of subscribers at the central office. A local exchange with 100.000 subscribers connected is a very large exchange, and by assuming that every subscriber is authenticated, the exchange must store 1,4 Mbyte key material. This is in reality a small amount, and for a telephone exchange which typically uses multiprocessor and distributed computing techniques, the 1,4 Mbyte can easily be subdivided and spread over all the line connection PCBs or similar devices. A line connection PCB typically serves 8, 16 or 32 subscribers, so the amount of key material in each gets very low. In reality, it is impossible to fix a general k_{max} for a modern telephone exchange, but that it is larger than 14 byte is certain.

8. Check solution for consistency

Since $k_{max} > 14$ byte, we conclude that $k_{min} < k_{max}$ is satisfied.

5. Implementation

In this section we will briefly discuss some aspects related to the practical implementation of the scheme.

6. Define as a security requirement the minimum accepted period for successful exhaustive key search. A number of assumptions have to be made about the available resources the attacker has at his disposal, and about the time it takes to check a single key. Using a well known algorithm where test data are available will make this task easier. This step will indicate the minimum key length k_{min} [bit].
7. Determine the available memory for key storage and estimate the maximum number of users in the system. This leads to a maximum key length k_{max} [bit].
8. Select a key length, k [bit], such that the condition $k_{min} < k \leq k_{max}$ is satisfied.

4.3. Application of procedure

1. Available bit rate calculation and maximal message length

The current QoS requirements should be met also by the protected access protocol, i.e. the modification should be transparent to the user. In existing networks, the telecom provider usually guarantees dial tone within 2s (or a similar number) of off-hook. The protocol must be executed within this interval. However, this interval is designed to allow the local exchange sufficient processing time to detect off-hook, allocate and set-up tone-analysers and tone-generators etc. In practice, we cannot allow more than a fraction of the interval to be used by the authentication protocol. A reasonable estimate is 1.0 second.

According to [6], the maximal signalling velocity of push-button dialling is roughly 100 ms per digit or 40 bit/s.

Assuming 200ms for total processing (both the LE and the AD) we find from (1) that the maximal message length

$$m_{max} = (T_e - T_c) \times \frac{q}{2} = (1s - 0.2s) \times \frac{40\text{bit/s}}{2} = 16\text{bit}$$

2. Random response and minimum message length

The intruder can guess a response and we require that the probability of correctly guessing the response to be small. Let us assume that this probability shall be less than $P_r = 10^{-4}$. This value is similar to what is currently used by challenge-and-response login solutions. Hence, we find that the minimum message length as defined in (2) gives

$$m_{min} = \log_2\left(\frac{1}{P_r}\right) = \log_2\left(\frac{1}{10^{-4}}\right) = 14 \text{ bit}$$

3. Usage frequency

Assume that each subscriber uses his/her telephone 10 times a day. Further, assume that the lifetime of a single key shall be 10 years¹. The number of challenge-response pairs observed is defined in (3) and is equal to

$$r = f \cdot T_k = \frac{10}{24 \cdot 3600} \times (10 \times 365 \times 24 \times 3600) = 36500$$

1. In order to minimize the key management, the lifetime of each subscriber key shall be as long as possible. We adopt a key lifetime of 10 years as a practical estimate of the maximum lifetime of a single subscription.

4.2. Procedure for Determining Message and Key Lengths

In this section we describe a procedure for specifying a particular challenge response authentication scheme, based on the external system constraints and security requirements, in more detail.

We make the idealistic assumption that cryptographic algorithms with flexible message length and key length are available, and analyse the constraints which will influence the message and key lengths.

