Secure Real World Interaction Using Mobile Devices

William R. Claycomb and Dongwan Shin

Department of Computer Science
New Mexico Institute of Mining and Technology
Socorro, NM 87801, USA
{billc,doshin}@nmt.edu

Abstract. As the capabilities of mobile devices grow, so do the potential uses for real-world interactions. Using commonly carried devices, such as cell phones and PDAs, to assist in routine tasks, such as using vending machines or ATMs, is an emerging area of ubiquitous computing. However, with the increase in potential uses, the potential for misuse and attacks also increase. In this paper, we present a secure method of using mobile devices to interact with real world objects. Our model uses two dimensional colorized barcodes to easily establish a secure link between mobile devices and service points. We discuss the model, its security, and potential uses for common tasks. We also discuss potential security concerns, and cover potential cases where this model could be used to both automate and control commonly performed tasks.

1 Introduction

Mobile devices, such as PDAs and cell phones, are becoming increasingly more capable of handling tasks such as interacting with other devices in ubiquitous and pervasive computing environments. Devices often have the ability to communicate using network protocols such as Bluetooth or 802.11x, and many now include additional multimedia features such as integrated cameras. The use of the combination of these features to establish communication between two or more mobile devices, or between a mobile device and some point of service, is an emerging and interesting research area in ubiquitous and pervasive computing environments.

The use of mobile devices to interact with real world objects will not only make certain transactions easier and more convenient, it may also open a new door to a realm of new conveniences. For instance, using a camera-enabled cell phone to establish a secure link with an ATM, and subsequently retrieve cash from the user's bank account, would present several advantages. Not only would this make the process easier from the user's perspective, it would also eliminate various types of attacks known to occur against ATMs. Using a mobile device equipped with an electronic wallet, or digital cash [9], would allow the user to make purchases from automated devices, such as vending machines, and do so without the use of actual cash, or even credit card information. Additionally,

using the functionality available to these devices, controls could be implemented to limit the type of, size of, or specific items related to transactions attempted using the mobile device.

Recently proposed solutions explore the tasks of discovering nearby connection possibilities, establishing a connection with a particular device, and using the mobile device to interact once a connection has been established [8, 16–19, 21]. One aspect that has not been explored in very much detail, however, is the security associated with these tasks. As the usage of mobile devices to interact with each other grows, so will the number and complexity of attacks upon these interactions. Therefore, there is an urgent need for better security associated with these relationships.

In this paper, we propose a solution to this issue, by presenting a security service model for establishing and using secure communication between two or more devices in a truly ubiquitous and isolated environment. Our model makes use of a two dimensional colorized barcode solution known as *Ubicolor* to establish a secure link between two devices. Once the link is established, secure communication can take place, which includes various types of interactions between mobile devices and service points. We discuss potential applications of this model, including improvements to previously proposed solutions, through the extension of the types of interactions that can occur and the limitations that can be placed on those transactions.

The remainder of this paper is organized as follows. Section 2 discusses background issues and related works. In Section 3 we detail our model, including specific examples of how it can be integrated into everyday use. Section 4 contains discussion of this model, including analysis of possible attack scenarios. Section 5 explores topics for future exploration and concludes this paper.

2 Background and Related Works

There have been several approaches to addressing the issue of establishing a secure channel in an ubiquitous computing environment. In this section we discuss them in conjunction with three different, but related aspects of ubiquitous computing. The first is authentication and key exchange in an isolated environment, the second is the use of camera-enabled devices to capture and analyze visual images, and the third is the combination of the two, using camera-enabled devices to authenticate and exchange keys in an isolated and ubiquitous computing environment.

2.1 Authentication and Key Exchange

In order to exchange a shared secret key between two previously unknown devices, it is necessary to first authenticate those devices so that the keys are exchanged between the intended devices only. This eliminates the chance that an attacker can successfully execute a man-in-the-middle attack, thus intercepting

and monitoring communication between the devices. One of the first works to address this issue was presented by Stajano and Anderson [20]. Their *Resurrecting Duckling* protocol performs a key exchange between two devices by initializing one of the devices and assuming that the first request for key exchange, known as "imprinting," originates from a trusted source. The imprinting in this particular protocol is done via physical contact.

