METHOD FOR SETTING UP A SECURE CONNECTION USING PUBLIC AND PRIVATE KEY GENERATED IN USER TERMINAL

he invention concerns the technology of setting up secure communication connections between a mobile client and a server. Especially the invention is applicable for setting up a secure communication connection for the utilisation of financial services offered by the server. With certain prerequisites the invention is likewise applicable in setting up a secure communication connection between a server and a fixed user's terminal, like a home computer or office workstation.

Digital cellular networks typically have certain built-in security features that are meant to enhance security at the step of setting up a communication connection as well as during the use of such a connection. Most mobile stations require the user to enter a PIN (Personal Identification Number) code at the switch-on stage, which should prevent unauthorised persons to use a certain mobile station. Digital cryptographic methods are employed to encrypt information that is transmitted over the radio interface, so that it would not be possible to listen to somebody else's call in the purpose of stealing information.

For high-security applications like accessing financial services with a wireless mobile station, however, the basic existing security measures are not enough. Most people keep their mobile stations switched on all the time, which means that if a mobile station is stolen, the thief can use it freely – thus impersonating the legitimate owner to the direction of all services used through the cellular radio network – until the eventual moment when the legitimate owner manages to report the theft to a telecommunications operator who closes down the user account and puts the IMEI (International Mobile Equipment Identifier) of the stolen device on a black list. Another reason why e.g. banks can not trust the basic security of cellular radio systems is the fact that they would have to rely on the willingness and capability of the telecommunications operator to implement and maintain extensive security. It is unacceptable for a bank or similar high-security service provider to leave any key elements of communications security outside its own direct control.

With fixed user terminals the situation is even worse, because they are left alone for lengthy periods of time for anyone to tamper with. Additionally the security of the Internet in its basic form from the viewpoint of confidential communication is close to nonexistent.

A widely used approach to providing end-to-end security between a user and a service provider is based on the use of disposable keys. The user has a user identifier, which is more or less permanent, and a list of keys, each of which can only be used once for setting up a communication connection. After having used a key once the user strikes it out and uses the next key from the list next time. The user must give the key (for example by keying it in with a keypad) at the very moment of setting up the connection, which means that a stolen mobile station alone is not enough for fully impersonating the legitimate user. The disposable nature of the keys ensures that even if someone managed to eavesdrop on the setup phase of a communication connection, the stolen information is of no use because the unauthorised party can neither reuse the same key nor guess the next key on the list.

The drawback of the above-described approach is the clumsiness and vulnerability of the list of disposable keys. Many users store (regardless of explicit contrary instructions) the key list together with their mobile station or at least in a place from which it is easily found, like in a wallet or a handbag or under the keyboard of a workstation. It is also easy to forget to strike out a key that has already been used, which results in a fruitless effort of using the same key again, which in turn causes an unnecessary "attempted break-in" alarm in the security provider's system.

Arrangements are also known in which a key or a list of keys are stored on a portable memory medium, like a smart card. The smart card can be the one used in a mobile station as a SIM (Subscriber Identity Module), or it can be a separate entity for which there is a smart card reader in the mobile station or computer. If such stored keys are any time readily available for the user, they again arouse the problem of stolen or seized devices that were already switched on at the moment when the unauthorised party took control. As a precaution it is typical that the smart card or similar storage means require the user to punch in a key each time when the contents thereof are accessed.

The drawbacks of the smart card based arrangements are related to the problem of distributing the keys. Even the best of cryptographic systems can be broken, if a single key or a limited set of keys are used for too long. Therefore it is necessary to every now and then provide the user with new key(s). The user may find it inconvenient if he has to repeatedly visit an authorised office or to change smart cards in order to maintain a topical set of keys.

Yet other approaches are known that are based on digital certificates issued by a third party generically known as the CA (Certificate Authority). These approaches rely on the use of pairs of keys, referred to as the public key and the private key. After a certain party has reliably identified himself to the CA, the latter issues a digital certificate, which is a piece of digital data that has been encrypted with the CA's private key and serves as a digital "ID card" of the party having obtained it. A certificate as such does not solve any problems of reliably encrypting payload data, because the certificate only identifies its user.

It is an objective of the invention to provide a method and an arrangement for setting up a secure communication connection between a user terminal and a server with few requirements for user effort. It is another objective of the invention to provide such a method and arrangement that make it possible for the user to later check, whether he communicated with the service-providing party he believed he was communicating with. A yet other objective of the invention is to obviate the need for a Certificate Authority in setting up a reliable communication connection between certain two parties.

The objectives of the invention are achieved by utilizing two pieces of program code, known as the first midlet and the second midlet, in the user's mobile station or computer. The first midlet sets up or configures the initial user interface through which the user announces his request for setting up a secure communication connection. A subsequent request from the user to the service provider triggers a process of downloading the second midlet into the user's mobile station or computer. The second midlet generates the necessary keys and other pieces of confidential information in the user's mobile station or computer.