1. Determine the available bit rate q in the communication medium where the authentication scheme is to be implemented. Define the maximum acceptable execution time $T_e = T_t + T_c$ where T_t expresses the total transmission delay, and T_c the total computation delay. We assume T_c given whereas the transmission delay is expressed as $T_t = 2 \cdot m/q$ where m is the message length. Since these timing constraints impose a maximum allowed message length m_{max} , which can be expressed as

$$m_{max} = (T_e - T_c) \times \frac{q}{2} \quad [\text{bit}] \quad (1)$$

2. Define as a security requirement the maximum probability P_r of correctly guessing the response when guessing at random. This is determined by the message space size M and can be expressed as $P_r = 1/M$. This imposes a minimum message length m_{min} which is expressed by

$$m_{min} = \log_2 \left(\frac{1}{P_r} \right) \quad [\text{bit}] \quad (2)$$

3. Estimate the usage frequency f as the number of authentications per second. Then define the maximum key life time T_k in number of seconds. The number r of observed challenge-response pairs that an attacker can observe during the lifetime of a key is then expressed by

$$r = f \cdot T_k \quad (3)$$

4. Define as a security requirement the maximum message space usage U_m as the ratio of maximum number r of observed challenge-response pairs over the total message space M . We must avoid exhausting the message space by a large margin. If challenge collisions can be avoided by e.g. using incrementing values, we no longer need to worry about replay attacks, and the whole message space can in principle be exhausted before a key is reused. If on the other hand random challenges are used, very low message space usage must be foreseen in order to get a very low probability of challenge collisions and low risk of replay attacks. Having defined U_m , we can calculate the required message length m

$$m = \log_2 \left(\frac{r}{U_m} \right) \quad [\text{bit}] \quad (4)$$

5. Check that the condition $m_{min} \leq m \leq m_{max}$ is satisfied.

4. Constraints Considerations for Determining Key and Message Lengths

Having decided upon an authentication mechanism and a protocol, we will in this section determine the appropriate message length and key length. For this purpose we use a simple model and employ a procedure which will give the required message and key lengths. The model and procedure are first explained and then applied to our application.

4.1. Constraint and Requirement Model

The influence that security requirements and system constraints have on message length and key length is graphically illustrated in Figure 1.

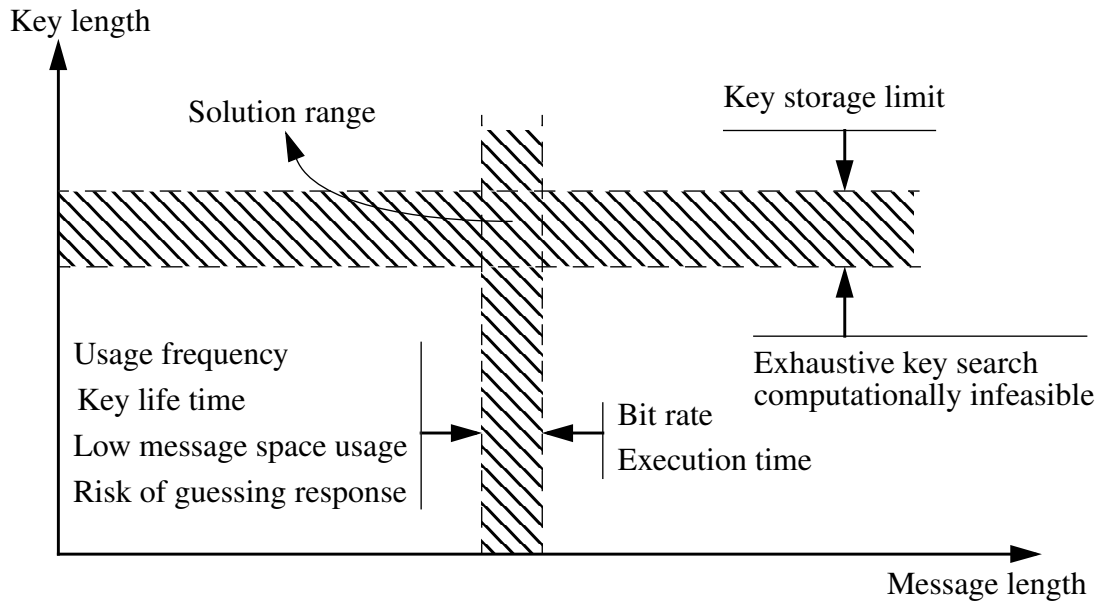


Figure 5. Constraint and requirement analysis

The main security requirements for the message length is to keep a low message space usage. The expected usage frequency together with the life time of a key determines the number of challenge-response pairs that an attacker can observe. This together with the defined message space usage determines the required message length. If the message is very short, there is a risk that responses can be correctly guessed at random. A maximum probability of that happening imposes a minimum message length. However, the available bit rate and execution time will impose a maximum message length.

