# Conditional Access for Mobile Location-Aware Business

Stig F. Mjølsnes
*School of Science and Technology,*
*Stavanger University College, Norway*

## BIOGRAPHY

Stig Frode Mjølsnes received his MSc (Physical Electronics) in 1980 and PhD (Cryptography) in 1990, both from Norwegian University of Science and Technology. For the last fifteen years he has carried out research in the field of cryptographic protocols applied to digital payment and ecommerce systems, mainly through European Community research projects. Since the Galileo project emerged in 1999, he has taken particular interest in the notion of location-awareness by mobile terminals in heterogenous networks. He is associate professor at Stavanger University College, Norway.

## ABSTRACT

This paper presents findings from a theoretical feasibility study carried out for GalileoSAT, ESA, on requirements and techniques for commercial access control of satellite navigation service. Both enhanced GPS and the Galileo satellite systems are scheduled to support commercial mobile terminals, networks and services with time-and-place information. The design and implementation of secure and cost-effective access control mechanisms could turn out to be a key component for direct commercial exploitation. The question is then how this can be achieved.

Currently, four broad categories of signal service are identified for Galileo; one free of charge, one for commercial access, one for safety-of-life applications, and one for public security. This will require certain provisions with regard to frequencies, signal structures and access messaging of the downlink. Regarding the commercial service, taking on and sustaining a business model that assumes the satellite navigation signal to be access controlled individually for a very large dynamic number of users will be quite demanding on a self-contained satellite system, and bear resemblance with the access control problem and user environment of satellite digital video broadcast systems. Essentially, the capacity limitation in the navigation service broadcast channel together with the lack of a cost-efficient integrated "return channel" imply that currently specified GNSS cannot provide a self-contained platform for user access restrictions to business-operated location-aware service creation on a European or world-wide scale. Thus the most significant aspect to successful commercial access control design in the Galileo context were found to be the availability and capacity of a bi-directional data communication channel, to effectively enable:

- Feasible cryptographic key distribution and management of location-aware (roaming and visiting location) access rights.
- Complete integration of payment with value-added service access.
- Efficient piracy and attack control.

Mobile bi-directional packet switched networks will not only be able to support the pull model for access rights distribution and management employing efficient and scalable security protocols, but this architecture will place commercial access control functionality exactly where it is needed; at the networked server entry points of value-added Internet-based geo-information, such as shared client-server computations for fast, accurate positioning. The confluence of several technology developments substantiates the proposal: The freely available satellite ranging signals, augmentation systems and services, 4G mobile networks and server-assisted computing.

## 1. INTRODUCTION

### Background

This paper is based on results from a study on "Controlled Access" carried out for GalileoSAT during the autumn of 1999 [1]. That report is mainly occupied with the problem of access control mechanisms for the navigation signals from the MEO broadcast satellite constellation and the user segment, though the proposals therein have consequences for the design of the other segments too. An implementation of cryptographic access control mechanisms that enables a flexible business model serving an expected very large number of commercial and private users will require certain provisions with regard to frequencies, signal structures and most importantly cryptokey distribution and management. The primary

Proc. 14th International Technical Meeting of the Satellite Division of The Institute of Navigation ION GPS 2001, Salt Lake City , pages 2600-2611, September 2001.

objective of that exploratory study was to assess the requirements and identify alternative options for this problem under the prerequisite of access restrictions to the signals from the navigation satellites.

Two initial observations were made with regard to this access control problem. First, there is a immediate and clearcut analogy between the problems posed by this navigational information channel and commercial satellite based video broadcasting [2]. Similar to GPS, satellite pay-TV distribution is traditionally based on a one way broadcast mode of communication. However, the emerging convergence and integration of digital communication technology based on packet-switched bidirectional networks, notably driven by the success of Internet technology, certainly makes an impact toward interactive modes of operation for typical stand-alone broadcast systems as well. Already, DVB satellite set-top decoder boxes include telephone network modems.

The second initial observation that was made in the study is that the GPS system started early to move toward auxiliary network integration, both terrestrial and satellite, thereby supplementing the basic functionality with augmented accuracy and integrity [3]. GPS-functionality merges with mobile cellphones for advanced emergency handling procedures. The concept that GPS terminals can obtain improved accuracy by employing differential data by Internet-based servers is already well established [4].

In this paper, I further the arguments developed in my Galileo report [1] for integrating fully bidirectional packetswitched networks that will accommodate, not only for solving the original problem in the limited context of technical requirements for secure access control, but the hypothesis is that this will provide a strong platform for flexible and open service creation of location-awareness. Although at that time being unaware about ongoing work on "server-assisted" GPS experimentations, this development has confirmed my independent reasoning for this direction as viable. For the last year, the industry has been moving quite rapidly towards this modality. (Ref Lucent, Snaptrack, SiRF [14])

**GNSS commerce and access control**

Both the emerging European GALILEO and the U.S. enhanced GPS GNSS are scheduled to support commercially available location-based services for networks and terminals, in addition to the fundamental strategic and institutional nature of such systems. The stated European policy is that the financing of GALILEO system should be based on the participation of the private sector as long as a guarantee of service utilisation comes from the public sector [5]. The U.S. policy is to 'encourage acceptance and integration of GPS into peaceful civil, commercial and scientific applications worldwide; and to encourage private sector investment in and use of U.S. GPS technologies and services.' [6]

Application specific civil and commercial augmentation and auxiliary communication channels are already established, such as WAAN EGNOS and LAAN/ILS.