An alternative arrangement according to the invention separates the tasks of the first midlet and the second midlet more clearly so that the second midlet is only responsible for generating the necessary keys and other pieces of confidential information in the user's mobile station or computer, regardless of the way in which it was obtained. The task of the first midlet is then just to automatize the process of making and maintaining connections to the service provider.

A method according to the invention is characterised by the features recited in the characterising part of the independent claim directed to a method.

An arrangement according to the invention is characterised by the features recited in the characterising part of the independent claim directed to a method.

The invention applies also to a user's communication device and a server side arrangement, the characteristic features of which are recited in the characterising parts of the corresponding independent claims.

In describing a first aspect of the present invention we assume that at some earlier stage the user has obtained from the service provider (or from some other party approved by the service provider) a first piece or program code, which when executed in the user's mobile station or computer causes a certain simple user interface to be offered to the user. For the sake of simplicity we will designate said first piece or program code as the first midlet. The tasks of the first midlet include setting up (or configuring) the user inteface so that triggering the use of the service in question will require only minimal interaction by the user. As a response to such interaction the first midlet uses the capability of the user's mobile station or computer to contact the service provider. The connection request to the service provider includes some kind of an identification of the user.

When the service provider receives a connection request from the first midlet, it checks the identity of the user and transmits as a response a second piece of program code, known as the second midlet, to the user's mobile station or computer. Along with the second midlet the service provider transmits also its own public key. When the second midlet has arrived at the user's mobile station or computer, it starts executing and generates, at the user's mobile station or computer, certain keys that are necessary the user's end for enabling secure communication connection with the service provider. The second midlet makes access to a secure storage area within the user's mobile station or computer and stores the generated keys there. It encrypts the user's newly generated public key with the service provider's public key and sends it to the service provider. After that secure communication between the two parties may commence.

As a part of its operation the second midlet (or any other part or program code that takes advantage of the keys generated by the second midlet) may generate a so-called salt, which is a piece of information the generation of which involves a suitable number of cryptographic operations to ensure that it can only have been generated by this particular user. As a part of the secure communication connection set up by the above-explained process, the user's mobile station or computer may send the

salt to the service provider – or what the user thinks at this moment is the service provider. Later if the user wants to ensure that the other communicating party was indeed the service provider in question, he may use other means to contact the service provider and request it to produce the salt the user has transmitted earlier.

In a second aspect of the invention the user may have obtained the "second midlet" into his mobile station or computer even before the first midlet, or simultaneously therewith for example at the stage when he registered himself as a registered user of the service for the first time. In this case the user initiates a connection to the service provider by using the user interface features modified by the first midlet. If at that stage the user is already in possession of the second midlet and the necessary keys, the connection request may already contain the user's public key, possibly encrypted with the service provider's public key. Alternatively the connection request may take place as was described earlier in association with the first aspect of the invention, and only the response from the service provider would trigger generating the keys or fetching previously generated keys from a secure storage area but this time with the second midlet that was already residing in the user's mobile station or computer, not necessitating downloading one from the service provider's side. After the keys have been generated or fetched, the process continues as in the first aspect of the invention.

The novel features which are considered as characteristic of the invention are set forth in particular in the appended claims. The invention itself, however, both as to its construction and its method of operation, together with additional objects and advantages thereof, will be best understood from the following description of specific embodiments when read in connection with the accompanying drawings.

Fig. 1 illustrates a method according to an embodiment of the invention,

fig. 2 illustrates a user's mobile station or computer according to an embodiment of the invention,

fig. 3 is a state machine representation of a certain first midlet according to an embodiment of the invention,

fig. 4 is a state machine representation of a certain second midlet according to an embodiment of the invention,

fig. 5 illustrates a service provider's side arrangement according to an embodiment of the invention,

fig. 6 is a state machine representation of a certain server process according to an embodiment of the invention, and

fig. 7 is a state machine representation of another server process according to an embodiment of the invention.

The exemplary embodiments of the invention presented in this patent application are not to be interpreted to pose limitations to the applicability of the appended claims. The verb "to comprise" is used in this patent application as an open limitation that does not exclude the existence of also unrecited features. The features recited in depending claims are mutually freely combinable unless otherwise explicitly stated.

Fig. 1 illustrates certain operations performed at a user's end and at the service provider's end as well as between the user and the service provider according to an embodiment that conforms to what has been said above about the first aspect of the invention. Step 101 is a preparatory step at which the user registers himself as a registered user of certain service. The level of certainty at which the service provider must identify the user at this step depends on the overall level of security that will be required for the use of the service in question. If the service is e.g. a recreational game with no monetary value involved, this step may be trivial or even dismissed with. On the other hand if the service is a financial one or otherwise involves considerable attractiveness for dishonest behaviour, very strict requirements may be placed. For example, the user may be requested to personally visit an authorised office and to produce an official identification document. Between these extremes there are intermediate forms of registration like using a cellular telephone to call a number, in which case the CLI (Caller Line Identity) number received by the service provider could be taken as a proof of identity. It is also possible that the user has a previous long-term relationship with the service provider, on the basis of which the latter considers the user to have been identified without the user himself actively doing anything.

