Energy-Efficient Mutual Authentication Protocol for Handhled Devices Based on Public Key Cryptography

Wilayat Khan, Habib Ullah, and Riaz Hussain

Abstract—**Mobility is one of the major features of wireless communication systems and handheld devices form a major part of these systems. The recent growth in the functionality and features of the handheld devices has increased their use and importance in different applications around the world. Unfortunately, the security threats and their intensities are also increasing with the passage of time. The limited resources like battery, memory, and computational power of these devices is a bottle neck in the security of such devices. Meanwhile, the focus of recent research is to propose better security solutions with limited resource usage. Authentication of handheld devices and servers with each other is one such security issue has not yet been solved satisfactorily. Many authentication protocols based on symmetric or asymmetric cryptography has been proposed. The former one is more feasible in terms of energy and speed but not as secure and reliable as the last one. The last one is more reliable but on the other hand it consumes more energy. In this paper, we propose enhancements to the existing asymmetric authentication protocol based on FIPS 196 standards. With these enhancements, the energy consumption in handheld devices is reduced during authentication process while the reliability and security level, inherent in the asymmetric cryptographic algorithms, remains the same.**

Index Terms—**Authentication, handheld devices, FIPS 196, security, public-key cryptography, asymmetric protocols.**

I. Introduction

The advancements in recent mobile development technology have increased the usage of handheld devices in serious business and financial applications. Furthermore, their features; smaller size, mobility, flexibility, ubiquity and convenience, make handheld devices and technology to become an indispensable part of every day's life. They are being used in medical care centers, real-time telecommunication fields and m-commerce applications. Current handheld devices are capable to perform security functions just like mini computers. Handheld applications improve overall benefits of the companies through increasing efficiency, effectiveness and customer satisfaction [1], [2], [3].

Like normal wired networks and applications, handheld devices are also facing serious security threats. Some of these security threats and risks for handheld devices connected to the Internet are mentioned in [4]. Due to its open nature, mobile communication is more prone to attacks like data interception, unauthorized access, and unauthenticated connectivity than wired communication networks. In normal networks, two-party communication system needs to authenticate each other before accessing different resources, services, and sharing secret information with each other [5], [6]. Many new security challenges arise due to the unique characteristics of the battery-powered handheld devices [7]. For battery-limited devices, perhaps, one of the foremost challenges is the mismatch between the energy and performance requirements of execution of security functions and protocols. These security functions and protocols are: symmetric key cryptography, asymmetric key cryptography, authentication protocols, secure session, and so on [8].

Currently, most of the handheld devices are capable to implement asymmetric key based authentication and secure session which provides reliable and secure authentication. On the other hand, handling of public key certificates, computational extensive cryptographic calculations and processing of security tickets to access services are the bottle neck in the performance of current handheld devices which have limited memory, processing power and battery.

In this paper we designed an authentication protocol for handheld devices which is based on the existing global security protocols, technologies and servers. This protocol reduces energy consumption in handheld devices during authentication process.

The rest of the paper is organized as follows: in the next section, we present an overview of the authentication protocols in handheld devices. The actual design of our proposed protocol is mentioned in the Section III. In Section IV, we analyze our protocol. In the last Section V, the conclusion will be presented.

II. Related Work

With the evolution of mobile communications, the network and communication security has gained significant attention of the research community. Many authentication protocols have been proposed for the handheld devices.

Simoes *et al.* [9] proposed an authentication protocol based on symmetric cryptography. It was designed to target the problem of high computational power of asymmetric cryptographic algorithms used in security protocols for resource-constrained handheld devices. This protocol needs an internal key created by the server and is sent to the device

Manuscript received November 28, 2012; revised February 26, 2013. This work was supported by Higher Education Commission (HEC), Pakistan.

Wilayat Khan was with the Department of Electrical Engineering, Comsats, Pakistan. He is now with Department of Computer Science, Ca Foscari University of Venice, Italy (e-mail: khan@ dsi.unive.it).

Habib Ullah was with the Department of Electrical Engineering, Comsats, Pakistan. He is now with Department of Computer Science, University of Trento, Italy (e-mail: habib@ciit-wah.edu.pk).

Riaz Hussain is with the Department of Electrical Engineering, Comsats, Pakistan. (e-mail: riaz.junejo@gmail.com).

through a client's PC as the proxy. The reliability of this protocol is based on two assumptions that the internal key travels from the server through a *secure channel* using the client's PC as a proxy and that the channel between the proxy PC and the device is *secure*.

