A comparison of mobile authentication methods

Tuomas Launiainen
Helsinki University of Technology
Tuomas.Launiainen@tkk.fi

Abstract

Authentication, or the proof of identity of one or both parties of a communication session, is often a crucial part of the session, and without it the session cannot proceed. Authentication is especially important in mobile payment scenarios, where it has immediate, monetary value. The authentication mechanisms must be resistant to attacks where someone tries to fool one or both parties of the communication session by intercepting and altering the messages between them. Privacy concerns must also be taken care of, so that unnecessary information about parties is not leaked to anyone who is not intended to receive it.

This paper reviews the current state of the art for authentication techniques used with mobile devices. Digital signatures are an important tool with authentication methods, and their generation on mobile devices is covered. Two specific authentication techniques are covered in greater detail, and their advantages and disadvantages are compared with each other and with general authentication methodology.

KEYWORDS: mobile authentication, mobile payment, digital signatures

1 Introduction

Reliable and trustworthy user authentication is essential in many electronic services today. From online banking and shopping to changing one's official postal address, it is vital to be sure of the identity of the user and the service provider. A hierarchical structure of cryptographically signed certificates is an efficient solution to identifying the service provider since there is a sufficiently small number of service providers and since they can be expected to apply and pay for a certificate. Identifying a user is more problematic, however, and the most used method is using a username and a password. This simple approach leaves room for improvement because the user needs to manage usernames and passwords for many services. Moreover, usernames and passwords are prone to so called phishing where a malicious party contacts the user pretending to be a representative of a service and asks for the authentication credentials.

Alternate methods of identifying the user have been developed, but none have replaced the password authentication as the most common. Cryptographic signatures are a promising technology on the user end as well, but many of the solutions developed so far suffer from portability issues or are too difficult to deploy. An authentication method that uses a mobile phone as an authentication provider is presented in [2]. It aims to be easy for the user and more secure than the traditional password authentication. An approach that focuses on mobile payment is taken in [3], and two payment protocols are introduced. Finally, [7] reviews the current state of digital signatures on mobile phones from a more general perspective.

The rest of this paper is divided as follows: Section 2 introduces the terminology and concepts used, Section 3 discusses different views of identity, Section 4 gives an overview of what the state of the art in mobile authentication is, Sections 5 and 6 review in more detail two technologies of mobile authentication, Section 7 compares their strengths and weaknesses, and Section 8 presents the concluding remarks.

2 Terminology and concepts

Authentication and Key Agreement (AKA) is the authentication mechanism used in 3G mobile networks such as the Universal Mobile Telecommunications System (UMTS). It is defined in the Internet Engineering Task Force's RFC 3310 [4].

Remote Authentication Dial In User Service (RADIUS) is a commonly used authentication protocol that uses a remote server to manage the authentication of users. It can be used for many kinds of authentication scenarios. RADIUS is defined in the Internet Engineering Task Force's RFC 2865 [5].

Diameter is a protocol that is designed to replace RADIUS. It addresses some security concerns in RADIUS, adds more flexibility, and has better support for roaming. It is defined in the Internet Engineering Task Force's RFC 3588 [6].

Public Key Cryptography is a method of encryption that uses a pair of keys: a public and a private one. The public key is used to encrypt the message that can then be decrypted only with the corresponding private key. Thus, the public key can be freely distributed, and anyone can use it to encrypt messages that only the holder of the private key can decrypt.

Digital signature is another use of public and private keys. A signature is created from the contents of a message using the private key, and then added to the message. The public key can then be used to verify that the signature was created with the corresponding private key, and that the contents of the message has not been tampered with.

Public Key Infrastructure (PKI) is a system for determining the trustworthiness of digital certificates. A Certificate Authority (CA) distributes its public key to all participants of the system and then signs certificates of trusted certificate holders with its private key. Other users can then verify the signature and decide to trust the certificate, even though they have no previous contact with the holder of the certificate.

A cryptographic hash function is an algorithm that takes an arbitrarily large amount of data and computes a short hash value from that data. With a good hash function, it should be unfeasible to obtain information about the data by just looking at the hash value, and it should be unfeasible to find two different data inputs that result in the same hash value. Cryptographic hash functions are often used with digital signatures since signing a message will result in a signature that is as long as the message. By computing a hash value from the message and then signing the hash value a significantly shorter signature is generated.