At step 102 the service provider or some third party approved by the service provider gives the user a certain first piece of program code, which is here designated as the first midlet. The word midlet (or MIDlet) as such is a general designation for Mobile Information Device Profile application and is understood to mean a piece of program code that can reside in the memory of a mobile station, enter an active state in which it performs certain operations utilizing the resources of the mobile station, and exit into a passive state after it has executed. A more Java-oriented description of a MIDlet is " a set of classes designed to be run and controlled by the application management software via an interface". The present invention is not limited to us-

ing midlets in their strictly standardised form. The word "midlet" is used here merely to reflect the fact that at the time of writing this description midlets were the most readily apparent widely known form of reducing the invention into practice.

The purpose of the first midlet is to make it as easy as possible for the user to start using a service offered by the service provider. In order to achieve this goal the first midlet comprises, as a certain part thereof, means for modifying the interactions between a user interface and a processor at the mobile station so that after such modification has been done, the mobile station knows that a certain start command given by the user, such as clicking on a certain icon, making a certain selection from a menu or pressing a certain key under certain circumstances for a certain time, means that the user wants to use a certain service. In fig. 1 step 102 comprises also substeps of using the first midlet to modify the user interface of the user's mobile station or computer, if such are needed.

At step 103 the user expresses his wish to start using the service of the service provider. Between steps 102 and 103 a long time may have been passed; after the first midlet has been loaded into the user's mobile station or computer it may reside there in a passive state for a long time before it is needed. In more detail, at step 103 the user gives the command that the mobile station or computer now knows to associate with an activation of the first midlet. For security reasons step 103 may involve a requirement for the user to enter a PIN number or other kind of identification.

At step 104 the first midlet controls the mobile station or computer to transmit a connection request to the service provider. This connection request contains an identifier of the user, which identifier the service provider will be able to associate with a certain user for which a registration according to step 101 has previously been performed. Contact information that is needed for properly directing the connection request to the service provider was typically included in the first midlet or in an information package downloaded simultaneously with the first midlet. In low-security applications where step 101 has been completely omitted and the user has e.g. downloaded the first midlet from a website or received it as a built-in feature of his mobile station or terminal, the connection request at step 104 may contain any user identifier, which the service provider will just accept and use as the "reliable" identifier of this user in later steps of the procedure. Other possibilities include, but are not limited to, making the transmission of step 104 as a telephone call and using the CLI transmitted therein as the user identifier, and using a certificate provided by an external Certificate Authority as the identifier of the user. The last-mentioned al-

ternative is naturally in conflict with one of the further objectives of the invention, i.e. not having to rely on external CAs.

At step 105 the service provider receives the connection request, reads the identifier contained therein and makes a verification in order to associate the user with the correct identity stored in the service provider's system in a previous registration step. Again in low-security applications step 105 may involve setting up a new user account that will be associated with an (made-up) identifier contained in a connection request.

After the user has been identified, the service provider makes a transmission at step 106, transmitting to the user a certain second piece of program code – known here as the second midlet – as well as a public key of the service provider. The last-mentioned is not absolutely necessary in all possible applications of the present invention, but it helps in achieving confidentiality and is mandatory if we require that the service provider must be able to digitally sign information so that the user may verify its origin. As an alternative to enclosing the service provider's public key into the transmission of step 106 the user may have received it together with the first midlet at step 102, or he may have obtained it from other, public sources.

At step 107 the second midlet starts executing at the user's mobile station or computer. The execution may begin automatically, if the device in question supports automatic execution, or it may require a brief interaction by the user in the form of e.g. clicking an icon or pressing a key or a few keys. The first midlet, which may still be running in the user's mobile station or computer, may have a part to play in starting the execution of the second midlet: for example it may prompt the user to perform certain action through which execution is started. A primary task of the second midlet is to generate certain keys that the user will need in communicating securely with the service provider. If we assume that PKI (Public Key Infrastructure) is in use, the second midlet generates a public key / private key pair for the user. It is also possible that the second midlet generates only one key. As seed information to the pseudo-random process of key generation the second midlet may use information it reads from somewhere in the user's mobile station or computer, like random key presses the user is prompted to make or noise received through a receiver of the user's mobile station or computer.

We assume that the user's mobile station or computer has a certain secure storage area where stored information is protected from unauthorised tampering. The sec-

ond midlet, while executing at step 107, stores the generated keys into such a secure storage area.

At step 108 one of the first and second midlets controls the operation of the user's mobile station or computer so that it makes a further transmission to the service provider, this time transmitting a newly generated cryptographic key of the user to the service provider along with an identifier of the user (preferably the same that was used at step 104) so that the service provider can recognise the sender. If a public key / private key pair was generated, the transmission of step 108 contains the public key. If a single key was generated, it is contained in the transmission of step 108. At least in the last-mentioned case it is mandatory that the user has previously received an encryption (public) key of the service provider and it is used for encrypting the transmission, because otherwise a valuable encryption key of the user would be transmitted in plaintext. Even if only a public key is transmitted at step 108, it is most recommendable to encrypt the transmission with the public key of the service provider.

