A Generalized Protocol for Mobile Authentication in Healthcare Systems

Eric Reinsmidt
The University of Tennessee
eric@reinsmidt.com

Li Yang
The University of Tennessee at Chattanooga
li-yang@utc.edu

Abstract

The trend of handheld, mobile devices being used increasingly in the collection and transmittance of electronic healthcare records (EHR) provides a particularly sensitive area in which data must be kept private and secure. This article discusses current methods for mobile authentication in EHR schemes. Their limitations in regards to EHR are examined. These methods are then contrasted against the current landscape of threats that are emerging in the realm of mobile computing. In addition, a generalized improvement over current approaches is introduced for further study.

Introduction

The computing landscape is shifting toward mobile platforms with an increasing number of smart devices used as the preferred method of computing as compared to traditional devices. Along with this shift, there has been a proliferation of healthcare and well-being applications being developed for mobile devices. Because of the ability of these devices to potentially collect or transmit EHR to a remote server, special care must be taken to ensure that the user of a mobile application does not risk having their EHR compromised through eavesdropping or alteration.

Several agencies around the world regulate exactly how a person's EHR, or sometimes more generally any personal information, must be protected. In the United States, regulation of EHR falls under the The Health Insurance Portability and Accountability Act of 1996 (HIPAA). In Canada, except where superseded by local privacy regulation, the Personal Information Protection and Electronic Documents Act (PIPEDA) regulates collection, use and disclosure of any personal information. Across the European Union, EU Directive 95/46/EC deals with the regulation of how personal data is processed. While some of these regulations are more generalized than dealing strictly with EHR, they all share in common that they require a person's EHR data to be kept secure. These regulations coupled with the necessity of a mobile device to communicate over a wireless network create some unique challenges in ensuring the secure transfer of EHR in mobile healthcare applications.

To keep an individual's EHR secure while using a mobile device, data transmission must take place on a secure channel. In addition it is necessary to not only authenticate the person using the device. The device itself as well as any server it is communicating with must be authenticated. This article focuses on current methods of authentication in mobile networks when dealing with EHR. We also discuss a generalized method whereby a mobile device, the user of the device, and any other devices the mobile device communicates with, e.g. a remote server, can all be successfully authenticated while creating a secure channel for communication at the same time.

Background

In the past few years the landscape of computing has seen a paradigm shift in the manner of devices consumers are choosing to purchase. There has been an increasing movement to mobile devices from more traditional devices such as the desktop PC or notebook. In 2013, 968 million smartphones and 195 million tablets were sold. That represents a 46% year to year increase with 2012 sales at 796 million units of smart devices in total. (Gartner, Inc., 2014a) Of that increase in sales of smart devices, tablets saw the largest growth year to year. With sales of 116 million in 2012, tablets increased by 68% reaching 195 million units sold throughout 2013. (Gartner, Inc., 2014b)

This increase in smart device sales is contrasted by declining sales in the traditional PC market, including desk-

top PCs, notebooks, and ultramobiles. In 2013 traditional PCs had declining sales, with a decrease in units from 341 million in 2012 to 299 million units in 2013, a 12.3% decline. (Gartner, Inc., 2014c)

Since mobile devices rely upon wireless network connections as their primary means of communication, there are two important attack vectors with which a mobile device must be able to protect against. The first type of attack vector is eavesdropping. As stated by Goodrich and Tamassia (2011) eavesdropping is "the interception of information intended for someone else during its transmission over a communication channel." (p. 14). In the case of a mobile device the most likely form of eavesdropping is through packet sniffing performed by a network interface card (NIC) that is set to promiscuous mode on the same network as the mobile device. (Ansari et al. 2003) The second attack vector, alteration, would most likely take advantage of a mobile device through a man-in-the-middle (MITM) attack. Alteration is the modification of data by someone who is not authorized to do so. (Goodrich & Tamassia, 2011)