The fact that GPS signals already are present and freely available makes it difficult to envision that the GALILEO system can charge for the basic ranging signal without added value. Of course, institutional regulations may provide juridicial assurance and liability as such. Nevertheless, suggested features of technical added-value are improved accuracy, controlled availability, higher integrity and continuity of signal service than GPS.

Currently, the requirement is that GALILEO shall provide four levels of service:

1. Open Access Service. A service to the mass market. The satellite position messages are freely available, with better than 10 meters 95 \% accuracy. Universal access to a basic signal for mass-market applications.
2. Commercial Access Service. A liable, certified satellite navigation message service with commercial added value compared to level 1. Guaranteed availability and accuracy with liability cover in case of system failure or warning.
3. Safety of Life Service. Safety of life and security-related services. A certified satellite navigation message service with added value to aviation, transport and community-critical and safety-of-life operations such as police, fire-, search- and rescue brigades etc, in particular with respect to integrity.
4. Public Security Service. This service is similar to the military part (PPS) GPS service with strong communication resistance to intentional interference and jamming, to be used by Member State security agencies.

The current official thinking is that service level two and three must only be available in return for payment, hence some sort of cryptographic techniques must be applied to limit the signal access only to eligible parties. No technical specifications of "controlled access cotnrol" has been adopted yet for GALILEO. Several commercial models for service access payment can be envisioned, the subscription model with flat rate being probably the most immediate and relevant. The discussion of this paper is mostly related to level 2 Commercial Access Service.

## 2. SCENARIOS

We have already noted the analogy between the problems posed by this navigational information channel and other systems, such as commercial satellite based video broadcasting. Similar to GPS, satellite pay-TV distribution is traditionally based on a one way broadcast mode of communication.

The emerging convergence and integration of digital communication technology mainly based on packet-switched two-way computer networks, notably Internet technology, certainly will make an impact on the modes of operation for typical stand-alone broadcast systems as

well. The unidirectional mode assumed of the downlink satellite communication could be enhanced with bidirectional satellite communication from the mobile navigation terminal. It is even more reasonable to introduce bidirectional channels by mobile cellular networks when available, that is, combine the Galileo navigational information channel with mobile access to terrestrial communications, such as UMTS with internet services.

This will introduce quite a large degree of freedom with respect to value-added services that will be supported by public-key based interactive cryptographic protocols implementing the access control policy. This architecture will provide a much stronger and more flexible access control security than ever be possible within the constraints of a "one-way" channel to the navigation terminal. The technical arguments for this will be presented further in Chapter 4.

The commercial success of the GSM is very well established by now. We note a potentially interesting analogy between the subscription model and user roaming techniques of digital mobile telephone systems like GSM and UMTS, and a similar subscription model applicable to a Galileo system with multiple operators and regional access control. The GSM is in many aspects an open system, with respect to network operators, to value-added service providers, and to equipment suppliers. The market competition between networked pocket computers and mobile telephones is happening. With the GSM, normally the user will hold a subscription with one of his local GSM operators. In principle, the user is able to roam through the realm covered by any other GSM operator, supported by the mechanism of visiting location registers and secure authentication service. Unfortunately, the GSM system specification work did not enter sufficiently into the challenges of automated clearing mechanisms of operators, nor did the standardisation process foresee the potential of other payment mechanism than the subscription model with its postpayment of usage. The unexpected popularity of the prepaid “cash” subscriptions shows that much is left to be understood within this business area.

Limiting mobile business merely to a subscription model will effectively be a "low-pass filter" on end-user service developments.Within the vision of a dynamic marketplace of communication services, where value-added services will come and go, the mechanisms of “cash” payments will be indispensable to cost-effective network content commerce.

Common to both a pay-TV system and UMTS is the utilization of a personalized token (subscriber identity module), currently with the physical design of a smart card. This formfactor is inherited from the banking industry and directly adopted in the telecom industry, althoug alternatives to cards exist. The essential property of the personalized token is as a security component for copy-protection, implemented by tamper-resistant hardware and secure software. The small computing device (32 bits microcontroller with integrated memory is becoming the industry standard), its content and computing behaviour are associated with access rights, hereby obtaining *conditional access* to the system services.

Although this general interpretation is quite reasonable in itself, `conditional access` is for historical reasons often linked to the specific use of the smart card in the DVB industry, characterized by satellite broadcast communication to the reference monitor (the point of access request decision, see next section for a more formal definition) distributed and installed in the user terminals. The access is effectively granted and performed in the user terminal end by the security token of a smart card.

A personalized security token is also part of the GSM access control system and installed in the user terminal. However, this access system is distinctly different from the DVB access control, because the access request is resolved on the network/service operator side.

The two access control solutions can to some extent be regarded as inversions of each other. In the DVB system, first the message is broadcasted and then the access decision is made at the receiver terminal. Whereas in the GSM system, first the access decision is made at the network end, and then the communication session is established. Basically this happens because the return channel was not originally available in DVB systems. Both DVB, DAB and other standardized broadcast network technologies are now heftily introducing "return channels". For DVB, the threat and activity relating to commercial piracy of subscription smart cards is a driving force for this. The current version of GSM subscription cards has also been hacked to some extent, but note that the impact is much less because of bidirectional data channel available between the mobile terminal and the service provider.

**Security-relevant characteristics**

Three significant characteristics or aspects have been identified in [1] to have major impact on the design constraints relevant to security and controlled access.

The *first and foremos*t is the existence and capacity of a bidirectional data communication channel between the receiver terminal and the system. Most modern cryptographic protocols assume a bidirectional channel of some sort. As will be shown, the bidirectional channel is important for the security of

- key distribution efficiency,
- piracy control,
- flexibility in implementing geographical access control policy, and
- integration of payment and service access.