At step 109 the service provider receives the key sent by the user and stores it securely for use in communication with this particular user. At step 110 the service provider responds with an acknowledgement message that tells the user that the initiating procedures have been completed, after which secure communication may begin at step 111. During the secure communication the user's mobile station or computer uses the public key of the service provider it received at some step of the process (e.g. step 106) as well as its own key(s) that the second midlet generated at step 107. The service provider uses the key(s) it received at step 108.

Certain important details of the process described above are the following:

1) Key generation at the user's mobile station or computer, with means obtained directly from the service provider. This way the service provider has direct control over the way in which keys are generated at the user's end. On the other hand generating all user's keys at his own device obviates the need of ever having to transport keys from outside to the user's mobile station or computer, either on portable memory means or through a communication connection. If the process of generating the keys can be made fully automatic, the user may even be kept unaware of the fact that encryption keys are generated and used. The last-mentioned feature is a definite advantage, since users have certain distrust of complicated operations when using their devices.

2) Nonsensitivity to features of the transmission channel or content of the information. There are no requirements for encryption or other security features being embedded in the transmission system, so for example the Internet can be used. On the other hand security features inherent in the transmission system do not complicate the procedure either. Additionally, once the key generation and exchanging procedure has been completed, the keys enable encrypting and decrypting any kind of information in all uni- or bidirectional communication between the user and the service provider regardless of its content or amount.

3) Use of a secure storage space at the user's mobile station or computer. We assume that the secure storage space has features that as such protect all information stored therein against unauthorised tampering. How such features have been implemented in practice is outside the scope of the present invention. Typically there is an API (Application Programming Interface) towards the secure storage space that requires the user to key in a PIN or similar code every time the secure storage space is accessed.

4) The possibility of using the generated keys for also other purposes than direct encryption and decryption at a transceiver interface. Once the keys have been generated and stored in the secure storage space, they are available for e.g. applications running in the mobile station or computer, so that they can use them for purposes of their own. One possibility is to use one secret key of the user to generate a so-called salt, which is a cryptographically protected piece of information that only the user himself can have created. If the user transmits the salt over the communication connection to what he thinks is a certain service provider and the last-mentioned stores it for future reference, the user may later use some other form of communication to request the service provider to produce a copy of the salt it received. This way the user may later ensure that he was communicating with the appropriate party.

All cryptographic applications involve the question of validity. In order to be secure, a certain generated key or certain used encryption method must expire, i.e. cease to be valid, after a certain number of times it was used and/or after a certain period of time. Regarding the validity of the keys generated by a second midlet as a result of going through steps 101-106 once, the parties may agree upon a common policy. For example according to a strict approach the keys are only valid for one connection, after which at least steps 103-106 must be repeated before a new connection can be set up. A key itself may also contain information about how long it

can be used, in which case a process that attempted using an expired key may halt and prompt the user for initiating the downloading of a new key-controlling midlet.

Fig. 2 is a block diagram of a user's mobile station or computer according to an embodiment of the invention. A transceiver 201 is arranged to be available for bi-directional communication with the service provider. A processor 202 constitutes the functional core of the apparatus, executing programs stored in a program memory 203. Payload data that are needed during operation are a stored in a data memory 204, and communication with a user is conducted through a user interface 205, physical part of which are output means (e.g. display) 206 and input means (e.g. keypad; possibly also a serial port, Bluetooth connection or the like) 207. One of the program memory 203 and data memory 204 includes an electronic safe storage (not shown). The tasks of the various parts of the user's mobile station or computer are highlighted in more detail in the following, where the operation of an exemplary first midlet and an exemplary second midlet is illustrated with reference to figs. 3 and 4.

Fig. 3 is a state machine illustration of the operation of a certain first midlet according to an embodiment of the invention. Installing the first midlet takes place after having received it either through a transceiver (201 in fig. 2) or through a local communications port included in the user interface (205 in fig. 2). The installation includes possible reconfiguring of the user interface features so that the possibility of giving a simple start command, like pressing a key or clicking an icon, is offered to the user. After installation the first midlet is in a passive state 301. A start command received through the user interface causes the first midlet to go into an ID request state 302, at which it prompts the user to give some kind of identification. The reason for including such a step in the operation of the first midlet is the need for ensuring that only the legitimate user can use this particular device and this particular midlet for commencing secure communication with the service provider. The identification given at state 302 may be a PIN code, some other character string type identification, a biometric identification such as a fingerprint or retinal scan result, to name a few. In low-security applications state 302 may be a trivial one where the mere start command is taken to represent "user identification". State 302 may even be completely omitted, if it suffices to rely on inherent "user ID" features of e.g. a cellular radio system, such as the CLI that will be transmitted in the telephone network without any action of the first midlet.

Receiving the user identification causes a transition to a communication state 303, in which the user's mobile station or computer now transmits a connection request of the type described at step 104 of fig. 1. In the block diagram of fig. 2 this means that the processor 202, executing a first midlet read from the program memory 203, instructs the transceiver 201 to transmit the connection request.