While there have been changes in users' computing preferences, there has at the same time been a push towards widespread adoption of EHR being implemented. In the US 44.4% of non-federal acute care hospitals had some sort of basic EHR system implemented in 2012 compared with 9.4% in 2008. In that same grouping of hospitals, certified EHR systems increased from 71.9% in 2011 to 85.2% in 2012. (The Office of the National Coordinator for Health Information Technology, 2013)

To deal with the rise of mobile computing and EHR, novel approaches have been devised to allow authentication of a healthcare system's users. Hsiao et al. describe a system that uses a two-factor authentication scheme with the use of a password and a smart card, both of which are used in conjunction with a mobile device. (Hsiao et al. 2012) A different form of multifactor authentication is discussed by Mirkovic et al. whereby a mobile device and its user are authenticated through the use of a personal identification number (PIN), and a short message service (SMS) message. In addition an identity server and authentication provider server are required as well as connection to the service provider. (Mirkovic et al. 2011)

A generalized protocol for mobile authentication in health care systems

Issues of existing authentication systems

The system proposed by Hsiao et al. is a robust system. The system is designed to deal with many types of attacks. It uses a one way hashing function that incorporates a timestamp, and so is protected against replay attacks as the timestamp must be recent to enter the system. It is also resistant to guessing attacks and impersonation attacks. However through its use of multiple physically accessible nodes to collect data, there is the possibility of data loss due to node theft. In addition the system requires a user to carry not only a mobile device but also an authentication token. Loss of either device leaves the user unable to access the system. Most importantly, the system is designed only for healthcare providers such as physicians or clinicians to access the system. There is no mechanism to allow a patient to access the system.

The secure solution provided by Mirkovic et al. is a very well-thought-out system. It provides multifactor authentication and can be used on any mobile platform and with any mobile provider, which is a distinct advantage. The system also provides for a high level of usability. The solution is also robust against session hijacking attacks and in general is secure due to the multifactor authentication. However the system also uses a multistep process for authentication using multiple servers and requires an SMS message to be sent. Because of this there are increased monetary and temporal costs.

The proposed protocol for mobile authentication

A generalized secure protocol is proposed below that allows authentication of all actors in a transaction of EHR between a mobile device and a remote server. A secure channel is also created during the authentication process. It does not require any additional physical hardware such as an authentication token. In addition the transaction takes place between a single remote server and a mobile device and so has a reduced cost associated with the proposed system.

The protocol uses a Diffie-Hellman (DH) key exchange for creation of an encryption key. However the DH key exchange does not provide for authentication; it is an anonymous exchange. To add authentication into the system, a Rivest Shamir Adleman (RSA) public-key cryptosystem is introduced. This allows both the mobile device and remote server to authenticate the identity of the other and at the same time create a private symmetric session key that can be used for the encryption and decryption of data that is transmitted. On every connection between a mobile device and remote server a unique session key is created, preventing replay attacks. In addition impersonation attacks are not possible unless either the mobile device or the remote server has had their private RSA key compromised.

The protocol has multiple steps involved in creating this secure channel while also authenticating the mobile device and the remote server:

(1) After a mobile device opens a socket with the listening server, the server responds with its public portion of the DH exchange, $g^y \bmod p$, which is calculated with its secret y.

(2) The mobile device takes this value, raises it to the power of its secret x, and uses the resulting value to calculate the key K, a SHA-256 hash. K is then used as a symmetric encryption/decryption key using the advanced encryption standard (AES) employing any of the modes stronger than electronic codebook (ECB) as it is susceptible to side channel attacks. The mode chosen may depend on the data being sent. For example, if the mobile device is being used to send streaming sensor data from the device as part of a diagnostic tool, then a mode such as counter (CTR), output feedback (OFB), or cipher feedback (CFB) may be desirable. The mobile device then signs the server's public DH value with its private key after which it encrypts the resulting value with K. This is then sent to the remote server along with the mobile device's public portion of the DH exchange, $g^x \bmod p$, which has been calculated with the mobile device's secret, x.