The main security requirement for the key length is that exhaustive key search must be computationally infeasible, and this imposes a minimum key length. On the other hand short keys are easier to manage than large keys, and the available key storage space may impose an upper limit on the key length.

Figure 1 shows that there is a solution range for both the message and key lengths, and this is a condition for the implementability of an authentication scheme. If the constraints lead to incompatible upper and lower bounds for either the message length or the key length, the application can not be implemented. If that is the case, the figure clearly illustrates what the problem might be and the possible modifications to the system specification that will solve the problem

very vulnerable to replay attacks.

Class 3 provides protection against replay attacks. More precisely it uses a challenge-response scheme, and is specified with 3 message transfers; 1) Authentication Request, 2) Challenge, 3) Response. In our case, the physical access port in the exchange identifies the customer. Hence, we do not need to send the first message to communicate the identity of the customer as part of the authentication.

Zero-Knowledge techniques is useful in situations where the verifier could misuse the authentication information to masquerade as claimants. In our case, the operator is the verifier, and he has no interest of masquerading as customer. Zero-Knowledge would therefore be inappropriate in our case.

Authentication by challenge-response is generally useful when authentication takes place across an insecure channel because it provides dynamic passwords and thus resistance to simple password interception. When using cryptographic block algorithms in a challenge-response protocol, the equivalents of ciphertext and cleartext are both transmitted across the insecure channel. The cryptographic algorithm must therefore be resistant to known-plaintext attacks, and make exhaustive key search the best attack. Using a sufficiently large key spaces then finally makes exhaustive search computationally infeasible so that the algorithm effectively becomes resistant to all imaginable practical attacks. When selecting an algorithm, speed, price, availability and strength must be considered. A well known symmetric block cipher algorithm such as DES [1] or IDEA [7] is therefore a natural choice.

The proposed protocol between the customer, the authentication device and the local exchange is shown in Figure 1.