In the embodiment illustrated in fig. 3 a positive response, which is received from the service provider and contains the second midlet as well as the service provider's public key, causes the first midlet to go into a background state 304 in which it just waits for the second midlet to execute. The background state means that the processor (202 in fig. 2) is not actively executing any steps instructed by the first midlet, but a location in the program memory 203 is known from the which the execution of the first midlet will continue at the occurrence of a certain interrupt or other triggering condition. After the second midlet has been completed and the necessary keys are available, a return to the communication state 303 occurs. The states of the state machine are assumed to be aware of history, i.e. what caused a transition thereto, which means that this time dwelling in the communication state 303 involves sending to the service provider the user's public key and possibly the salt, which was discussed earlier. A positive response from the service provider means this time the reception of an acknowledgement message in accordance with step 110 in fig. 1, after which the first midlet may again go into the background state 304 for the duration of the secure communication connection between the user and the service provider.

In fig. 3 we assume that the first midlet may have some tasks related to the closing of a connection, such as storing log data or the like. Therefore the acknowledgement message mentioned above caused a transition to the background state 304 and not to the passive state 301, and also therefore the ending of the communication connection causes a further transition to a connection closing state 305, from which the first midlet returns to the passive state 301 after the closing time routines have been completed. If the responsibilities of the first midlet end altogether at the moment when the connection has been successfully set up, the state machine diagram included a direct transition from the communication state 303 to the passive state 301 at the moment of receiving the acknowledgement message.

An error state 306 has been defined into which a transition occurs if at any other state the normal operation does not proceed smoothly. For example every other state may include a timer that measures the time spent in that state in the absence of any

state transitions. If a timer expires before any new transitions occur, transition to the error state 306 results. From the error state 306 the process returns to the passive state 301 after having reported about the occurrence and reason for the error.

Fig. 4 is a state machine representation of a second midlet according to an embodiment of the invention. In the following description we first assume, in accordance with the first aspect of the invention, that the second midlet is downloaded into the user's mobile station or computer from the service provider as a response to the connection request. The installation of the second midlet thus causes immediately the necessary keys to be generated at state 401. The process of generating the keys may involve acquiring random data through some way known as such. When the keys have been generated, a transition to state 402 occurs, where the second midlet attempts accessing the electronic safe storage. This state typically involves prompting the user for a PIN code. Once the access is granted, the keys are stored into the electronic safe storage at state 403, after which the second midlet has completed its task and exits through a possible closing state 404 to a passive state 405.

Previously we already pointed out that the second midlet is not necessarily a disposable piece of code, which the user would have to order from the service provider every time before commencing secure communication. Therefore the state machine of fig. 4 includes the possibility of a transition from the passive state 405 to the key generation state 401 at the occurrence of a suitable starting command. An application program at the user's mobile station or computer, the user himself or even the first midlet can give such a starting command, although the last-mentioned alternative would require certain changes to the definitions of midlets customary at the time of writing this description, because at present a midlet is not allowed to start another midlet in a mobile information device.

An error state 406 has again been defined for the same purposes as described in association with fig. 3.

Fig. 5 is a block diagram of an exemplary service provider's side arrangement according to an embodiment of the invention. A basic approach behind the structure suggested here is that the actual service providing server, for example a server of a banking system offering users information about their bank accounts and possibilities for performing transactions, can be accessed through a variety of channels with various communication protocols, terminal capabilities and so on. In order not to burden the service providing server with the requirements of knowing the detailed

specifications of all possible communication channels, there is provided an intermediate server that sets up the specific channels towards e.g. various types of user terminals and standardises the interface towards the service providing server. The last-mentioned only needs to know one way of communicating, which it uses for communicating with the intermediate server. The intermediate server makes all adaptations transparently so that the service providing server does not even need to know, what kind of a terminal it is communicating with, while the user terminal can use the way of communicating which is most natural to it.

In the representation of fig. 5 the intermediate server comprises a network interface 501, a processor (or a more complicated array of processing entities) 502, a program memory 503 and a midlet database 504. It may also comprise a user database 505. The service providing server comprises a processor (or a more complicated array of processing entities) 511, a program memory 512, a user verification database and process 513, a cryptographic operations database and process 514 and a service database and process 515. Between the intermediate server and the service providing server there is a local interface 521, which does not need to be "local" in physical sense; the designation only highlights the fact that it is an interface between entities that both belong logically to the service provider's side e.g. in the viewpoint taken in fig. 1. There may be also other interfaces 522 to the service providing server. These may be further coupled to other intermediate servers that make adaptations in accordance with some other channel offered for the users in utilizing the service providing server.

Fig. 6 is a state machine illustration of the operation of an intermediate server according to an embodiment of the invention. In the idle state 601 the intermediate server waits for transmissions from users. At the reception of a connection request a transition to state 602 occurs, in which the intermediate server extracts the user identification contained in the connection request. If some formal modification is required in order to make the presentation of the user identification conform to the way in which the service providing server wants to receive user identifications, these are performed before passing on the used identification to the service providing server. At that moment the intermediate server goes into a wait state 603 at which it waits for a positive response from the service providing server, indicating that the user has been identified and accepted for continuation of the process. An acceptance from the service providing server causes the intermediate server to go into a midlet lookup state 604, in which it selects a suitable "second" midlet to be transmitted to the user. The suitable midlet can be selected e.g. on the basis of pre-

viously stored information about the user or information received in the connection request. The criteria for suitability include factors like terminal type that affect the user's capability of generating and using keys, or explicitly expressed preferences.