As a consequence, the availability and capacity of this bidirectional channel will be a major force in the rapid development of location-aware value-added services based on Internet-technology.

The *second* aspect concerns the concept of subscription roaming, similar to "visiting location" service available in GSM. A similar notion of "realm access policy" could be enforced by the smart card, or a combination of smart card and a interactive network service. The GPS differential system can be considered to be a first generation of such services.

The third aspect is the access control mechanism support of a variety of payment options and tariffs. Payment is "the other side of the coin" of a commercial access control system, and provisions for completely digital payment functionality are mandatory for future digital services. An integrated approach to digital payment systems will carry implications to the access control system, both at the user terminal level and the system management level. It will impact the level of smooth service integration and combination.

It becomes clear that the Galileo system could span access control mechanisms and management from a range of scenarios, the main categories being:

a) Off-line with physical predistribution of smartcards, updated over navigation satellites.

b) Auxiliary system, for instance "preauthorization" and update of smartcards via Internet service.

c) Navigation satellite integrated data communication for online access control.

d) Realtime online client-server mode of operation employing mobile networks for augmentation, for instance with commercial access control to server-assisted fast, accurate positioning computations.

## 3. ACCESS CONTROL

Recall the standard cryptographic threat models of passive and active attacks, that is, the signals can be tapped, generated, modified or deleted without physical restrictions. This leads to consider the basic information properties of required of the communication channel.

*Confidentiality:* only authorized users can acquire the information.

*Authenticity:* the recipients can verify the origin and integrity of the information.

Access control theory models the access problem as a triple *(S,O,R)* where *S* is the set of subjects, *O* is the set of objects, and *R* is the set of access rights or operations defined on the objects. This can be depicted as an access matrix coordinating subjects and objects, and with entries being subsets of *R*. Examples of fundamental access rights of information objects are *Read* and *Write*.

Consider the information security of the downlink channels carrying navigation messages that shall be access controlled within, say, three levels of service. The satellite channel is a broadcast channel, each satellite covering a fourth of the earth's surface at any time. Protecting access to information objects on an open communication channel available for both passive tapping and active attacks can only be done by cryptographic techniques. Restricting *read* access can be done by cryptographic coding (encryption), hence creating confidentiality. Restricting *write* access to the channel is done by authentication coding, hence creating verifiable integrity and origin of data.

The information objects are primarily the navigation data and ranging signal in our context here, but must also include the access rights themselves, in the representation of cryptographic keys. The cryptographic key management is the mechanism for implementing the access rules.

The concept of a *reference monitor* in access control theory is an abstraction that is postulated to control all references from subjects to objects according to the rules of the access control policy and the specific access rights granted. Basic security properties for a reference monitor are:

- Enforcing a complete separation between subjects and objects, so that it is always invoked,
- Complete and correct operation according to access control rules and policy.
- Tamperproof functionality.

As it is an abstraction, a reference monitor is not necessarily implemented by a single piece of hardware and software. In a distributed system, it rather represents the collection of access control devices.

The basic idea is to design an access policy model and implement that by protecting the information by encryption such that only authorized users can gain access to the navigation information and value-added messages, that is, being able to decode and interpret such messages. Authorizing read access is carried out be distributing decryption keys to eligible receivers.

The informational authenticity and integrity property, that is, security against sender impersonation ("spoofing") can be solved, for example, with public-key techniques of one-way hash function and digital signatures on the dataframes, where copies of the system's public key reside in the mobile receiver terminals, encapsulated by the smart card chips. The sender is authorized for write access by assignment of a secret key, the receivers are enabled to check the authenticity by using a corresponding public key, for instance distributed by regular broadcasting.

## 4. ACCESS CONTROL MECHANISMS

### Reducing accuracy

"Selective availability" is the GPS system term for controlling access to full system accuracy of the "standard positioning service". The mechanisms used for reducing accuracy were by introducing intentional errors to the satellite's clock and the navigation data parameters. Ingenious methods of circumventing this have been developed over the years and should make one cautious

about the design challenges in general for securing access to a broadcast ranging signal, particular having in mind side-information that may be available over mobile networks. GPS selective availability was switched off last year.

Of course, switching off the signal broadcast from the satellites achieves immediate access denial. Jamming can be used for regional and local cancelling of the navigation service. However, these mechanisms are hardly sensible in a commercial setting where one wants to maximize revenue!

**Spreading cipher**

A signal design based on spreading code phase measurements requires a distinction between the channel providing the ranging (*R*-channel) and the acquisition channel (*A*-channel), because in order to synchronize efficiently to the current state/position in a *cryptographic* spreading sequence of the ranging channel, some information about system time, shift register states etc. are needed up front. (This must be done even for a long noncryptographic spreading codes.) The necessary "access" data must either be received by some acquisition channel or internally computed.

It follows that this information must be established *prior to* using the ranging channel, hence the ranging channel itself cannot be used to transmit the necessary synchronization data. Thus this mechanism can be employed to achieve access control. In principle, there exists several possible sources of approximate synchronization information that needs to be input to the demodulator.

The receiver, when already having access to the cryptographic key and time-count, can carry out a search for the correct state by means of correlation techniques between the approximately known position in the sequence, and the incoming signal. If the time-count is not kept, it can be manually input based on synchronization information received somehow outside the system. Or, more efficiently, it can be acquired from another communication channel made available.