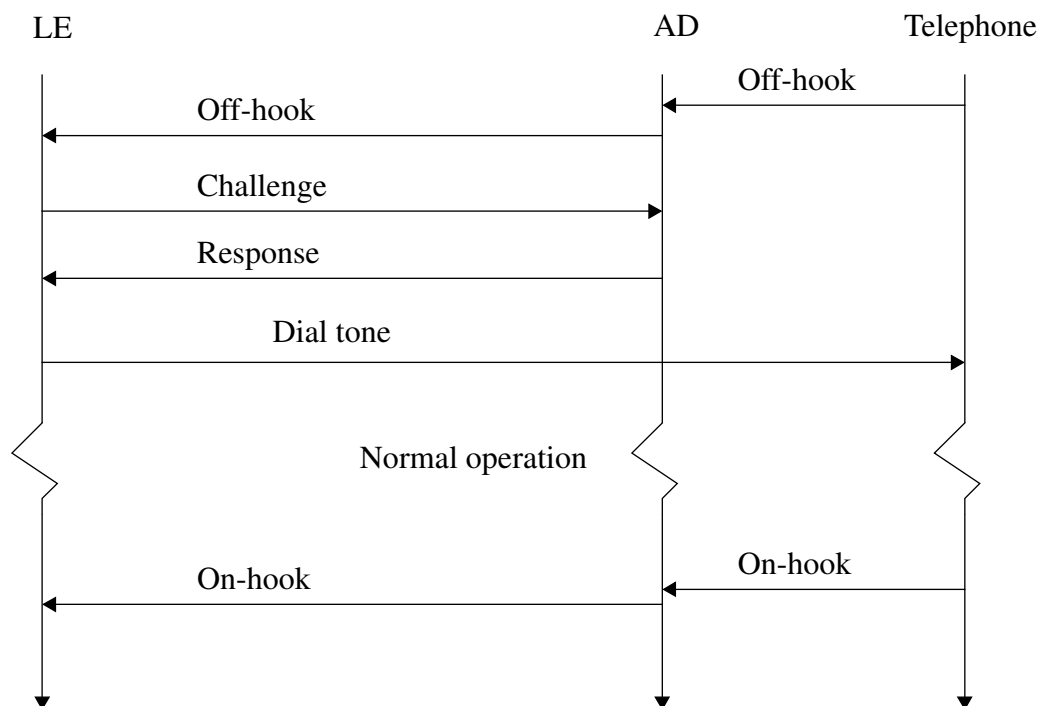


Figure 4. Authentication protocol

possibility, the authentication function can be integrated into new telephone sets. Both solutions can in fact be used in the same network.

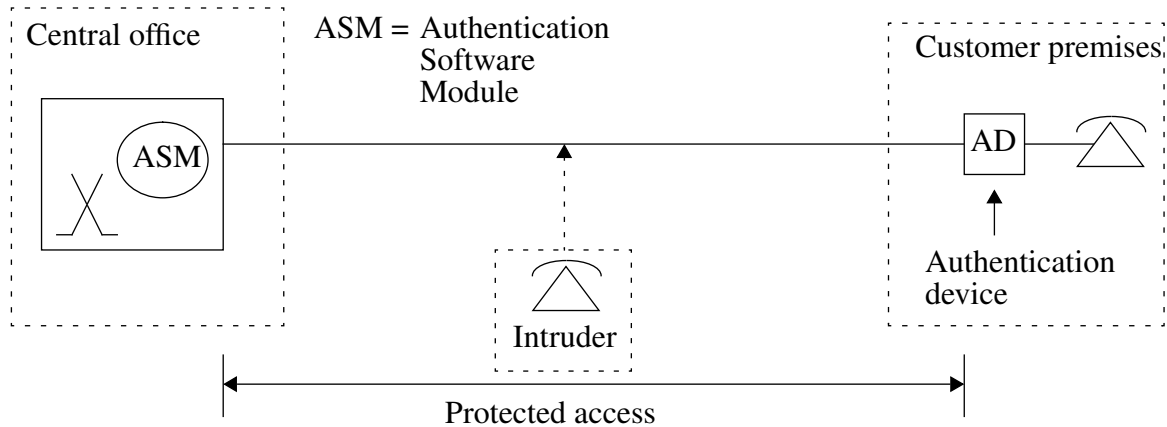


Figure 3. Protection of analogue access network

This solution restricts use of resources to authorized entities. An intruder connecting to the line anywhere between the customer premises and the central office will be refused service by the central office. From a security model point of view, each customer is authorized by the existence of a subscription.

It will also be impossible for a customer to falsely claim theft, because such claims no longer are credible.

In this scenario, customers will no longer be overcharged because of intruders and operators know that only legitimate customers receive service. As a secondary effect they will have satisfied customers and gain a good reputation in the market.

With this solution, we are thus able to eliminate the risk described in the previous section. It can be mentioned that an intruder still will be able to take over an established session, or telephone call, but this risk is considered minimal, because he can not control the routing of the call, so that it has minimal value to the intruder.

3. Choice of Mechanism

The DTMF (Dual Tone Multiple Frequency) inband signalling as defined in ITU-T Rec. Q.24 [6] is natural to use because every modern telephone exchange already contains the necessary functionality for sending and receiving DTMF tones. Other types of signalling for data transfer would require a modem, which certainly would be a too expensive solution.