After selecting a midlet and sending it to the user the intermediate server returns to the idle state 601. If now a second transmission from the user occurs, containing the user's key, the intermediate server again reads the user identification at state 602, sends it together with the key to the service providing server and enters the wait state 603. When the service providing server announces having received and accepted the key, the intermediate server goes into a tunnelling state 605 in which it only passes on all encrypted information transmitted between the user and the service providing server. When the connection is closing there may be a specific closing state 606 at which the intermediate server performs any necessary cleaning-up routines before returning to the idle state 601. An error state 607 takes care of any unexpected conditions as was the case with the midlets in the user's mobile station or computer.

How the tasks illustrated in fig. 6 are distributed among the functional blocks illustrated in fig. 5 is self-evident. The user database 505, if it exists, is for storing information about the users e.g. for the purposes of associating each individual user with the correct type of second midlet to be delivered.

Fig. 7 is a state machine representation of the operation of a service providing server according to an embodiment of the invention. It too has an idle state 701 in which it waits for transmissions from users or from the intermediate server. When it receives an identification checking request, which is the connection request from the user possibly modified in the intermediate server, it goes into a verification state 702 in which it checks, whether the user can be identified and accepted. When the service providing server has announced its acceptance, it goes back into the idle state 701. Receiving a key causes again a transition into the verification state 702, in which the received key is associated with the appropriate user and stored. Thereafter the service providing server goes into a service state 703, in which it provides the requested service by communicating with the user's apparatus. At the time of closing the connection the server goes through a closing state 704 back into the idle state 701. An error state 705 has again been defined for the same purposes as above.

The description above has mainly assumed that at the moment of transmitting the connection request the user's mobile station or computer does not possess the sec-

ond midlet or the necessary keys generated by the second midlet. Next we may consider, how would the operations described in association with figs. 3, 6 and 7 differ from that described above if the second midlet already existed in the user's mobile station or computer at the moment of starting the first midlet. This assumption forms the basis of the second aspect of the present invention.

Let us first assume that one full round through the state machine diagrams of figs. 3, 4, 6 and 7 has been made. As a result, all processes are back in their passive or idle states, a second midlet exists stored within the program memory of the user's mobile station or computer, and certain keys exist stored within an electronic safe storage within the user's mobile station or computer. Now the user gives a start command that causes a transition from state 301 to state 302 in the state machine of the first midlet (fig. 3). The purpose of state 302 was to enhance security, and no changes are required there. A received piece of identification information from the user causes a transition to state 303, in which the first midlet prepares and transmits a communication request to the service provider.

Communication resources are usually not to be wasted, which means that the service provider should not send an unnecessary copy of the second midlet if the user already has one. So in order to account for the already existing second midlet we may require that before sending the connection request at state 303 the first midlet checks from a register of some kind, whether a second midlet is needed from the service provider's side. It is not needed if at least one of the following conditions are true:

1) a valid and usable copy of the second midlet already exists in the user's mobile station or computer

2) the necessary keys exist in an electronic safe storage of the user's mobile station or computer and are valid.

Thus the communication request transmitted at state 303 should contain an information element that tells the service provider's arrangement, whether a second midlet is actually needed or not. A positive response from the service provider causes a transition to state 304 like in the basic case discussed earlier, but the criteria for a transition back to state 303 must be augmented as follows:

- A second midlet freshly obtained from the service provider has executed (like in the previously described case) OR
- a second midlet that already existed in the user's mobile station or computer is called and executed OR

- it is noted that all necessary keys exist already and no second midlet is needed at the moment.

Thereafter the process continues in accordance with what has been described earlier in association with the first aspect of the invention.

Regarding the operation of the intermediate server illustrated in fig. 6, only the descriptions of state 604 and the transition therefrom back to state 601 need to be augmented. At that state (or based on a decision made already in some previous state in which the intermediate server analysed the connection request from the user) the intermediate server decides, whether it transmits the second midlet to the user or whether it only transmits an indication about the connection request having been accepted. The second midlet is transmitted in those cases where the connection request did not contain an indication about the second midlet not being needed.

The operation of the service providing server, as illustrated in fig. 7, stays the same regardless of whether the first or the second aspect of the invention is discussed.

Finally, regarding the second aspect of the present invention, we may loosen the assumption according to which the second midlet had been received specifically by going through one round of operation according to the first aspect of the invention. The first midlet may be completely independent on the way in which the second midlet has been obtained, as long as there exists an easy and flexible way of utilising the process for obtaining one in case none is currently available or a new copy is needed because the old one has become obsolete.