Key exchange is expanded to include other methods besides physical contact by Balfanz et al., who refer to this technique as using a *location limited channel* [2]. Specifically, they note that "Location-limited channels have the property that human operators can precisely control which devices are communicating with each other". For usability purposes, as well as to reduce possible errors in data entry, electronically controlled methods are mainly employed as the location limited channels used to exchange information for secure key exchange. Such channels could include physical contact, infra-red communication, audio transmission and detection, or displaying images for capture by digital cameras [2, 6, 8, 10, 22].

2.2 Using Digital Cameras

Several works deal with the topic of using camera-enabled devices to capture and analyze images of barcodes, or *visual tags*. The physical aspects of using cameras to recognize and capture visual tags are addressed in [7, 14]. Others address embedding information about products or services in visual tags, for recognition and potential communication redirection by the camera-enabled device [11–13, 15–17].

The use of digital cameras to capture visual tags for establishing communication between devices is explored by several papers [8, 18, 19, 21, 22]. The importance of using digital cameras to capture visual tags is that they provide *demonstrative identification* [2, 22]. This is the ability of one user to visually determine that the device displaying the visual tag is the device intended for communication.

2.3 Using Visual Tags for Authentication and Key Exchange

The use of visual tags to enable secure key exchange between mobile devices is a relatively new area of study. [8] proposes the exchange of a public key using a sequence of visual tags displayed on one device, and captured on another. This model does not use information contained in the visual tag to establish a network connection, however, it only uses visual tags to verify the authenticity of a public key. [18] proposes using visual tags to identify the address of a Bluetooth-enabled mobile device in a pervasive environment. The approach detailed in [18] is concerned with eliminating the time it takes for Bluetooth device discovery. This approach does not address the security of the communication between the devices, except to note that native Bluetooth encryption could be used. Additionally, authentication is addressed by using public keys signed by a certificate authority, which then could be either accepted or rejected by the other user.

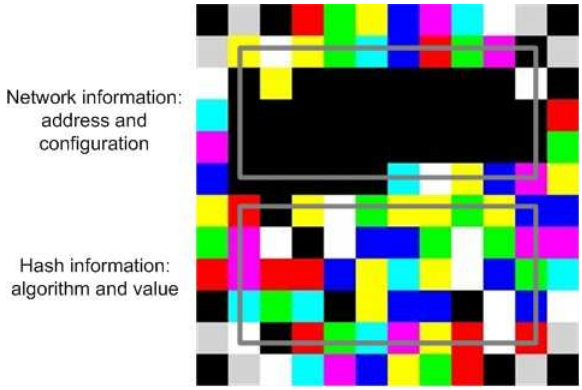

Fig. 1. UbiColor Visual Tag

[19] proposes a method of conveying both network and public key information in a single visual tag. Called *UbiCode*, this method of key agreement does not require an existing network connection between devices, nor does it require the transmission of the entire public key in the visual tag. Rather, the hash code of the public key is used to verify the validity of the public key itself, once it is provided by the device at the network address also included in the visual tag.

An extension of this idea involves using color in the visual tag. Called *UbiColor* (Figure 1), this approach increases the amount of data that can be contained in the visual tag, while also reducing the number of squares in the tag needed to convey the necessary information. Both UbiCode and UbiColor are capable of operating via various networking protocols and between two or more camera-enabled devices. We use these methods as key components the model of interaction presented in this paper.

[21] presents a "Mobile Service Toolkit, which is used to facilitate mobile devices accessing site-specific services. This model uses visual tags containing a network address to direct a mobile device to connect to a service-providing device. Once connected, personal information on the device may or may not be shared, and user interfaces can be pushed from the service device to the user device. This model builds upon [18] to establish network connections, but does not specifically address the security of the connections established.