(3) After the server receives this information from the mobile device it can calculate the symmetric encryption/decryption key K by hashing the value resulting from raising the mobile device's public portion of the DH exchange to it secret y with SHA-256. After the server has calculated K it can use the key to decrypt what was sent from the mobile device. The decrypted value is the server's public portion of the DH exchange signed with the private RSA key of the mobile device. By using the public RSA key of the mobile device to decrypt this, the server should be able to then compare its public portion of the DH exchange with the resulting value. If the two values match, the server knows two things: the encryption/decryption key K which can be used for secure channel communication, and it knows that the mobile device is indeed who it says it is and so is authenticated. If however the values do not match, it can then be assumed that an impostor was contacting the server and hence the connection is dropped.

Assuming the values did however match, the server signs the mobile device's public portion of the DH exchange with the server's private RSA key, and encrypts that data with the symmetric encryption/decryption key K and sends this to the mobile device.

(4) The mobile device uses encryption/decryption key K to decrypt the data received, after which it uses the server's public RSA key to decrypt the result from the previous step. If the final result of this matches the mobile device's public portion of the DH exchange, then the mobile device now knows that the server is who it says it is and hence has been authenticated. If the two values do not match, then the mobile device knows there is a malicious user on the other end impersonating the remote server and so closes the socket, terminating communication.

If everything did match though, both devices have been authenticated. In addition there is a one-time symmetric encryption/decryption key K, which can be used only during this session. This provides a means for the two authenticated devices to communicate on an encrypted, secure channel. It is important to note that both the server and the mobile device have to have knowledge of the other's public key. As with any public/private key system, if the private key of either has been compromised, impersonation can occur by an attacker. Once the mobile device and cloud server have been authenticated and a secure channel has been setup for communication, a human user can enter their password. Figure 1 shows a sequence diagram of the secure channel creation as well as the authentication of both the mobile device and remote server.

Only after the mobile device and remote server are authenticated can the user authenticate them self. Instead of a typical text based password a visual password can be used. Visual passwords are advantageous in that they can be more easily retained mentally compared with a text-based password and PIN. (Duncan et al. 2004) A study done at Carleton University showed that a visual password system had login accuracy rates of 96%. (Chiasson et al. 2007)

There are many types of visual password systems. An image can simply be mapped to a numerical value and the resulting array of numerical values can be compared against a stored password array. A visual password system can consist of a small number of images, but by allowing password elements to consist of multiple images a large alphabet size can be achieved. Using this technique Jansen et al. were able to create an alphabet size of 930 from a field of 30 images. Using this method an eight entry password would have the equivalence of a 12 character text password using the 95 ASCII printable characters. (Jansen et al. 2003) Another method of a visual password system was described by Chiasson et al. whereby a single image was used and the user would have to click on different positions within the image. Depending on the resolution of the image and the screen size of the device this could lead to a very large alphabet from which to create a password. However visual passwords and similarly textual passwords can be susceptible to shoulder-surfing attacks.

Our protocol would use a different type of visual password system. Wiedenbeck et al. proposed and implemented a visual password system that is resistant to shoulder-surfing attacks. (Wiedenbeck et al. 2006) The password system works by presenting users with multiple challenge-response rounds. A user is shown a screen with multiple images displayed. Of those displayed images only a small

number, n, are actual elements of the user's password. The user must click within the n-gon created by the user's password elements. By having several of these rounds in a row the system is using zero-knowledge authentication by the user never actually giving away the secret that they know. While this does require a longer login time the added security against a should-surfing attack makes it worthwhile.