The invention has been described above with reference to examples, which should not be construed as restrictive. For example, the designation "midlet" can be generalised to cover any forms of program code that cause a programmable device to perform the required operations. For example in the future it is well possible that the functionality referred to herein as the first midlet will be implemented in a browser program tailored for mobile terminals, configurable features of which include but are not restricted to the possibility of associating certain simple commands from the user with transmitting connection requests to certain well-defined service providers as well as the possibility of requiring the user to produce some kind of user identification before transmitting such connection requests, which user identification will then be placed into the connection request for the service provider to be able to positively identify the user. Another point at which the literal expressions used in

the description should have a restrictive meaning is the question of an intermediate server and a separate service providing server. In some embodiments of the invention the described processes may well be implemented in a single server.

A class of interesting high-security variants of the embodiments discussed so far can be obtained by not storing at all the public key generated by the second midlet at the user's mobile station or computer. Let us recall that the public key will be transmitted to the service provider, which will use it to decrypt messages that the user encrypted with his corresponding private key. Let us then assume that the user's mobile station or computer will store the private key in encrypted form, from which it can only be decrypted with the help of a correct PIN code given by the user. Now, if the mobile station or computer ends up in dishonest hands, a hacker may try a brute force attack against the encrypted private key by sequentially going through all possible PIN codes, thus obtaining all possible decrypted forms, one of which is the correct, decrypted private key. If the mobile station or computer also contains the public key, or the public key can otherwise be reliably obtained, the hacker may sequentially use each of the decrypted "private keys" he obtained for encryption and the public key for decryption of a model message, until he finds the correct private key by noticing that the model message encrypted therewith allows itself to be correctly decrypted by the public key.

After thus having found the correct private key, the hacker could contact the service provider, use the private key he found for encryption, and get acknowledged by the service provider as the legitimate user.

An effective way of blocking the way of the hacker is to ensure that the mobile station or computer does not contain any information with which the hacker could check, whether he has obtained the correct private key. In other words, after the second midlet has created a public key / private key pair and the public key has been sent to the service provider, all traces thereof must be erased from the mobile station or computer. Also the second midlet and the encryption algorithm used to encrypt the private key must be made irreversible, meaning that it is not possible with reasonable resources to use them to decrypt the encrypted private key or to obtain the correct form of the public key. To be exact, the "public" key must not become public at all; it must remain strictly at the possession of the service provider and nobody else. The technology of irreversible encryption and key generation algorithms has been thoroughly investigated, and both general principles and detailed implementations thereof are widely available.

Not storing the plaintext private key or the public key at the mobile station or computer does not keep the potential hacker from performing his brute force decryption attack, i.e. sequentially trying through all possible PIN codes. However, the hacker has no other way of verifying the result than using a decrypted form of the private key for signing a message and transmitting that message to the service provider. The last-mentioned will only acknowledge the message if the hacker managed to guess right in selecting the PIN, and the decryption was successful. Most likely it will take him a large number of attempts to guess right. The service provider must keep track of all connection set-up attempts associated with a certain user account, and refuse all further attempts after a small number (e.g. 5) of unsuccessful ones.

Claims

1. A method for setting up a secure communication connection between a user's mobile station or computer and a service provider's arrangement, comprising the steps of:

- transmitting from the user's mobile station or computer to the service provider's arrangement a connection request that contains an identification of the user placing the connection request,
- responding to the connection request by transmitting from the service provider's arrangement to the user's mobile station or computer a response message,
- transmitting from the user's mobile station or computer to the service provider's arrangement an encryption key to be used in communication between the user's mobile station or computer and a service provider's arrangement, and
- thereafter commencing secure communication between the user's mobile station or computer and the service provider's arrangement;

characterised in that it comprises the steps of:

- conveying from the service provider to the user's mobile station or computer a computer program for generating cryptographic keys and
- using said program in the user's mobile station or computer to generate cryptographic keys for use in the secure communication between the user's mobile station or computer and the service provider's arrangement.

2. A method according to claim 1, **characterised** in that the step of conveying from the service provider to the user's mobile station or computer a computer program for generating cryptographic keys is performed as a response to the connection request transmitted from the user's mobile station or computer to the service provider's arrangement, after the service provider's arrangement has identified the user on the basis of the identification of the user contained in the connection request.

3. A method according to claim 1, **characterised** in that

- the step of transmitting from the user's mobile station or computer to the service provider's arrangement a connection request comprises the substeps of:
  - checking, at the user's mobile station or computer, whether a computer program for generating cryptographic keys is needed from the service provider's arrangement and
  - inserting into the connection request an information element that tells the service provider's arrangement, whether a computer program for generating cryptographic keys is needed; and

- the step of conveying from the service provider to the user's mobile station or computer a computer program for generating cryptographic keys is performed as a response to the connection request, if said information element tells the service provider's arrangement that a computer program for generating cryptographic keys is needed.

4. A method according to claim 3, **characterised** in that in the step of checking whether a computer program for generating cryptographic keys is needed it is decided that said computer program is not needed if at least one of the following statements is true:
- a valid and usable copy of a computer program for generating cryptographic keys already exists in the user's mobile station or computer
- certain necessary keys exist in an electronic safe storage of the user's mobile station or computer and are valid.