This synchronization data acquisition channel could employ spread spectrum modulation, so that the very same channel could provide some means of ranging measurements too. Apparantly we have now entered into some kind of circular thinking. Why is it easier to synchronize the receiving end of this channel? Of course, the trick normally done with the spread spectrum channels is that the spreading code applied is short and publicly available, so that it can be predistributed and embedded into the receivers. Hence the correlation technique can be applied successfully in short time without extra input.

So despreading of the ranging channel *R* is dependent on the synchronization data to be received by the data channel *A*. If the synchronization data provided over the *A*-channel is encrypted, the access to the crypto-key will control the access to high-precision pseudoranging and coded jamming. This is (probably) similar to the mechanism of PPS of GPS.

Assume some crypto key hierarchy (described in the next section) is applied. The access right key $k_r$, which indirectly enables the pseudorange measurements, is distributed on the *A*-channel. This key gives access to the necessary and sufficient synchronization and ephemeris data, also to be distributed on the *A*-channel.

The "handwaving" claim here is that a spreading code based on a strong cryptographic sequence (non-cycling spreading code) will be orthogonal with high probability to a short cycling Gold-code [12]. Conversely, we can achieve channel code orthogonality (with high probability) by employing a spreading cipher for the ranging channel. Any strong cryptographic pseudorandom bitsequence can be used as a spreading code because it will, among other characteristics, satisfy the correlation properties necessary for a good spreading code. Being a cryptographic bitsequence, it is indistinquishable from a uniformly distributed random sequence.

**Signal authentication**

The mechanism of encrypting the spreading code, as introduced into GPS, is directed at the threat of active disruption, for instance replay of a signal already transmitted. The claim is that the signal achieves "anti-spoofing" properties, the signal reception becomes more robust against deceptive signal generation, delay or modification attack, such as intentional signal jamming.. More precisely, one achieves *signal authentication* of the ranging channel by employing a symmetric cipher.

Recall that the GPS data rate is 50 bps, the encryption rate is 500 000 bps, and the spreading code rate is 10.23 Mbps. If the main security concern is for the authenticiy of the source data, then it is very inefficient to encrypt the channel code in this respect. Obviously, this is not a sufficient requirement. Probably, the reason for using distinct rates for the cryptographic sequence and the spreading code sequence was one of economy of design and implementation, trading off security.

Hence the spreading cipher mechanism is not employed foremost for restricting “read” but “write” access to the navigation message. The read access control to precision ranging measurements is carried out by encrypting the data necessary for the receiver’s PVT computations, for instance some (or part) of the ephemeris parameters. However, it is inherent impossible to distinguish between “read” and “write” access rights by using a symmetric cipher, thus the one comes with the other in the spreading cipher mechanism.

Authentication by public-key techniques does not require cryptographic secrecy (in the classical sense) at the receiver end. In other words, the authentication service can be made available to all, or on some differential basis, with little extra cost by employing this mechanism. This is relevant to for example the aviation sector demands, because it does not require access to a secret key, only a

public key. Unfortunately, public key techniques cannot possibly be used to generate spreading cipher because the despreading activity of the receiver requires essentially *the same* capabilities as the sender, contradicting the asymmetric properties of public-key systems.

The immediate security implication of having the same cryptographic capabilities for both the sender and the receiver is that any user that holds a legitimate receiver, in principle, is able to spoof the navigation signal because he can, for instance, delay and retransmit the signal with a better S/N so that the receivers will lock onto this instead of the original navigation signal. It is not feasible to screen and classify users in a commercial setting, so the service provider cannot assume trusted users. Thus by this thinking, any normal user is now potentially enabled to carry out signal spoofing. The threat of signal impersonation and interference is reportedly perceived to be realistic.

**Trusted computer token**

Smart cards are readily available technology designed to satisfy the demands of high-volume mass market applications. The most important functionality of a smart card, which basically is a microcontroller glued to a plastic card, is to function as (part of) the access reference monitor in the system. This small computer must be able to protect access parameters, such as cryptokeys, ideally rendering it impossible for nonauthorized to read, write or modify the protected data. Further, the trusted computer must be able to perform the computations where secret keys and data are part of the input, in such a way that it is impossible for non-authorized to inspect, influence or modify the protected computation. This put forth the requirement of *tamper-resistant hardware*.

By making this distinction between the general terminal/receiver as such, and a special computing module, it becomes practical to personalize the tamper-resistant computing module independent of the receiver. Thus separating the customer management from the manufacturing process of the general receiver.

Currently, the smart card is the most popular formfactor of a trusted computer token, and is expected to be applied in quantities of billions in the coming years. The GSM and UMTS developments of subscriber modules have made the smart card a vital part of the mobile telecommunication business. Other formfactors are available though, which provide a better and more robust encapsulation. For instance, the unavailability of a constant power supply makes the card microcontroller vulnerable to physical attacks of reverse engineering. For the implementation of active tamper-resistance, various electronic sensor and alarm technologies can be added to the encapsulation, thus providing for a total erasure functionality conditioned on tampering.

Several types of attacks are possible. Cryptoanalytical attacks that break the cipher algorithm will imply that all smart card must be exchanged for new ones with stronger cipher algorithms. Other attacks will try to read out the secret keys. There are physical reverse engineering attacks on the VLSI circuitry. And there are potential electrical attacks, like voltage, current and clock tampering, trying to achieve some unexpected circuit behaviour. And there are logical attacks breaking some assumptions about the communication protocols. Normally, smart card hacking is not easy and cheap. In fact, well equipped labs are used in an organized way in the pay-TV business.

The race between the makers and the breakers results in the need for continuously exchanging old technology smart cards with better tamper-resistant technology at a reasonable pace. This fact must be carried in the system management plans, which should find a balance between level of tamper-resistance, cost, and functionality.