A framework for mobile interactions with the physical world is presented in [17]. This framework describes interactions between users, devices, and various objects and service in the physical world. We build upon this framework to create our model for real-world interaction, focusing on the security of the transactions and the methods by which those transactions can be instantiated, monitored, and limited.

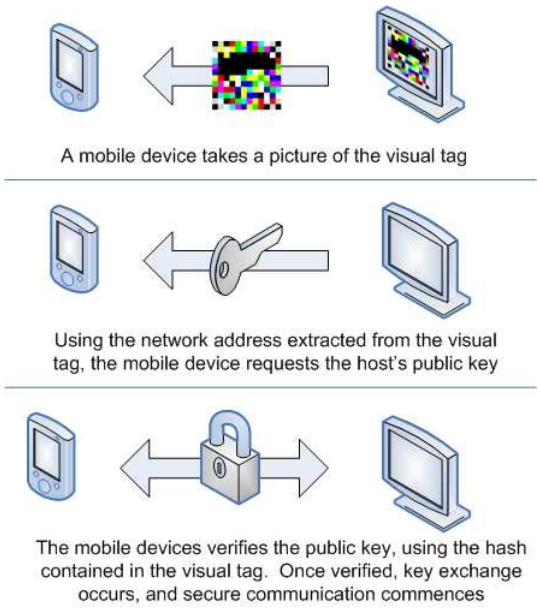

Fig. 2. Establishing Secure Communication

3 Security Service Model

Our approach is concerned primarily with establishing a secure communication channel, as quickly as possible, with no support required from outside entities, such as certificate authorities. We propose a model which builds upon the work done in [17–19, 21]. Specifically, we address the security of the connections established, and expand upon the possible uses of the devices involved in the transactions. The model is comprised of two distinct parts, the first is to establish a secure channel, and the second is to use that secure channel to access a service of some kind.

3.1 Establishing a Secure Channel

We make use of UbiColor to establish a secure channel between a camera-enabled mobile device, the client, and a screen-enabled service provider, the host. This process is shown in Figure 2. The first step is for the client to capture an image of the host device's screen, which displays a UbiColor visual tag. Once captured, the client processes the visual tag to extract two pieces of information:

- Network information: the network type and address of the host device
- Hash information: the hash algorithm and hash value of the host device's public key

Using the network address extracted, the client then sends an unencrypted request to the host for its public key, which is the same public key used to generate the visual tag displayed by the host[1]. The public key is sent, unencrypted, from the host to the guest. The guest then hashes the public key based on the hash algorithm retrieved, such as SHA-1, and compares that hash value to the hash value retrieved from the visual tag. If the two hashes match, then the public key is verified as having originated from the host device. This type of authentication does not use a trusted source or certificate authority, but rather makes use of demonstrative identification [2, 22] to verify the identity of the parties involved.

Once the public key is verified, a shared secret key can be generated using a number of various well-known authenticated key agreement methods, such as the Diffie-Hellman key agreement protocol [5]. Using this shared secret key, a secure communication channel is established, on a randomly selected port, using a symmetric key encryption method, such as Rijndael [3].

3.2 Using the Secure Channel

With a secure channel established, the next step in the process can begin, which is using the secure channel. We propose interacting with some sort of service-providing device. For instance, using an ATM to withdraw money or purchasing a snack from a vending machine are two possible uses of this model.

Security Service Scenario: ATM. In the first case, using an ATM, suppose Alice would like to withdrawal money from her bank account. First, Alice takes a picture the visual tag from the screen of the ATM, using her camera-enabled and bluetooth-enabled cell phone. Using the visual tag and the process described above, she establishes a secure channel. Using that channel, Alice's cell phone then authenticates, via the ATM, to her bank, using credentials provided in one of several ways:

- Manually keyed in by Alice at the time of the transaction
- Stored securely on Alice's cell-phone, and accessed by a pin keyed in at the time of the transaction
- Stored securely on Alice's cell-phone, and provided automatically during the transaction, without any input from Alice
- Stored securely on an attached device, such as a USB key or flash card, and provided at the time of the transaction. This option enables additional features which are discussed later.