3 Identity

Before examining in detail what solutions exist for mobile authentication, we must define the desired end result, establishing the identity of the user. The identity of the user can be an ambiguous concept, however. In the case of making a payment, for example, the identity could refer to the person making the payment, or it could refer to a credit card or a debt card with which the payment is made. In the first case the result of the authentication must be some proof that the user is who they claim to be, but in the second case it is sufficient to have some proof that the user is someone who is entitled to make payments with the card in question. Usually this is equivalent to knowing the PIN number or security code of the card.

An identity can also be viewed from the perspective of information routing. In this case the identity that needs to be authenticated is the identity of the device. When a device initiates a connection to a wireless network, for example, it is often required to prove that it is entitled to use the resources of the network, and therefore must authenticate itself. The connection can also be used to deliver messages to the device after the authentication. Email notifications, for example, can be sent to the device after it is known that the device can be reached through that connection. This kind of scenario is a common one with mobile devices because their connections can change along with their physical location.

In this paper an identity is not limited to any single scenario. An identity is some property of one participant in a communication session that entitles that participant to something: access to a service, a piece of information or a resource that is not available to everyone. Authentication means proving that the participant has that property. Sometimes non-repudiation of a message is required in addition to authentication. This means that the sender of a message cannot later deny that it has sent the message. Non-repudiation is very closely related to authentication, as it can be achieved by making the identity unique, and making the proof of identity universal, meaning that anyone can verify the identity from that proof. Therefore, generally non-repudiation techniques can also be applied to authentication.

Another issue that needs to be considered with identities is privacy. Many details of a participant in a communication session should be kept secret from anyone but the actual receiver of the information, and some details should be kept secret from anyone but the participant itself. Examples of such details are the authentication credentials, payment information, and activities that are not related to the particular communication session.

4 State of the art

Digital signatures are a very useful tool for authentication and non-repudiation. Both can be achieved by having a pair of public and private keys, the latter being accessible only to a single party, and then doing the authentication by having that party generate a signature of some data. Provided that the keys are managed properly, and that the public key is globally accessible, anyone can verify that a public key is the correct one and that the signature was generated by that key. This scheme is also very flexible, since the actual meaning of a signature can be defined by the protocol being used between the communicating parties.

An exhaustive review of signature generation methods with mobile devices is given in [7]. Particular attention is paid to the requirements set by the European Parliament and the Council to digital signatures that can be viewed as equivalent to hand-written signatures. The requirement are that the signature must be based on asymmetric cryptography or elliptic curve cryptography, the signature must be created with a Secure Signature Creation Device, meaning that the private keys must be protected so that they cannot be extracted from the device, and finally that the certificates for the signature must be obtainable from a qualified certificate authority.

The paper concludes that while there are many good ways of generating digital signatures with mobile devices, the ones that can be compared to hand-written signatures and that are most convenient for application providers to use are the solutions that store the private keys on the SIM card of a mobile phone. Because the SIM cards are tamper-resistant devices, the private keys are secured in the SIM card's memory. The signature is also created on the SIM card itself, so the private key cannot be extracted in any phase of the process. Moreover, the interface to use the key pair on the SIM card has been standardised, making it relatively easy to develop applications that utilise the process on multiple mobile platforms. Further, because the public and private key pair can be issued by a qualified certificate authority and shipped together with the SIM card, this approach is clearly the best available way of handling mobile signature generation, since no other way provides all the aforementioned features, nor do they provide any features that offer a clear benefit over this.

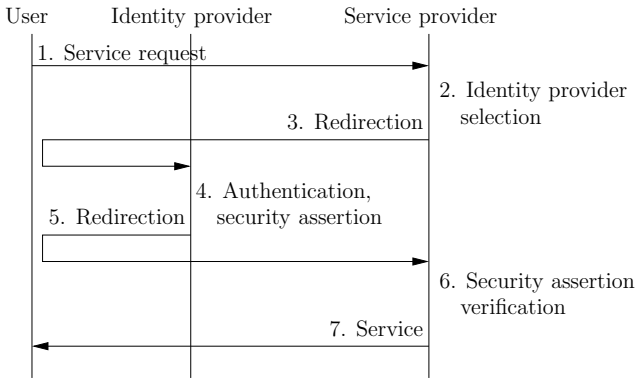

Figure 1: The authentication procedure with the mobile identity provider.