It is very important to note the implications of signal authentication, as discussed in the previous section, with respect to the smart card. The decryption facility must be part of the despreading operations, with high speed requirements. For GPS, the *P*-channel demodulator must operate at the chipping rate of 10.23 MHz. Current standard smart card technology cannot support such high-speed stream computations. After despreading, the processing rate per satellite seldom will exceed 1 kHz, typically in the order of 200 Hz or lower. The smart card ISO standard today is half-duplex input/output with datarates restricted to about 5 -- 50 kbps, and likely this will not change in the near future. The consequence is that the current smart card technology is not sufficient to implement signal authentication, but is adequate for message decryption and key management.

**Multi-application tokens**

Smart cards today are security parts of closed systems where the issuer is equal to or closely associated with the service provider. The user ends up with a card for each service. The reason for this is rooted in the security and technical management architecture of the systems.

The problems with usability, management cost, and service flexibility are clearly acknowledged, and the industry is trying hard to move towards solutions based on multiapplication/service cards.

The forces against this are at least threefold: All service providers want to be in control of the card and their branding, the sharing of resources and smart card creates new security problems not solved yet, and the general task of revenue clearing among the service providers requires established infrastructure cooperations.

This problem will not show up for a "closed" unidirectional GNSS access control system, where a centralized hierarchical subscriber organization can be set up. The problem will emerge as soon as the commercial navigation service is to be integrated with another service, say cellular phone service. How many card slots will be available? Or how can the commercial service access be put on and securely managed on for instance subscriber cards for UMTS. Or the other way around, how can the GNSS card accomodate for more security applications?

Even so, new concepts of signed applets downloadable by the owner of the smart card are being purported by Microsoft, Sun and others. This will create an opportunity for a totally new way of organizing "the booting" of security applications involving smart cards. For instance, the user will be able to download a subscription "cardlet" and mount this on his or her card. The underlying assumption is that the navigation service provider is able to rely on trusted public-key infrastructure already in operation.

## 5. UNIDIRECTIONAL BROADCAST ONLY SYSTEM

This section will look into what access control on the navigation satellite broadcast channel will require in terms of cryptographic key distribution to the receivers.

### Key management complexity

A first sketch of a key distribution solution entails cryptographic parameters and keys associated with user subscription to be stored securely in smart cards, and to be attached to the the mobile receiver terminals. The service access right, the smart card, and the subscribing person are linked by some kind of identity verification, such as password or biometrics. The subscription roster will be maintained in a key management center in the master control.

The key distribution and management system must:

- Serve in a global commercial environment, accordingly, only public-key based distribution techniques can offer sufficient efficiency, flexibility and scalability.
- Provide sufficient flexibility to accommodate a variety of access policies with respect to content, levels and geographical areas.
- Use protocols with personalized receivers, including tamper-resistant storage and computing

The scaling complexity of users is the most important issue in cryptographic key distribution and management. Satellite DVB systems constitute a pertinent case study for analysis in this respect. We shall argue that key distribution by broadcast satellites (unidirectional on-satellite key-distribution) neither scale well with increasing subscribers/users, nor with respect to dynamic service provision. Hence, a distribution mode similar to the "traditional" DVB push solution is not suitable, especially if the business model needs access rights granted on a customer-by-customer basis.for secret keys and computations.

### Broadcast key management for conditional access

The main focus here is the design problem of access control of positioning service utilizing only the satellite broadcast down-link for key distribution and management.

Access control could be based on several conditions, such as geographic area of subscription (Lat,long), time of subscription (Expires end of ..), service level, etc. The dynamic management of a subscriber database will include operations and procedures for **new** users, **update** users, **delete** users, **modify** access rights, and so on.

The very large scale of the system, both in the number and variety of users with its global coverage advices employment of a distributed database system. The initial access right of the user (for instance a flate rate subscription) must be set up and pre-distributed by some auxiliary system that can exchange either with authenticated identity or digital payment or both.

The trust relation between issuer of the access right and the holder of the access right is asymmetric; the issuer does not trust the holder not to bring the keys to the knowledge of a third parties. Hence the need for the service provider to issue a "tamper-resistant" storage and computing device, for instance a smartcard. However, commercially available tamper-resistant hardware can be broken if sufficient reverse engineering resources are put on the job, so we need "break-tolerance" built into the system, thereby avoiding totally breakdown of correct access control if a break should occur. A trade-off exists between cost and tamper-resistance.

**Example:** The message $m$ is encrypted under key $k$ giving $E_k(m)$. Every person that possesses access right to $m$ is given the key $k$. If $k$ or some bits of $k$ are disseminated/leaked/published, then the correctness of the access control mechanism in principle breaks down immediately.

Break tolerance can be alleviated by a more elaborate key distribution system. Let $k_i$ be assigned to authorized person $P_i$. The message is encrypted under key $k_i$ giving $E_{ki}(m_i)$ . Then only $P_i$ has access to $m_i$. However, if for all $i$ and $j$, $m_i = m_j = m$, that is, the message sent to all is the same, then for any $i$, a leak of $k_i$ will result in total breakdown of the correctness of the access control mechanism. In principle, with no return channel from the mobile terminal to the authorization system, no evidence can be signalled that an access key must be revoked.

In practice, the "pirate" has a key distribution problem of his own if he wants to redistribute the forged access rights. He could simply publish the key, but this will reveal *which* key is broken, so the compromized key can be revoked by the key management system. This will not generate any direct revenue stream, but can disrupt normal operation with a costly revoking game. If $k_i$ is revoked, then $k_{i+1}$ can be found and published. Another approach is to reduce the cryptoperiod, that is, a frequent renewal of valid keys. This puts more effort in the redistribution game of the forgery, in particular if it is a "commercial redistribution". Incidentally, this is basically the game played in the DVB industry.