Another protocol for mobile network authentication and security was designed by Yi et al. [10]. This protocol is based on ElGamal signature scheme and Diffie-Helman key distribution protocol. The protocol is used to authenticate a mobile user with a base station and vise versa in mobile networks. The protocol requires the existence of a Certification Authority (CA) in the mobile networks to create public key certificates for mobile users and base stations. This protocol is very useful as the resource restricted handheld devices do not generate random numbers; instead they are generated by the base station. Also only the public parameters; a large prim p and the public key of the CA (y_{ca}) are required to be known by all of the network participants.

Although, this protocol allows efficient computation and less storage requirement in resource-limited handheld devices, but, Martin et al., Wong, and Laih et al. [11]-[13] showed that this protocol is vulnerable to some attacks. Latter, Lee et al. proposed an improved version of this protocol in [14] to remedy these attacks. The authors have claimed that the session key used is not possible to be obtained by attacker because it is changed constantly and is encrypted with the secret key of the mobile user and the base station. Furthermore, this protocol is also able to prevent an attacker from replaying or forging certificate because time-stamps are used.

Another well known protocol is Federal Information Processing Standard (FIPS) 196 protocol [6] where the entity authentication is based on the public key cryptography. Being an asymmetric cryptographic protocol, it provides the full security features like authentication, digital signatures, non-repudiation and secure key distribution.

A computationally efficient protocol for handheld devices based on PKI, PKASSO was presented in [15]. PKASSO include the Single Sign-On (SSO) and delegation technologies which is used especially for handheld devices with limited resources. The SSO capability relieves the poor device from repeating the sign-on procedure. This infrastructure mainly is based on Kerberos, CA and LDAP protocol. The main aim of PKASSO was to provide non-obstructive authentication latency in a handheld device with restricted computing power to support digital signature and non-repudiation. The two main components, a delegation and a referee server, are responsible to perform extensive PKI operations on behalf of handheld device and to provide computationally efficient digital signature and non-repudiation.

Symmetric key cryptography, despite of its fast processing, is not liked due to its low reliable nature and the key distribution issue. In the protocol proposed by Simoes *et al.* [9], the key distribution challenge has not been addressed appropriately. Also both of the parties, the device and the server, have to share some shared secrete; the internal key. The main benefit of the Yi *et al.* protocol is its computational efficiency for the resource-limited devices. One of the reasons stated for this efficiency is the location of the random number generator function at the base station. But in the Lee et al. [14] protocol, the mobile user is responsible to generate random number. So one of the feature mentioned in the Yi *et al.* protocol is conflicted in the Lee *et al.* protocol. Also both of the protocols proposed by Yi *et al.* and Lee *et al.* uses time-stamps to prevent an attacker from replaying or forging certificate. In a widely distributed environment where two entities do not necessarily know each other prior authentication, they need to maintain synchronized clocks [6]. In the world wide applications like e-commerce this often happens that the two entities willing to communicate do not know each other before authentication.

The FIPS 196 based authentication protocol provides the main security features like digital signatures and non-repudiation but the computationally extensive cryptographic calculations during digital signature creation/verification and certificate validation are its big drawbacks for the handheld devices with limited computational power, memory and battery life. PKASSO could be a very feasible option in terms of efficiency and reliability as it provides digital signature and non-repudiation through using SSO and delegation but it has a limitation as it works only in a local environment. Kerberos provides a foundation for PKASSO which can provide authentication facility in a limited area. Our proposed protocol is mainly designed to work in a global environment with the same benefit of efficient authentication with digital signature and non-repudiation. The proposed protocol and the underlying design principles are discussed in the next section.

III. The Proposed Protocol

The main aim of the proposed protocol is to achieve three objectives. First, the handheld device will be able to authenticate the server and vice versa effectively with minimum usage of resources. Second, the infrastructure include the main security features digital signatures and non-repudiation. The third objective is the system will be able to work in a global environment.

As our proposed protocol is based on FIPS 196 standard, therefore, we first explain the main steps of FIPS 196 mutual authentication protocol. This protocol consists of two main parties, the initiator 'A' and the responder 'B'. The following messages exchange between the initiator and the responder.

1) In the first message MessageAB1, the initiator sends an authentication request message to the responder.
2) The responder, if willing to proceed sends back a message MessageBA1 which includes a random challenge R_B.
3) After receiving the message in step 2, the initiator creates its own random challenge R_A, combines it with R_B and creates a digital signature by signing the combination of both R_A and R_B. The initiator sends the signature, its public key certificate (or public key) and some other optional data in MessageAB2.
4) The responder, after receiving the MessageAB2, first verifies the initiator's certificate or the public key, then verifies the digital signature using the initiator's public key. If the verification succeeds, the initiator is

authenticated and the responder creates signature of both R_A and R_B. It sends the signature, its certificate (or public key), and some other optional data towards the initiator in MessageBA2.