5 Identity provider on the mobile phone

The prototype implementation in [2] utilises the signature capabilities of a SIM card. The primary idea of the implementation is to replace an ordinary username and password as an authentication method on a web page. The main improvements over password authentication are that the user need not remember their password, which improves usability in addition to security, and that authentication can be done even in scenarios where the service provider has no prior knowledge of the user. Some additional security features are also discussed in [2].

The system works by acting as an identity provider component that is defined in the Single Sign On procedure of the Liberty Alliance [1]. The HTTP POST profile of the identity provider authentication is used, which consists of seven steps:

1. The user initiates a connection to the service provider, requesting some service.
2. The service provider selects an identity provider by some means.
3. The service provider redirects the user to the identity provider.
4. The identity provider authenticates the user, if it has not already done so, and creates a security assertion that functions as a proof of identity of the user.
5. The identity provider redirects the user to the service provider and adds the security assertion from the previous step to the request.
6. The service provider verifies the security assertion.
7. If the security assertion was correct, the service provider grants the requested service to the user.

Figure 1 illustrates the transaction.

At the heart of the authentication procedure is the security assertion. It contains a digital signature of the identity provider, and the service provider validates the assertion based on the signature. This means that the service provider must have access to the public key of the identity provider and some assurance that the public key is the correct one. The signing is done with a private key that is pre-loaded on the SIM at the time it was manufactured, and the corresponding public key on the SIM card is signed by the operator that grants the SIM card. Therefore, the public key can be sent to the service provider along with the security assertion, and the service provider can make sure that the public key is the correct one by verifying the operator's signature on it.

The authentication system does have a few clear problems. One is that the mobile phone needs to function as an HTTP server, which is generally not possible with the network connections provided by operators. The solution proposed in [2] is to have a separate computer that functions as a proxy, acting as the server and communicating with the mobile phone to get the security assertion. The other problem is that the service provider needs to know how to contact the identity provider service on the mobile phone. The solution proposed to this problem is to have the proxy computer recognise special web addresses that contain the name of the user, and then having the user type such an address on the web page of the service provider. This seems to be a bit cumbersome process, but it does enable the service provider to identify the user without any prior knowledge of them.

The additional security features discussed in [2] are related to the scenario where a malicious party tries to trick the user into generating a security assertion that would authenticate the wrong person for the service provider. One possible way around such an attack is to display information on the mobile phone screen about the party for which the security assertion is made. This might not be very effective though, because many users ignore warnings that are displayed to them by computers or mobile phones. A better way is also presented: to have the mobile phone remember details about the user's computer, and when a security assertion that does not match those details is about to be generated, have the user directly access the identity provider on the mobile phone to verify the changed details.

6 Mobile payment with digital signatures

The authentication scheme in [3] is meant for making mobile payments. Therefore, in addition to proving the identity of the user, the system must take care of privacy issues and provide enough evidence about the transaction to all parties. To achieve that, public and private keys granted by the Finnish Population Register Centre are used. All parties of the payment, namely the customer, the merchant, and the customer's bank, are expected to have keys granted by the Population Register Centre, thus making sure that there is no question about the identities of the parties. The customer's key pair is stored on a SIM card, and the key pairs of the bank and the merchant can be stored however they see fit. The messages transmitted in a payment transaction are both encrypted and signed with the relevant keys in each case, and different parts of the message are encrypted with different keys to limit the amount of information each party gets to that which is actually needed by the party.

The main difference to other, existing mobile payment solutions is that this one does not require a separate mediator for the payments. A mediator usually handles a pre-paid account for customers and transfers the funds from that, also charging some amount of the payment to cover its own operational expenses. This increases the total cost of the transaction, so it is desirable to handle payments without a need for a separate mediator.