Once authenticated, Alice interacts with her cell phone, either by manual entry using the key pad, or via speech recognition, to direct the ATM to distribute the desired amount of cash. Once authorized by the bank to do so, the ATM

[1] The host generates a new public/private key pair for each visual tag, to reduce the chance of spoofing, or a replay attack [4].

then dispenses the cash to Alice. It is at this point, and this point only, that Alice actually physically interacts with the ATM. Once the cash is retrieved, the secure connection is terminated, and the transaction is complete.

This approach provides several advantages to traditional ATM interaction. Some have previously been noted, such as the following [15]:

- Less physical use of the service device, thus keeping it cleaner.
- More options for physically locating the service device. It no longer has to be in an area immediately accessible by the user.
- Personalization, based on information stored in the user's device

Additionally, the following advantages exist:

- Better physical security when interacting with the device; the user can perform interactions from within a locked vehicle, for instance.
- No need for an ATM card, which can be lost or stolen.
- Protection from "skimming" scams [1], which use barcode readers and hidden cameras to capture user's card information and PIN codes. In fact, no PIN code need be entered at all during this process, eliminating even a casual observer from capturing sensitive data.
- Better sanitation, since the only thing the user touches is the money they withdraw.

Security Service Scenario: Vending Machine. Suppose Bob would like to purchase a snack from a vending machine at his school. Unfortunately, Bob does not have any cash, but he does have his camera-enabled and 802.11x enabled PDA. Additionally, he has a secure cryptocard containing his parents' credit card information, which he can access by entering a PIN code.

Bob decides which snack he would like to purchase, and initiates a communication request with the vending machine. He takes a picture of the visual tag displayed by the vending machine, and establishes a secure 802.11x connection[2]. Once established, Bob uses the secure connection and the display on the PDA to request the item he would like to purchase. The vending machine sends information about this product to Bob's PDA, and requests payment for the item. Normally, Bob would now enter the PIN code to access his parents' credit card information, but unfortunately for Bob, in this case his parents have specified which types of goods can be purchased using their credit card information. The criteria for the snack, provided electronically by the vending machine, does not meet the criteria set forth by Bob's parents, and he is forced to choose a more nutritious snack before the transaction can proceed. Once an acceptable snack is chosen, and the correct PIN code is entered, the credit card information is sent securely to the vending machine, which then provides an electronic receipt to Bob's PDA, dispenses the snack, and terminates the connection.

[2] Note that the use of 802.11x protection capabilities, such as WEP or WPA, do not need to be used - the connection is secure through the use of symmetric key cryptography between the client and host.

This example has many of the advantages of the previous example, in terms of lack of user interaction and device placement. But it also provides other distinct advantages to the model:

- Payment information does not necessarily need to be stored on the device itself.
- Controls can be implemented to limit the type and/or the amount of purchases that can be made.
- Electronic receipts can be provided at no cost to the service provider.

4 Discussion

The model described in this paper provides a means of establishing a connection between devices that is immediately secure. This eliminates any guesswork or additional steps with relation to the connection itself. By providing a secure connection, sensitive transactions, particularly those involving the exchange of personal information or financial information, can be performed without risk of the data being intercepted during transmission.

This model provides several advantages over traditional device interaction. Many of these advantages were listed in Section 3. In addition, this model extends previously proposed models by means of its versatility and ease of use. Our approach is not dependent on a particular type of network, and can easily be adapted to various network environments. Additionally, it does not require the presence of a certificate authority (or access to one) to establish a secure channel. In the particular examples provided, an external connection is necessary to validate user credentials, but is not necessary to establish a secure connection. If a financial transaction is not taking place, such as when user simply wants to retrieve information securely, then a connection to an outside network is not necessary.