**Protection hierarchy**

The classical solution to the above stated problem is a cryptographic protection hierarchy, say of three key levels. Let us call the first level the *subscriber keys*. According to normal use of public-key cryptography, a subscriber *s* will associate with a public key $K_s$ and a corresponding secret key $k_s$. These parameters represent the long-term key information associated with each subscriber or user. The subscriber keys are stored securely in tamper-resistant hardware and pre-distributed to the user, the public key is shared with the key distribution system.

The next level will be the *accessgroup keys*. Accessgroup keys will be assigned and distributed to eligible subscribers, protected by individual public subscriber keys. An access group *i* will be assigned an access group key $k_g(i)$, where a subscriber *s* to access group *i* will receive the key distribution message $E_{K(s)}(k_g(i))$.

*Access right keys* will be protected by access group keys. An access right *j* is assigned to an access group *i* by encrypting the access right key $k_r(j)$ under the access group key $k_g(i)$. The resulting message $E_{kg(i)}(k_r(j))$ is broadcasted and received by all users. However, only the members of the access group determined by the access group key are able to read the protected access right key. The fourth protection level is the satellite navigation data itself. This navigation data *m* is protected by one or more access right keys, broadcasted by the satellite as $E_{kr}(m)$. Hence this data can be accessed if and only if an access right key is available. The assumption is that the range measurements and position, velocity and time computations cannot be carried out without the correct reception of the actual navigation data. A further assumption is that these data can only be received by the satellite broadcast channel where they are protected by the access right key.

The general concerns of this distribution scheme are:

- The authenticity of the key messages
- Synchronization and error propagation
- Cryptoperiods and key revocation techniques, in particular in relation to the message key.

**Revocation**

A basic assumption for the application of the proposed key hierarchy is *active authorization*. The service access of the user expires unless renewal messages of some sort are received. The proposal is to achieve this by enabling communication. Unless a correct and authentic message containing a valid access right is received by the smart card within some deadline, the service becomes unavailable, simply by a change of coding. A simple mechanism implementing this would be according to some pseudorandom sequence generated by a keyed one-way function, but the best solution is to use the key distribution of the access right and user group keys. When the satellite changes encryption key, a “non-enabled” smart card simply cannot do its message decryption task because the new key instance is not replaced by the previous one. Further detailed consideration on which keys should be specific to each satellite and which keys could be systemwide keys are left open here.

*Active deauthorization* is an alternative approach, where the key distribution center is to (attempt to) turn off the signal service at the recipient end. Modification and revocation of access rights are signalled to the receiver and interpreted by the smart card as a "turn yourself off" message. Actually, a combined enable and disable scenario might be envisioned too. Anyway, one fundamental security assumption of this approach is that the “off message” is processed by the smart card, and that the smart card cannot be controlled by the user/subscriber because of tamper-resistance.

The approach of active authorization requires that the authorization must be renewed by a synchronized update of keys. This approach of deactivation requires that the authorization revocation message must be synchronized with the user being “on-line”. None of the approaches are feasible with the fundamentally "asynchronous listening" receiver. Both alternatives therefore require continuous “carroussell” broadcasting of messages until satisfying high likelihood for receivers being online is reached. The “activate” solution requires a list of activation messages, the “deactivate” solution requires a list containing deactivation messages. Several schemes of carrying out the “turn-off” service at the recipient end can be sketched:

- A “synchronized” revocation will destroy the very intention, because now the user will know when to shut down the receiver listening.
- Continous broadcasting of the deactivation message over time is already discussed.
- Delayed until the next renewal period of access right grant, but this is very close to a variation of the “activation” approach.
- Providing a local expiration time in the smart card, relying on the feed of correct UTC time into the smart card. But on closer inspection, it can be realized that this is a variation of the "activation" approach too.

Nevertheless, both approaches need some scheme for updating access keys, at least in the long run. It could be argued that this can be done by distributing replacement cards, relying on physical means of distribution outside the satellite navigation system. Or it can be argued that this should be done by a the presumably cheaper and faster method of digital distribution. It can also be argued that break-tolerance must be part of the design criteria. Therefore, the active authorization mode becomes the recommended design approach for the long term view.

**Key distribution**

Given the general scheme of a protection hierarchy of cryptographic keys, as outlined in the previous section, the next challenge is to investigate how the cryptographic keys can be distributed in the system, projecting the message complexity that this key hierarchy implies for GNSS. This

cannot be definitely answered before the channel and signal architecture is determined. On the other hand, the key distribution is of fundamental importance to the flexibility of the access control, and therefore to the viability of the commercial exploitation of the system. Hence the key distribution complexity must be a determining factor in the design of the system, if the mechanism of access control to the navigation signal is accepted.

Each subscriber will possess a subscriber key pair, the secret key "imprinted" in the smart card. Of course, the number of keys will increase linearly with the number of subscribers, say $n$. The distribution is by personalizing the smart card and secure physical transfer to the subscriber. Efficient procedures for this issuing already exists commercially today.

The initial access group keys $k_g$ can be distributed with the issuing and transfer of a smart card, but distribution by communication should be available in order to create a economic and flexible distribution system. The number of messages needed for a full update of one access group key is linear in the number of group members, and so is in the order of $n$. How frequently these messages have to be distributed and renewed depends on the validity period duration set by the subscription service. There are various methods of key generation schemes that will handle smooth overlapping of validity periods.