Two different scenarios are considered in [3]: a real point-of-sale, where the customer is physically present at a merchant's site, and a virtual point-of-sale, where the customer is not physically present, e.g. when paying for an order made on the Internet. In a real point-of-sale the customer connects to a merchant's device with Bluetooth, a short-range general purpose wireless network technology. The device could be a payment terminal or a vending machine. A notable feature of the payment system is that the merchant's device does not need to have any connection besides the one to the customer's mobile phone. The proof of validity for the client's public key along with the proof of a completed payment can all be forwarded to the merchant's device by the customer's mobile phone. All that is needed by the merchant's device to verify those proofs are the public keys of the banks that use the mobile payment system and the public key of the Population Register Centre, and those can be pre-loaded into the machine when it is first set up. The transaction for a real point-of-sale consists of eight phases:

1. The customer initiates a connection with the merchant's machine.
2. The merchant's machine offers a set of choices that are available from it to the customer.
3. The customer sends its choice to the merchant's machine.
4. The merchant's machine sends a request for payment to the customer's mobile phone.
5. The customer sends a payment order to their bank.
6. The bank sends a payment receipt to the customer.
7. The customer sends the receipt to the merchant's machine.
8. The merchant's machine delivers the product to the customer.

Each message is signed and encrypted by the sender with the public key of the receiver, and after the message sequence, each party has sufficient proof about the transaction that any disputes that may arise about the transaction can be solved.

Figure 2 illustrates the session.

The scenario of a virtual point-of-sale means that the customer is connected to the merchant's machine remotely, meaning that the merchant's machine obviously has a remote connection. Because of this, the message exchange between the customer and the merchant can be a bit simpler. Only six steps are required in total, four of which are between the customer and the merchant:

1. The customer initiates a connection to the merchant's machine.

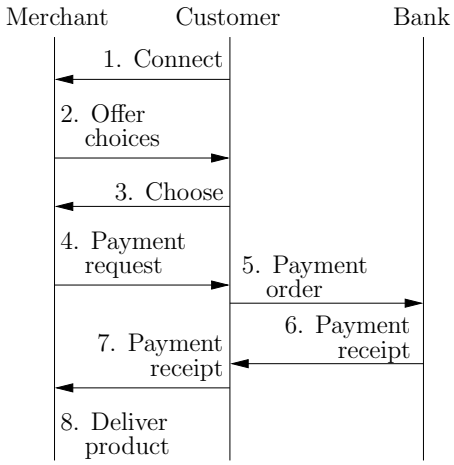

Figure 2: The transaction steps of a real point-of-sale purchase.

2. The merchant's machine offers a set of choices available from it to the customer.
3. The customer sends its choice to the merchant's machine, including a signed authorisation of payment.
4. The merchant relays the payment authorisation to the customer's bank.
5. The bank confirms the payment to the merchant.
6. The merchant delivers the product and a receipt for the transaction to the customer.

Again, each message is signed and encrypted by the sender with the public key of receiver, and in this case the payment authorisation is encrypted with the bank's public key, so the merchant cannot decipher it. Each party has sufficient proof of the transaction in this case as well, except for the situation where the merchant does not deliver the product or the receipt for the customer. The bank has proof of the transaction, however, and the customer can use that in case of a dispute. The bank gets information about the customer and the merchant, which is necessary to make the payment, but other details about the transaction are not given to the bank. The merchant gets information about the customer and the customer's bank, which is the same information that it gets from a debt card payment.

7 Comparison

The two solutions for mobile authentication are evaluated here and compared to each other. The evaluation criteria are the requirements set for the user of the mobile phone and the requirements set for other parties of the authentication for the authentication to work properly. The possibility to use existing solutions are also considered, and the security features provided by each solution are evaluated.

In [2], the requirements for the user are to have a SIM card with an operator-issued pair of public and private keys and the identity provider software installed on the device. These requirements are fairly light, but additionally the user must remember the address through which their identity provider

can be contacted, which can be a bit of a burden, especially if the user has to use many other authentications with other services as well, e.g. password-based ones. Still, it is not a requirement that would seriously hinder the adoption of this authentication method, since it does remove the need to remember a password for the service.

The requirements for the service provider are the ability to verify signatures made by private keys that are issued by a mobile operator, and the ability to use the single sign-on procedure of the Liberty Alliance with an arbitrary address for the identity provider component. While they both require custom software to be implemented, tested, and deployed, much of the work could be handed to already existing libraries that can be used to perform the tasks. Therefore the requirements set for the service provider are not overly big either.