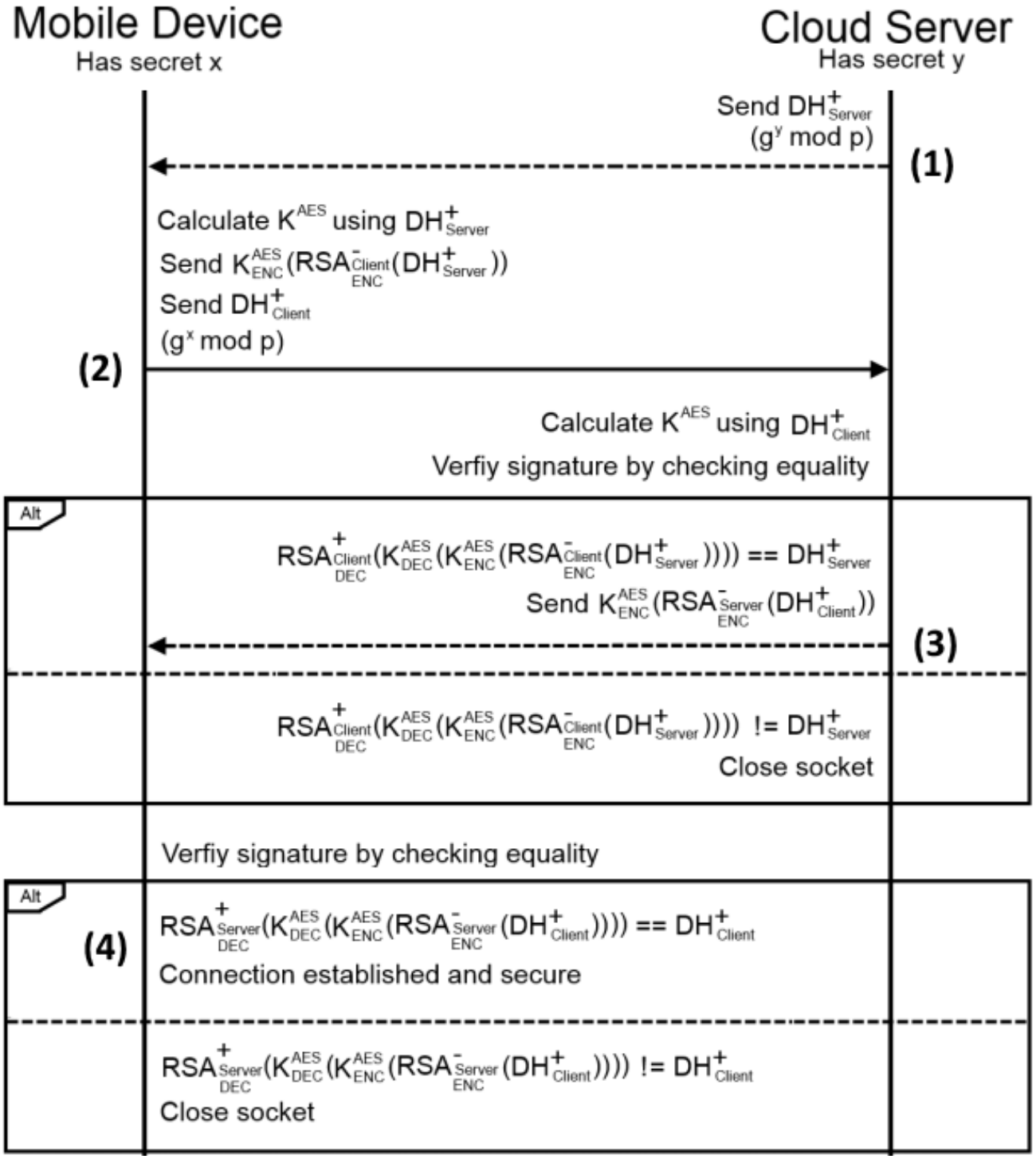

Figure 1. Secure Channel and Authentication Protocol Sequence Diagram

Future Research Directions

There will be continued growth in the mobile healthcare field. As this field continues to expand research opportunities will continue. Smart devices continue to increase in popularity, and hence there is a wide swath of demographics that use these smart devices. Because of this, usability will continue to be a key component of mobile healthcare applications. Further research may include an actual implementation of this proposed system as well as a study of its usability and performance. In addition there is growing interest in using mobile devices as sensors. Smart devices are able to capture a wealth of information and so data collection is another area of growth within mobile healthcare.