With DTMF signalling in mind, we must try to define the best protocol. The International Standard ISO10181-2 [4] gives a good overview of the possible mechanisms. Class 0 which uses unprotected static passwords would be vulnerable to interception and replay attacks. Class 1 and 2 which provides protections against disclosure of authentication information are still

but we shall consider this as outside the scope of study.

It can also be imagined that genuine subscribers falsely claim theft, and thereby unrightfully may be given compensation or reduced telephone bills.

The most obvious vulnerability is the huge network of partly unprotected copper cables between the switching sites and the customers premises, combined with the analogue interface which provides no protection against masquerade.

The lack of proofs in the billing system is also a vulnerability, because the operator then has no irrefutable argument to refuse false claims from customers that their bills are too high. Detailed billing, i.e. where all B-numbers are specified on the telephone bill, provides some evidence, but it can only show that theft has taken place, but is not able to prevent it. It can also not provide evidence against customers giving false complaints.

The security breaches caused by the above mentioned threats and vulnerabilities can thus be resumed as; 1) Theft of telecommunication resources, and 2) False complaints.

2.2. Risk: Affected Assets and Adverse Business Impact

The asset most directly affected by theft is the revenue from subscribers. Cynically it can be said that when the operators were in a monopoly situation and were in a position to refuse all complaints of too high bills, theft was to their advantage because of increased revenue. In today's situation where customers require more justification, the operators may not be paid for the services they provide to intruders. Seen from the inter-operator point of view, illegal international traffic between operators can benefit some operators and cause extra charges for others, depending on their inter-billing agreements.

False claims can result in subscribers obtaining free communication service, and this can therefore also affect the operator revenue directly.

Another very important operator asset is a good image in the market. If that is destroyed by dissatisfied customers, because the operator is not able to solve the theft problem, it can cause a serious negative business impact.

The goal is to reduce the consequences and adverse business impacts described here, and the next section will discuss how this can be done.

2.3. Risk Resolution

Because we are unable to eliminate the threats or to make assets immune against security breaches, we will concentrate on reducing the system vulnerability. It would be an immense task to physically protect the copper cables, and the obvious solution is therefore to modify the interface so that it provides protection against masquerade. Authentication service protects against masquerade, and is therefore what we need.

A modified access network providing protected access is shown in Figure 1. Protected access is achieved using an authentication device at the physically secure customer premises together with appropriate software in the local exchange. The reason for using a separate authentication device is that it can be introduced without any modification at all to the existing installations, and compatibility with the existing network was one of the top level requirements. As a second

2. Risk Assessment and Countermeasures

Existing analogue access networks are comparatively easy targets for misuse of resources while we may assume that the central office and the customer premises are physically secure for our purposes. The unprotected access network is shown in Figure 1.