Performance is not explicitly discussed in the scope of this paper. However, an implementation of UbiColor has been thoroughly tested. Generating and displaying a visual tag, including generating a public/private key pair, takes about 0.1 seconds on a laptop, and just over 1 second on a PDA. The average time to process images of visual tags on a camera-enabled PDA (specifically an HP iPAQ rx3715) is 1.28 seconds. Establishing a secure channel, based on the information retrieved, is heavily dependent on the type of network used. Symmetric keys are generated in less than 1 second, but it takes on average six seconds to actually establish a secure channel between devices, using an IP based 802.11g wireless network. Even with the extra time necessary to capture the image itself, this entire process still falls within an acceptable time limit.

Previous works address the issue of device discovery, and present methods of avoiding the delays related to device discovery, particularly when the user is visually aware of the intended device. Our approach eliminates the need for these techniques entirely, because no automated device discovery is necessary at any point. The network address contained within the visual tag is the only

information necessary for the client device to locate the host device in the wireless environment.

Questions may arise concerning the security of this approach. By using a location-limited channel, we eliminate the risk of a man-in-the-middle attack. Even if an attacker were to somehow capture the same image on his device, he could not intercept secure communication between the guest and the host; the attacker would have no knowledge of the shared secret key generated between them, because it is encrypted using the host's public key. Another potential concern is for the possible mis-identification of information contained within the visual tag. If the network information is misread by the client, then no communication attempt will be established with the host, and the process will fail. Even if a connection is made to an incorrect host, that host can not provide the correct public key, and the process will fail. Similarly, incorrectly identifying the public key hash will not allow the client to establish a secure link with any host. Because the visual tag's hash component changes with each new connection attempt, it would be infeasible for an attacker to replace the visual tag with an altered one. One possible way to alter a visual tag displayed by the host in a predictable manner would be to compromise the host itself, and replace the displayed visual tag with one directing the guest to a malicious host instead. In this case, since the host has already been compromised, the security of the transaction is no longer intact anyway. Alternatively, a pre-printed visual tag could be physically placed over the hosts display to redirect users to a malicious host, but this also seems rather far-fetched, as the necessary proximity of the user in order to capture the visual tag suggests the user would be able to detect a superimposed visual tag, versus one presented on the host's electronic display.

5 Conclusion

In this paper we have presented a model for establishing and using a secure link between two devices in a ubiquitous computing environment. Using UbiColor tags to facilitate the creation of a secure channel between the host and guest, and the capabilities of the devices involved to manage and interact with each other, we have explored several advantages to using such a technique. Additionally, we have discussed performance issues and related security issues, including examples of how using this model eliminates existing security risks associated with current interaction between users and service devices. This method of establishing and using secure communication between mobile devices in an ubiquitous environment should help to improve performance time and the overall user experience, which is key to the adoption of such techniques in real-world situations.

Better privacy control and more fine-grained access control on the information exchanged for the interactions in an ubiquitous environment should be an interesting topic. For instance, privacy or access control based on the context of those interactions can provide more dynamic and flexible security services. Our immediate future work will investigate those issues.

Acknowledgement

This work is partially supported by a startup grant from the research and economic development office of the New Mexico Tech.