5) Similar to the responder, the initiator also first verifies the public key certificate or the public key of the responder and then verifies its digital signature using responder's public key. If both of the verifications succeed, the responder is authenticated.

In FIPS 196 based authentication protocol, the responder and the initiator both may terminate the authentication exchange for at least one of the reasons: failure of the verification of the entity's binding with its public/private key pair and failure of the verification of its digital signature on a random number challenge.

The termination decision is taken after they receive the third and the fourth messages respectively of the authentication exchange. After the message MessageAB2 is received (step 4), the responder verifies the binding of the initiator with the public/private key pair and verifies the initiator's signature. If any of these verifications fail, the authentication exchange is terminated by the responder. Similarly, the initiator verifies the binding of the responder with its public/private key pair and its digital signature after it receives the last message MessageBA2 (step 5). The initiator terminates authentication process if any of the verifications fail.

In any case, the initiator and the responder exchange at least three messages prior authentication termination. In the case when the authentication is terminated due to the failure of signature verification, the time/energy spent on these messages (three for the responder and four for the initiator) exchanged cannot be avoided because they must share these messages for the signatures to arrive at the responder and the initiator. If the authentication termination is due to the failure of the verification of the entity's binding with the public/private key pair, even then they have to share at least three messages because the certificates are exchanged in the third and fourth messages. Unfortunately, in the situation where at least one of the entities is fraudulent, its binding with the public/private key is verified after they share at least three messages. If the server is fraudulent, the poor device spends time and energy to send/receive four messages, create digital signature, and so on. The situation becomes worse if the frequency of the fraudulent authentication efforts is high. In this paper, the timing of the entity's binding verification with the public/private key pair is modified to reduce the number of messages exchanged and the security processing. These improvements are shown in the proposed protocol by bringing some enhancements in the FIPS 196 based mutual authentication protocol.

There are two possible enhancements. The first enhancement is brought by modifying the contents of the first and the third authentication messages MessageAB1 (authentication request) and MessageAB2. The second possible enhancement is made by modifying the contents of the second and the fourth authentication messages MessageBA1 and MessageBA2. These enhancements in FIPS 196 protocol as used in our proposed protocol for authentication between client and server are discussed below.

1) The verification of the client's binding with its public/private key at the server is done just after receiving the first message. For this purpose, the public key certificate of the client and its user ID is included in the authentication request message MessageAB1. When the server receives this message, it verifies the certificate of the client. In case certificates are not used for signature verification, the identifier for the initiator is used to retrieve the public key of the client.

   The server decides either to terminate or continue the authentication process based on the result of certificate verification. The server takes a decision exactly on time to either terminate or continue authentication. If the certificate of the client is not verified, unlike the standard FIPS 196, the server does not send the second and the fourth message (MessageBA1 and MessageBA2) of the authentication exchange. Also, the client neither creates digital signature nor it sends the third message MessageAB2 of the authentication exchange.

2) The server's certificate verification at the client can be performed earlier. In FIPS 196, this verification is done after the client (initiator) receives the fourth message MessageBA2 which includes public key certificate of the responder Cert B. If the certificates are not used, then the server's identifier is used to retrieve its public key. Unlike FIPS 196, the certificate of the server Cert B is included in the second authentication exchange message MessageBA1. The client verifies the certificate after it receives MessageBA1. Based on the result of verification, it decides whether to create digital signature to authenticate itself with the server or not.

The sequence of the four messages exchanged between the client and the server while authenticating each other are shown in the Fig. 1. The Token ID is the identifier of the token TokenBA1, Token BA2 or Token AB. These tokens contain random number challenges, digital signatures, and/or some optional data as defined in the FIPS 196 standard. The certificates (Cert A and Cert B) of the client and the server are optional. The corresponding public key can also be used or the user ID is used to get the user's public key certificate from the certificate directory like X.500 directory server.

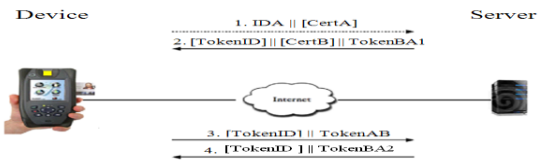

Fig. 1. Mutual authentication between client and server

IV. Analysis of the Protocol

As was mentioned in the previous section, this protocol is efficient because it reduces the security processing and data transmission. In communication networks, a packet experiences a number of delays while moving from source to destination [16].