**Example:** Say $10^5$ users must have their parameter update within 10 minutes, where the user parameters total a length of $10^4$ bits. (Currently, this can likely be reduced to a fourth or fifth, but one should be conservative and provide for more taking into account the expected life-cycle of Galileo to be 15-20 years.) This results in a requirement for 1 Mbps data rate distribution channel. A preliminary estimate [13] of the user population tells $10^7$ subscribers of Commercial Access Service, and $10^5$ subscribers of Safety of Life Service. This little example indicates that it is in particular crucial to the system dimensioning in the unidirectional key distribution system.

The access right key distribution aggregates only a few messages, proportional to the product of the number of access groups and the number of access right types used in the system. A first estimate would be less than 100. If these messages could be satellite broadcasted more or less continuously on a revolving basis then the start asynchronicity of the users/receivers will probably result in an acceptable average waiting time to service access. Nevertheless, the access right keys should be made available by an auxiliary network as well in order to accomodate time-critical applications. The validity period of the access right keys could be a multiple of the satellite upload rate, determined by the fastest required renewal frequency of the access rights.

The general problem of "broadcast encryption" has been dealt with in the theoretical litterature, for instance where the model allows a dynamically changing partitioning of the user set of $n$ members [7]. Of course, this results in an exponentially bounded number of subsets in $n$, but cryptographic schemes can be constructed that create resilience against any coalition of $k$ out of $n$ users, where every privileged user has to store $O(k \log k \log n)$ keys and the key distribution center has to broadcast $O(k^2 \log^2 k \log n)$ messages. Unfortunately, this is for every new billing/key period. There are trade-offs in key distribution between the transmission length and the storage size at the user, but theoretical studies have shown prohibitive lower bounds [8], either the transmission will be very long or a large number of keys that need to be stored in the receiver. The practical concern of terminals being off-line and not being able to acknowledge if the keys have been successfully received has not been included in the models and adds to the practical concern already discussed in the foregoing.

The user requirements that put constraints on the solution for key distribution mechanisms for access control need to be identified, such as acceptable "log-on" time, resynch-time with key renewal, and other system parameter updates. The system planned lifecycle of 15 -- 20 years requires careful attention to the scalability with respect to users and service providers. The dimensioning of channel capacity for distribution of access group keys is of immediate concern.

This results in the proposition that capacity limitation in the navigation service channel together with lack of a cost-efficient "return channel" imply that currently specified GNSS cannot provide a self-contained platform for user access restrictions to business-operated location-aware service creation on a European or world-wide scale.

**Side-information effects on access control**

It is important that the security analysis of the acess control mechanisms and system designed make the assumption that one or more "independent" monitoring networks will be available to users and could be used in attacks on the security of the access mechanisms. It is very likely that, for instance, dedicated amateurs enter to organize a worldwide monitoring network coordinated and provided by Internet services, if such a service will provide added value to such a group and beyond.

## 6. PUSH OR PULL

One fundamental problem formulation with respect to key and parameter distribution is to distinguish which model to go for of the following two:

- **Pull model**, where users pull new access rights and keys (asynchronously) from the system database when required. This is optimal with respect to communication and scales well on a distributed system basis, but requires bidirectional communication.
- **Push model**, where keys, certificates and other parameters are broadcasted on a revolving basis ("carousel mode"), to all users whether they need it or not. This is best suitable for a closed system with a priori knowledge of the maximum number

of users, but can be realized within a unidirectional broadcast communication.

The recommended design for a global access restricted GNSS will be in accordance with a pull model. One solution is to let this key distribution be determined by "user pull" via auxiliary bidirectional mobile networks (Internet technology). This will accomodate the asynchronicity of the users log-on. Personal computers and mobile communicators already come standard equipped with smart card readers. GSM terminals are on the market now with dual slots for smart cards, enabled for WAP based applications.

## 7. BIDIRECTIONAL AUGMENTATION

Consequently, the proposal is to adopt the client-server computational model with access control to mobile value-added geo-information. This model of twoparty client-server computations communicated via mobile bi-directional packet switched networks will not only be able to support secure and flexible access protocols, but will place access control functionality exactly where it is needed: at the entry points of the value-added services.

### The future mobile environment

Contemporary terrestrial digital cellular networks labelled second-generation (2G) networks were designed mainly for voice communications employing traditional line switching technology. The data rate of 2G networks is up to 14.4 kbps. Ongoing upgrade activity to 2.5G networks may increase the data rate up to ten times that speed, but more importantly 2.5G add-on packet radio switching (GPRS) providing "always-on" service. The 3G network solutions, in particular UMTS, specify radio access data rates from 384 kbps up reaching 2 Mbps for low-velocity and a few hundred meters distance to base station. The recent financial down-turn of mobile business has spurred quite some uncertainty with respect to the schedule of expected worldwide migration to 3G, said to be through year 2005.

Nevertheless, 4G wireless network research is ongoing (reference in Computer etc) taking on the technical challenges of global roaming support across heterogeneous wireless and mobile networks [9]. Observe that the deployment schedule for 4G coincides roughly with the Galileo project (ca. 2008). Therefore, it is reasonable to anticipate 4G networks to be *the coexisting mobile environment* for future GNSS (Enhanced GPS and Galileo). The 4G mobile network is envisioned to provide IP packet-switching interoperability for seamless mobile Internet access offering available bit rates of 50 Mbps. The 4G network architecture will have to resolve issues of multimodal access by integrating cellular, wireless LAN, satellite and fixed wireless data transmission technologies [8]. Note also that the optimal topographical configuration of a cellular communication system constitutes a very poor platform even for 2 D navigation/positioning according to recent research [11]. Following this line of thought, the interesting question within this paper's context becomes: How can future GNSS services take advantage of 4G mobile networks?