5. A method according to claim 1, **characterised** in that the step of conveying from the service provider to the user's mobile station or computer a computer program for generating cryptographic keys involves conveying a midlet for controlling the operation of a mobile information device.

6. A method according to claim 1, **characterised** in that the step of using said program in the user's mobile station or computer to generate cryptographic keys involves encoding a validity period into the generated cryptographic keys, indicating a period during which it is allowable to use said cryptographic keys in secure communication, thus implicitly determining when a next time of using a program in the user's mobile station or computer to generate cryptographic keys will become topical.

7. A method according to claim 1, **characterised** in that the step of conveying from the service provider to the user's mobile station or computer a computer program for generating cryptographic keys involves encoding a validity period into the conveyed computer program, indicating a period during which it is allowable to use said computer program in the user's mobile station or computer for generating cryptographic keys, thus implicitly determining when a next time of conveying from the service provider to the user's mobile station or computer a computer program for generating cryptographic keys will become topical.

8. A method according to claim 1, **characterised** in that it additionally comprises the steps of:
- generating at the user's mobile station or computer a cryptographically protected piece of information, the cryptographic protection of which is of a kind ensuring that only a user possessing a certain key may have generated it,
- transmitting said cryptographically protected piece of information from the user's mobile station or computer to the service provider's arrangement and
- storing such transmitted cryptographically protected piece of information in the service provider's arrangement.

9. A method according to claim 1, **characterised** in that it comprises the steps of:
- receiving the connection request at a first functional entity within the service provider's arrangement,
- reading a user identification from the received connection request,
- transmitting the user identification to a second functional entity within the service provider's arrangement,
- verifying a user identity at the second functional entity and announcing acceptance to the first functional entity, and
- converting the announced acceptance into a positive response to be transmitted to the user's mobile station or computer, and
- transmitting the positive response from the first functional entity to the user's mobile station or computer.

10. A method according to claim 9, **characterised** in that the step of converting the announced acceptance into a positive response to be transmitted to the user's mobile station or computer comprises the substeps of reading a computer program for generating cryptographic keys from a database of such computer programs and including the computer program for generating cryptographic keys into the positive response.

11. A method according to claim 10, **characterised** in that the substep of reading a computer program for generating cryptographic keys comprises selecting said computer program on the basis of information that was contained in the connection request.

12. A method according to claim 9, **characterised** in that it comprises setting up a standardised interface between the first and second functional entities, so that the

operation of the first functional entity comprises making transformations between a communications format used over said standardised interface and a communications format used between the first functional entity and the user's mobile station or computer.

13. A method according to claim 1, **characterised** in that after generating a pair of interrelated asymmetric cryptographic keys it comprises the steps of
- sending one of said interrelated asymmetric cryptographic keys to the service provider,
- erasing all traces of said one of said interrelated asymmetric cryptographic keys from the mobile station or computer,
- enctrypting the other of said interrelated asymmetric cryptographic keys and
- storing the encrypted one of said interrelated asymmetric cryptographic keys in the mobile station or computer in encrypted form.

14. An arrangement for setting up a secure communication connection, comprising:
- a user's mobile station or computer and a service provider's arrangement,
- within the user's mobile station or computer means for transmitting to the service provider's arrangement a connection request that contains an identification of the user placing the connection request,
- within the service provider's arrangement means for responding to the connection request by transmitting to the user's mobile station or computer a response message, and
- within the user's mobile station or computer means for transmitting to the service provider's arrangement an encryption key to be used in communication between the user's mobile station or computer and a service provider's arrangement;
characterised in that it comprises:
- means for conveying from the service provider to the user's mobile station or computer a computer program for generating cryptographic keys and
- within the user's mobile station or computer means for using said program to generate cryptographic keys for use in the secure communication between the user's mobile station or computer and the service provider's arrangement.

15. A user's mobile station or computer equipped for setting up a secure communication connection with a service provider's arrangement, comprising:
- means for transmitting to the service provider's arrangement a connection request that contains an identification of the user placing the connection request, and

- means for transmitting to the service provider's arrangement an encryption key to be used in communication between the user's mobile station or computer and a service provider's arrangement;

characterised in that it comprises:

- means for receiving from the service provider a computer program for generating cryptographic keys and
- means for using said program to generate cryptographic keys for use in the secure communication between the user's mobile station or computer and the service provider's arrangement.

16. A service provider's arrangement equipped for setting up a secure communication connection with a user's mobile station or computer, comprising:

- means for receiving from the user's mobile station or computer a connection request that contains an identification of the user placing the connection request,
- means for responding to the connection request by transmitting to the user's mobile station or computer a response message, and
- means for receiving from the user's mobile station or computer an encryption key to be used in communication between the user's mobile station or computer and a service provider's arrangement;

characterised in that it comprises:

- means for conveying to the user's mobile station or computer a computer program for generating cryptographic keys for use in the secure communication between the user's mobile station or computer and the service provider's arrangement.