DF = Distribution Frame

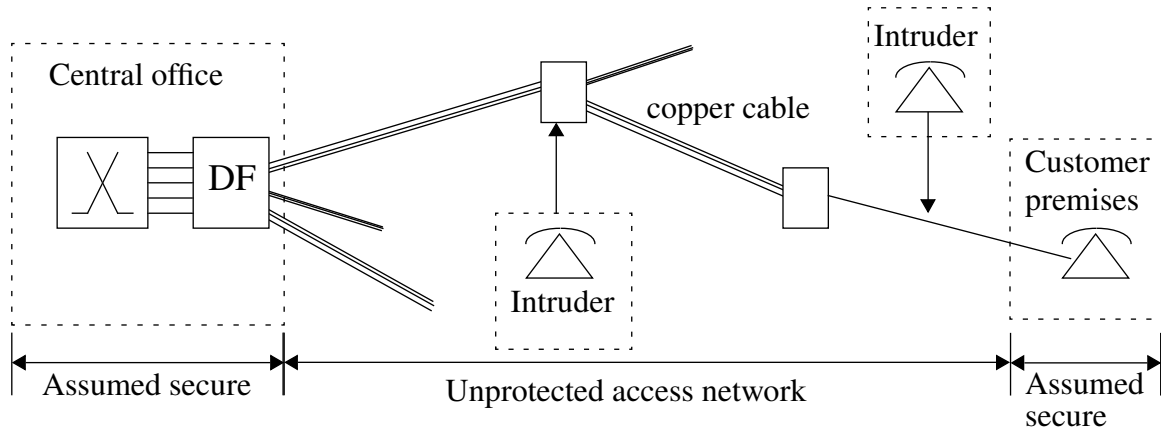


Figure 1. Access network fraud scenario

We will here try to analyse the problem from the operators side. The risk model illustrated in Figure 1 is used. Threats together with system vulnerabilities can lead to security breaches, but a security breach in itself does not cause any damage before assets are involved. The combination of security breach and assets then finally causes the damage, or what is here called risks. The goal is to reduce or eliminate the risks, and there are three possible ways to proceed: 1) Reduce or eliminate the threats. 2) Reduce or eliminate the system vulnerabilities. 3) Make the assets immune against the security breaches. A combination of the three methods will of course produce a greater effect than any of them alone. We will apply this model to our case, and see where efforts give the greatest effect.

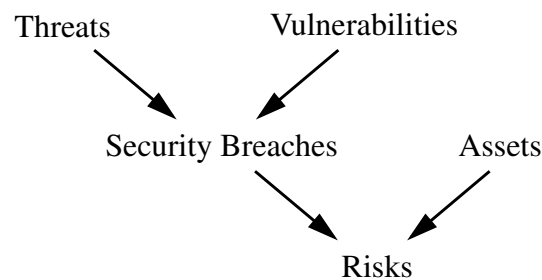


Figure 2. Risk Model

2.1. Threats, Vulnerabilities and Security Breaches

The threat causing theft in the access network is simply the group of dishonest people who see an advantage in stealing communication resources from others. It is theoretically possible to solve the problem by persuading everybody of the advantage to society of honest behaviour,

Authentication in Analogue Telephone Access Networks

Audun Jøsang¹ and Knut Johannessen²

Abstract

Theft of communication resources in public telephone access networks is a problem for many operators and their subscribers. In this paper we describe a practical solution to the problem of intruders connecting to other peoples analogue telephone lines and making calls on their charges. As main criteria, the solution must be simple to implement and cheap to introduce on a large scale, while giving strong protection. A challenge-response authentication protocol is found to be the most suitable mechanism, and the implementation is described in some detail

Key words

Authentication, Access Network, Cryptography, Telecommunication Crime

1. Introduction

The losses to customers and teleoperators due to theft of billable teleservices is not publicly known, but is estimated to 2-3% of the total world-wide revenue. This loss is partly due to theft occuring in the access network, and we will here present a solution to that problem.

The telephone operator's switching equipment and billing system is presently not able to distinguish between illegal and legal analogue calls, and the customer is initially charged for the theft. Only when the customers complain, does this become relevant to the operators. In the past, when most countries had state owned monopoly operators, it was very difficult for customers to prove the theft, and they had nowhere else to go if they were dissatisfied with the operator's service. This is rapidly changing because of liberalisation, and customers will in a competitive marked with multiple operators be in a better position to win disputes over theft, so that it is more and more becoming the operators problem. We therefore feel that operators now have an interest in solving the problem of theft in the access network.

The analogue telephone access network was not designed for providing security services, and the first top level requirement for any solution must therefore be that it is compatible with the huge amount of existing installations. Since there still is, and for a long time still will be, a large number of analogue subscribers, it can be characterised as a mass consumer market. This leads to the second top level requirement that the solution for each subscriber must be cheap to implement. As a third top level requirement we will add that the solution must give strong protection, and we will show that these three requirements can be combined.

1. Dept. of Computer Systems and Telematics, Norwegian University of Science and Technology,
N-7034 TRONDHEIM, NORWAY
Email: ajos@idt.unit.no

2. Telenor Network AS, P.O.Box 6701, St.Olavs Plass, N-0130 OSLO, NORWAY
Email: Knut.Johannessen@s.nett.telenor.no