Having presented good arguments for employing a bidirectional data communication networks for security and service-augmentation still leaves the concrete construction up in the air. How can this be realized? The problem becomes a special case of realizing access control in mobile distributed system, and as such not at all solved in its generality for the time being. However, the GSM system represent a practical and successful first design iteration for how this can be done. Moreover, the UMTS system is a followup on this design, improving on many of the shortcomings of GSM security, such as mutual authenticity in access and provision for end-to-end security.

Location-based services are probably best considered as part of the "content-provision", rather than part of the access network service. On the other hand, it could turn out that the UMTS security architecture can be employed directly and efficiently for access control. Note that several industrial consortia are already busy working on similar problems with respect to mobile commerce [10], so having pointed in this direction we leave this challenging question here for further research for 4G systems.

### The thin client and the fat server

Important considerations for the GNSS receiver equipment that could take advantage of a thin client include:

- the cost of the receiver equipment,
- time to first position computation after start,
- computation time of the receiver,
- storage requirements,
- time validity and transmission timeof the positioning parameters,
- ephemeris and position computation accuracy.
- power consumption and battery duration in wearable equipment.

Lowering mobile user terminal cost and power consumption can be achieved by delegating much of the computational effort from the mobile user terminals to networked servers. Already projects are underway to develop single-chip GNSS receivers to be embedded in wearable terminals such as mobile phones and personal digital assistants. Multimodality of antenna, RF front end and digital processor are research issues, whereas integration with general microcomputer, input and output components of the handheld device is a matter of software integration.

At least four different modes of augmented communication can be identified:

- Receive broadcasted or multicasted augmentation data.
- Request augmentation data on demand.

- Request computation on demand.
- Request a continous service on demand.

## CONCLUSIONS

During my design study [1], I became convinced that commercial access control made best sense at the point of online *value-added service*, and not at the user terminal. This paper has given an analysis of the reasons for this with respect to access control mechanism implementations. Unidirectional broadcast channels within extant and projected capacity restrictions do not provide sufficient support and flexibility for access rights management with cryptographic access control in the terminal. Instead, the pull model of key management is recommended for full flexibility of world-wide scale provision of location-based services, but this requires a bidirectional data channel. The technical obstacles of designing a low-cost terminal-integrated bidirectional (satellite) data channel of sufficient capacity for cryptographic key management is acknowledged. The proposal becomes to employ UMTS and emerging 4G mobile networks for this purpose.

Now four comcomitant observations create the following reasoning. First, GPS ranging signals are already and very likely will continue to be freely available and enhanced, so commerce must be found in value-added services. Second, the validation of this is already established in the successful GPS augmentation systems and services, either already operating or scheduled for deployment. Third, 4G mobile networks planned to coexist with GNSS-2 will provide Internet access, implying broadband access to an open system of dynamic service creation and provisioning. Four, wireless access by handheld and wearable devices will take advantage of the notion of client and serverbased computations to obtain small, inexpensive devices with low powerconsumption. This applies in particular to GNSS receivers, where for instance the basic position computations easily lend itself to be carried by a protocol between client and server.

All this leads up to the recommendation that mobile location-aware GNSS-2 business should not be based on conditional access control in the user terminals, but be based on end-to-end access control mechanisms between the user terminal and the mobile Internet point of value-added service.

## ACKNOWLEDGEMENT

Thanks to Børje Forssell and Knut Grythe for supplying knowledge and discussions about GPS and UMTS.

## REFERENCES

[1] S.F.Mjølsnes. *A study of implementation aspects of GALILEO controlled access service*. SINTEF Technical Report STF40 F00045, Trondheim, 15 May 2000.

[2] S. F. Mjølsnes. *Security of Access Service for GNSS.* Proc. 4th European Workshop on Mobile and Personal Satellite Communications (EMPS 2000), Eds. Gardiner and Glover, Wiley 2000, pp. 185-190. ISBN 1857900774.

[3] Special Issue on GPS. *Proceedings of IEEE*, January 1999.

[4] Internet Based Global Differential GPS experimentation. http://gipsy.jpl.nasa.gov/igdg

[5] European Commission of the Communication on Galileo, on February 10th, 1999

[6] March 1996 U.S. Presidential Decision Directive

[7] A. Fiat and M. Naor, "Broadcast encryption", in Advances in Cryptology –CRYPTO'93, LNCS 773, Springer-Verlag, 1994, pp.480-491.

[8] M. Luby and J. Staddon, "Combinatorial bounds for broadcast encryption," in Advances in Cryptology – EUROCRYPT'98, LNCS 1402 Spirnger-Verlag, 1998, pp.512-526.

[9] U. Varshney and R. Jain: *Issues in Emerging 4G Wireless Networks*. IEEE Computer, June 2001 page 94 - 96.

[10] For example: Mobile electronic Transaction - MeT http://www.mobiletransaction.org/

[11] G. Hein, B. Eissfeller, V. Oehler, J.O. Winkel: *Synergies Between Satellite Navigation and Location Services of Terrestrial Mobile Communication.* Proceedings of ION GPS 2000, Sept. 2000, pages 535 – 544.

[12] B. Forssell. Personal communications and cooperation, Jan 1999- Jan 2000.

[13] R. Lucas. Personal communication. Intermediate meeting at ESTEC. 9 November 1999.

[14] An industryreport can be found on http://www.mobic.com/news/2000/01/qualcomm_acquires_wireless_locat.htm