References

1. ATM Skimming. http://www.globalcu.org/newsarticles/n_fraud1205.htm
2. D. Balfanz, D. K. Smetters, P Stewart, H. C. Wong. Talking To Strangers: Authentication in Ad-Hoc Wireless Networks. In *Proceedings of Network and Distributed System Security Symposium 2002 (NDSS02)*, San Diego, CA, February 2002.
3. J. Daemen and V. Rijmen. AES proposal: Rijndael. Advanced Encryption Standard Submission, 2nd Version, March, 1999.
4. M. Danzeisen, T. Braun. Secure Mobile IP Communication. *Workshop on Wireless Local Networks at the 26th Annual IEEE Conference on Local Computer Networks (LCN'2001)*, November 2001.
5. W. Diffie and M. Hellman. New Directions In Cryptography. In *IEEE Transactions on Information Theory*, vol. IT-22(6), pages 644654, November 1976.
6. M. T. Goodrich, M. Sirivianos, J. Solis, G. Tsudik, and E. Uzun. Loud and Clear: Human-Verifiable Authentication Based on Audio. Cryptology ePrint Archive, Report 2005/428.
7. T. R. Hansen, E. Eriksson, and A. Lykke-Olesen. Mixed Interaction Space Designing for Camera Based Interaction with Mobile Devices. In *CHI05 Extended Abstracts of the Conference on Human Factors in Computing Systems*, ACM Press, pp.19336, 2005.
8. J. M. McCune, A. Perrig, and M. K. Reiter. Seeing-is-believing: Using camera phones for human-verifiable authentication. In Proceedings of *The 2005 IEEE Symposium on Security and Privacy*, May 2005.
9. P. Panurach. Money in electronic commerce: digital cash, electronic fund transfer, and ecash. Communications of the ACM 39 (6), 1996, pp. 45-50.
10. I. Pavlosoglou, R. Ramirez-Iniguez, M. S. Leeson, and R. J. Green. A Security Application of the Warwick Optical Antenna in Wireless Local and Personal Area Networks. In Proceedings of *The London Communications Symposium 2002*, pp. 225-228, Sep 2002.
11. J. Rekimoto and Y. Ayatsuka. CyberCode: designing augmented reality environments with visual tags. *Proceedings of DARE 2000 on Designing augmented reality environments*, Elsinore, Denmark, 2000.
12. M. Rohs. Real-World Interaction with Camera Phones. *Ubiquitous Computing Systems: Second International Symposium, UCS 2004*, Tokyo, Japan, November 8-9, 2004.
13. M. Rohs and B. Gfeller. Using camera-equipped mobile phones for interacting with real-world objects. *Advances in PervasiveComputing*, pages 265271, April 2004.
14. M. Rohs and P. Zweifel. A Conceptual Framework for Camera Phone-Based Interaction Techniques. In Proceedings of *Pervasive Computing: Third International Conference, PERVASIVE 2005*, Munich, Germany, May 8-13, 2005.
15. E. Rukzio, A. Schmidt, and H. Hussmann. An Analysis of the Usage of Mobile Phones for Personalized Interactions with Ubiquitous Public Displays. Workshop on Ubiquitous Display Environments, UbiComp 2004, Nottingham, UK, 2004.

16. E. Rukzio, A. Schmidt, and H. Hussmann. Physical Posters as Gateways to Context-aware Services for Mobile Devices. In *Proceedings of WMCSA 2004*, English Lake District, UK, December 2004.
17. E. Rukzio, S. Wetzstein, A. Schmidt. A Framework for Mobile Interactions with the Physical World. Wireless Personal Multimedia Communication (WPMC'05), Aalborg, Denmark, September 2005.
18. D. Scott, R. Sharp, A. Madhavapeddy, and E. Upton. Using Visual Tags to Bypass Bluetooth Device Discovery. *The ACM Mobile Computing and Communications Review (MC2R) Special Section on Discovery and Interaction of Mobile Services*, January 2005.
19. D. Shin and S. Im. Visual Device Identification for Security Services in Ad-Hoc Wireless Networks. In *Proceedings of 20th International Symposium on Computer and Information Sciences (ISCIS 05)*, Istanbul, Turkey, October 2005.
20. A. Stajano and R. Anderson. The ressurecting duckling: Security issues for ad hoc networks. In *Proceedings of the International Workshop on Security Protocols*, 1999.
21. E. Toye, R. Sharp, A. Madhavapeddy, and D. Scott. Using smart phones to access site-specific services. IEEE Pervasive Computing, 4(2):6066, April 2005.
22. F. Wong and F. Stajano. Multi-channel protocols. B. Christianson et al. (Eds.): *Security Protocols* 2005, LNCS.