The source has to break the data down into packets to be transmitted. It brings some delay at the origin of the message

(e.g. handheld device). This is *packetization* delay (dpack) which depends on the size of the packet. The packet travels through many nodes (e.g. routers) in the network to reach its destination.

At each node, the header of the packet is examined to determine where to route it. This is *processing* delay (dproc). The packet has to wait in a queue to be transmitted onto the link which is *queuing* delay (dqueue). Besides that there is *transmission* delay (dtrans) required sending bits into link and *propagation* delay (dprop) required for bits to travel from router to router. The sum of all these delays is *authentication transmission* delay (dauthtrans).

According to a survey for handheld Symbol PPT2800 Pocket PC [17], for a secure wireless session where 64 kB of data is transmitted, which involves 3DES for bulk data encryption and SHA for message authentication, nearly 21% and 44% of the overall energy is consumed for security processing and data transmission respectively. The results for different size of data and algorithms will be different (higher for large data sizes and asymmetric algorithms).

As is mentioned in FIPS 196, the authentication of an entity depends on two successful actions: the verification of the entity's binding with its public/private key and the verification of its digital signature on a random number challenge.

In the proposed design, we have changed the timings of these verifications. The device and the server takes a termination decision well on time just after each receive the first message (server takes a decision after it receive first message and device after it receive the second message of the authentication process), unlike the FIPS 196 based protocol where the server waits until it receive the third message and the device until it receive the last message.

Ravi *et al*. [8] calculated energy consumption of the digital signature algorithms. According to their calculations, RSA algorithm with 1024 bits key size consumes 270.13mJ, 546.5mJ, and 15.97mJ energy respectively for key generation, signing, and verification.

In a scenario when a server and/or client attempt to authenticate each other by a fraudulent way, the device and the server do not wait until the end to terminate authentication. In case the server is fraudulent, the resource-limited device does not send the third message (saves transmission time and energy). It does not need to create digital signature to be sent in the third message. As authentication is terminated, so it does not need to verify the digital signature of the fraudulent server (saves security processing time and energy). If we only consider the signature creation and verification, the energy saved is approximately 562.47mJ. Similarly, if the client is fraudulent, the server, after it receives the first message, will not verify client's binding with its public/private key pair. It neither sends any message to the device nor does it create/verify any digital signature. In both cases, it saves security processing time and energy.

As a result even if we assume the *priori generation* of the parameters used in the key generation process, still energy and time required for security processing and transmission can be saved. The energy and time saved are of great importance for the busy servers and battery-limited handheld devices.

V. Conclusions

The importance and use of the handheld devices is rapidly increasing in the wireless applications like e-commerce. Unfortunately, the security threats and challenges are also increasing due to the wireless nature and limited resources like battery, memory, and processing power of the handheld devices. One of these security challenges is the authentication of two entities where one of the entities is a resource-limited handheld device. The researchers have been trying for many years to design efficient and secure authentication protocol for the handheld devices. The symmetric algorithms provide efficient alternatives to be used in the authentication protocol but the preference is given to use established and higher assurance algorithms instead of faster one with security risks unknown because efficient security protocols are of limited use if they compromise on the security.

The proposed protocol in this paper is an asymmetric-based authentication protocol which inherits the reliability and security features of the asymmetric algorithms. In this protocol, the client and the server takes an early decision to either terminate or continue with the authentication process when any of the parties is trying to authenticate itself in a fraudulent way. This reduces the energy and time consumption during the authentication process which is very important for the battery-powered handheld devices. The protocol is applicable for both PC-based cleint/server application and when one of the parties is a resource-limited handheld device.

Although, the energy and time saved has not been calculated experimentally for different key sizes, algorithms, and frequency and types of attacks, but the approximate mathematical analysis shows that the design is more efficient and reliable to be used for client/server applications especially when one of the parties is with limited power. Researchers are invited to measure the time and energy saved during transmitting authentication data, signature creation and verification using different algorithms and key sizes, and in the presence of different types of attacks with different frequency.

References

[1] K. A. Banitsas, P. Georgiadis, S. Tachakra, and D. Cavouras, "Using handheld devices for real-time wireless telecommunication," in *Proceedings of the 26th Annual International Conference of the IEEE*, pp. 3105-3108, 2004.

[2] O. O. Obe and V. F. Balogun, "Practice, trends and challenges of mobile commerce in nigeria," *Information Technology Journal*, vol. 6, no. 3, pp. 448-456, 2007.

[3] A. Herzberg, "Payments and banking with mobile personal devices," *Communication of the ACM*, vol. 46, pp. 53-58, 2003.