Conclusion

Because of regulation such as HIPPA, PIPEDA, and the EU Directive 95/46/EC mobile applications that deal with EHR must ensure that a user's data is kept secure from endpoint to endpoint as well as during storage. In regards to healthcare systems, authentication of mobile devices and users is of great importance. In this article, we have examined some existing methodologies of authentication on mobile devices in regards to EHR, and how those methodologies can improved upon. In addition, a generalized solution for creating a secure channel for a mobile device to send EHR to a remote server has been introduced that is robust against many types of attacks.

References

Ansari, S., Rajeev, S., & Chandrashekar, H. (2003, January 22). Packet sniffing: a brief introduction. IEEE Potentials, pp. 17-19.

Chiasson, S., van Oorschot, P. C., & Biddle, R. (2007). Graphical Password Authentication Using Cued Click Points. 12th European Symposium On Research In Computer Security, (pp. 359-374). Dresden.

Duncan, M. V., Akhtari, M. S., & Bradford, P. G. (2004, May). Visual Security for Wireless Handheld Devices. unpublished.

Gartner, Inc. (2014a, February 13). Gartner Says Annual Smartphone Sales Surpassed Sales of Feature Phones for the First Time in 2013. Retrieved October 20, 2014, from http://www.gartner.com/newsroom/id/2665715

Gartner, Inc. (2014b, March 3). Gartner Says Worldwide Tablet Sales Grew 68 Percent in 2013, With Android Capturing 62 Percent of the Market. Retrieved October 20, 2014, from http://www.gartner.com/newsroom/id/2674215

Gartner, Inc. (2014c, January 7). Gartner Says Worldwide Traditional PC, Tablet, Ultramobile and Mobile Phone Shipments On Pace to Grow 7.6 Percent in 2014. Retrieved October 20, 2014, from http://www.gartner.com/newsroom/id/2645115

Goodrich, M. T., & Tamassia, R. (2011). Introduction to Computer Security. Boston: Pearson Education, Inc.

Hsiao, T.-C., Liao, Y.-T., Huang, J.-Y., Chen, T.-S., & Horng, G.-B. (2012). Secure Authentication Scheme for Supporting Healthcare in Wireless Sensor Networks. 2012 26th International Conference on Advanced Information Networking and Applications Workshops (WAINA), (pp. 502-507). Fukuoka.

Jansen, W., Gavrila, S., Korolev, V., Ayers, R., & Swanstrom, R. (2003). Picture Password: A Visual Login Technique for Mobile Devices. National Institute of Standards and Technology Interagency Report, National Institute of Standards and Technology, Gaithersburg,.

Mirkovic, J., Bryhni, H., & Ruland, C. (2011). Secure solution for mobile access to patient's health care record. 2011 13th IEEE International Conference on e-Health Networking Applications and Services (Healthcom), (pp. 296-303). Columbia.

The Office of the National Coordinator for Health Information Technology. (2013, March). Adoption of Electronic Health Record Systems among U.S. Non-federal Acute Care Hospitals: 2008-2012 . Retrieved October 4, 2013, from http://www.healthit.gov/sites/default/files/oncdatabrief9final.pdf

Wiedenbeck, S., Waters, J., Sobrado, L., & Birget, J.-C. (2006). Design and Evaluation of a Shoulder-Surfing Resistant Graphical Password Scheme. 2006 Proceedings of the working conference on Advanced visual interfaces (AVI), (pp. 177-184). New York City.