The requirement for a third party, e.g. the mobile operator, to have a computer that acts as an HTTP server and conveys the authentication messages to and from the user's mobile phone is a bit troublesome, however. It requires both hardware resources and custom software development, testing, and deployment from the third party, who is likely to want some compensation for it. This makes the solution somewhat unattractive compared to the simple username and password authentication, which is completely free for both the service provider and the user. The method does offer some security benefits over the traditional password authentication, namely making phishing attacks more difficult, but it is hard to see this as a benefit that would outweigh the fact that not insignificant resources are required from a third party.

The requirements for the user in [3] are again the authentication software installed on their mobile phone and a SIM card with a public and private key pair, granted by the Finnish Population Register Centre in this case. This time there is no need for the user to remember any additional addresses or other non-obvious details, however.

The requirements for the merchant are the ability to verify signatures made by private keys that are granted by the Finnish Population Register Centre and the ability to communicate with the described protocol. The signature verification would probably not be a problem in this case either, but the software to handle the protocol would probably have to be built at the expense of the merchant. Moreover, since the protocol is not yet standardised in any way, only abstractly described in [3], there is currently no possibility to build such software. Still, if the protocol is ever standardised, it is likely that different implementations, even free ones, become available.

The requirements for the bank are the ability to verify signatures made by private keys that are granted by the Finnish Population Register Centre and the ability to communicate with the described protocol, making money transfers as required by the communication. Should the protocol become an accepted form of payment, however, most banks would probably implement it like they currently do with Internet banking software, and cover the costs from their normal revenue. Therefore the requirements for a third party in this case could be seen as less severe than in the previous case. The security features in the protocol are quite refined: every authentication phase is verified by a digital signature, only the necessary pieces of information are given to each party, and each party has some form of proof that the transaction took place. The only real problem with the authentication method is that it is not yet standardised. Still, it appears to be a very promising system with a clearly defined potential usage.

8 Conclusions

The current state of mobile authentication and identity management methods seems to be that some promising systems based on digital signatures are being developed, but none are quite ready for actual deployment yet. Techniques like RADIUS and Diameter are already deployed, but neither has received very much focus, apparently because the need they fulfil is not universal: they are not suitable for any authentication situation. In the view of the author, a successful authentication method should place as little burden on the parties of the authentication, but some compromises are always necessary, e.g. establishing trust via a third party.

References

[1] The Liberty Alliance Project. `http://www.projectliberty.org`.

[2] T. Abe, H. Itoh, and K. Takahashi. Implementing identity provider on mobile phone. In A. Goto, editor, *Digital Identity Management*, pages 46–52. ACM, 2007.

[3] M. Hassinen, K. Hyppönen, and K. Haataja. An Open, PKI-Based Mobile Payment System. In G. Müller, editor, *Emerging Trends in Information and Communication Security, International Conference, ETRICS 2006, Freiburg, Germany, June 6-9, 2006, Proceedings*, volume 3995 of *Lecture Notes in Computer Science*, pages 86–100. Springer, 2006.

[4] A. Niemi, J. Arkko, and V. Torvinen. Hypertext Transfer Protocol (HTTP) Digest Authentication Using Authentication and Key Agreement (AKA). RFC 3310, The Internet Engineering Task Force, September 2002. `http://ietf.org/rfc/rfc3310.txt`.

[5] C. Rigney, A. C. Rubens, W. A. Simpson, and S. Willens. Remote Authentication Dial In User Service (RADIUS). RFC 2865, The Internet Engineering Task Force, June 2000. `http://ietf.org/rfc/rfc2865.txt`.

[6] C. Rigney, A. C. Rubens, W. A. Simpson, and S. Willens. Diameter Base Protocol. RFC 3588, The Internet Engineering Task Force, September 2003. `http://ietf.org/rfc/rfc3588.txt`.

[7] A. Ruiz-Martínez, D. Sánchez-Martínez, M. Martínez-Montesinos, and A. F. Gómez-Skarmeta. A survey of electronic signature solutions in mobile devices. *JTAER*, 2(3):94–109, 2007.