[4] W. Susilo, "Securing handheld devices," in *Proc. 10th IEEE International Conference*, pp. 349-354, 2002.

[5] M. S. Hwang, C. C. Lee, S. K. Chong, and J. W. Lo, "A key management of for wireless communications," *International Journal of Innovative Computing Information and Control*, vol. 4, issue 8, pp. 2045-2056, 2008.

[6] *Entity Authentication Using Public Key Cryptography*, FIPS PUB 196, U.S. Department of Commerce/National Institute of Standards and Technology, 1997.

[7] S. Ravi and A. Raghunathan, "Security in embedded systems: Design challenges," *ACM Transactions on Embedded Computing Systems*, vol. 3, no. 3, pp. 461-491, 2004.

[7] S. Ravi, A. Raghunathan, and N. K. Jha, "Analyzing the energy consumption of security protocols," in *Proceedings of the International Symposium on Low Power Electronics & Design*, pp. 30-35, 2003.

[9] P. Simões, P. Alves, J. Rogado, and P. Ferreira, "An Authentication Protocol for Mobile Devices," *International Workshop on Internet 2000 (integrated in the Internatioanl Conference on Distributed Computing Systems-ICDCS'2000)*, Taiwan, April 10-13, 2000.
[10] X. Yi, E. Okamoto, and K. Y. Lam, "An optimized protocol for mobile network authentication and security," *ACM Mobile Computing and Communications Review*, vol. 2, no. 3, pp. 37-39, 1998.
[11] K. M. Martin and C. J. Mitchell, "Comments on an optimized protocol for mobile network authentication and security," *ACM Mobile Computing and Communications Review*, vol. 3, no. 2, pp. 37, 1999.
[12] D. S. Wong, "An optimized authentication protocol for mobile network reconsidered," *ACM Mobile Computing and Communications Review*, vol. 6, no. 4, pp. 74-76, 2002.
[13] C. S. Laih and S. Y. Chiou, "Cryptanalysis of an optimized protocol for mobile network authentication and security," *Information Processing Letters*, vol. 85, no. 6, pp. 339-341, 2003.
[14] C. C. Lee, I. E. Liao, and M. S. Hwang, "An extended certificate-based authentication and security protocol for mobile networks," *Information Technology and Control*, vol. 38, no. 1, pp. 61-66, 2009.
[15] K. W. Park, S. S. Lim, and K. H. Park, "Computationally Efficient PKI-Based Single Sign-On Protocol, PKASSO for Mobile Devices," *IEEE Transactions on Computers*, vol. 57, no. 6, pp. 821-834, 2008.
[16] W. Stallings, *Data and Computer Communications*, 8th ed., Prentice Hall, 2008.
[17] R. Karri and P. Mishra, "Minimizing energy consumption of secure wireless session with QoS constraints," in *Proc. of the International Conference on Communication*, vol. 4, pp. 2053-2057, 2002.

Wilayat Khan earned his B.Sc. in computer systems engineering from NWFP University of Engineering & Technology, Peshawar, Pakistan in 2007 and master degree in information and communication systems security from the Royal Institute of Technology (KTH), Stockholm Sweden in 2009. He started his carrier in 2009 as a lecturer at the Department of Electrical Engineering, COMSATS Institute of IT, Wah Campus, Pakistan. He is currently pursuing PhD degree in the Ca Foscari University of Venice, Italy. He has published four papers on wireless and mobile networks and security in international journals and conference proceedings. His research interests include information security and using formal methods and interactive theorem provers for the security analysis of systems, in particular, web browsers and mobile systems. Mr Khan is the recepient of "Best Research Work" award from Ca Foscari University of Venice for his work on web browser security.

Habib Ullah received B.Sc. degree in computer systems engineering from NWFP University of Engineering & Technology Peshawar, Pakistan in 2006 and M.Sc. degree in electronics and computer engineering from Hanyang University, Seoul Korea. He started teaching as a lecturer in 2009 at the department of Electrical Engineering, COMSATS Institute of IT, Wah Campus, Pakistan. He is now a PhD student in the University of Trento, Italy. He has six publications focusing on background modeling and shadow detection. His research interests include information security, pattern recognition, behavior modeling and object segmentation.

Riaz Hussain got his B.Sc. in computer engineering from Quaid-e-Awam University .of Engineering, Science & Technology, Nawabshah in 2005 and master degree in telecommunication from Blekinge Institute of Technology (BTH), Sweden in 2010. He is currently serving as a lecturer since 2010 at the department of Electrical Engineering, COMSATS Institute of IT, Wah Campus, Pakistan. His research interests include mobile networks, image segmentation, image compression and signal processing.