The Untrusted Computer Problem and Camera Based Authentication

Matt Burnside, Dwaine Clarke, Blaise Gassend, Thomas Kotwal, Marten van Dijk, Srinivas Devadas, Ronald Rivest

Introduction: The use of computers in public places is increasingly common in everyday life. In using one of these computers, a user is trusting it to correctly carry out her orders. For many transactions, particularly banking operations, blind trust in a public terminal will not satisfy most users. Our aim is therefore to provide the user with authenticated communication between herself and a remote trusted computer, via the untrusted public terminal.

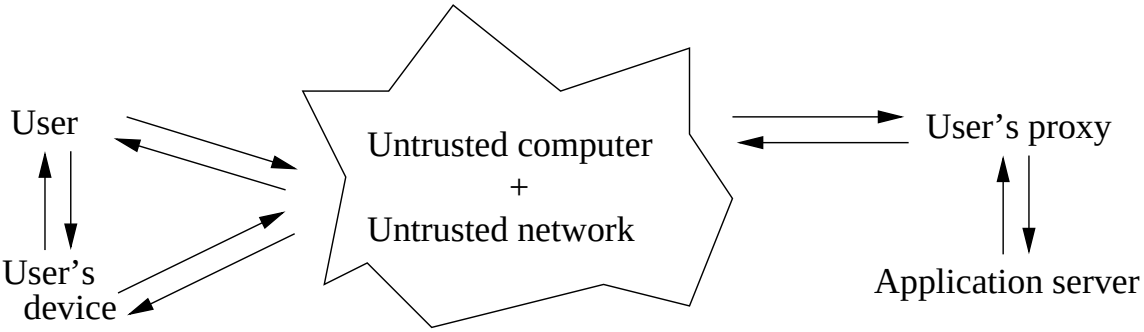

Figure 1: The untrusted computer model : The user, equipped with her device, wishes to establish authenticated communication with her proxy over an untrusted channel. Once this is done, she can safely use any application on her proxy, or on some remote application server (such as her bank)

Approach: We will assume that the user is using an untrusted computer to contact a trusted computer over a network. The user will have a proxy running on the trusted computer, which will then contact other parties, such as a bank. Our goal is to present methods to provide an authenticated bidirectional channel from the proxy to the user, through the untrusted computer. Authenticated means that messages received through the channel are guaranteed to be unmodified copies of messages that were sent by the party on the other side of the channel.

In order to accomplish this goal the user will have a trusted device. There are two basic designs for this device. The first is that the device receives encrypted messages sent through the untrusted computer, decrypts the messages, and displays them to the user. In the second design the user sees the messages displayed on the untrusted computer while the device monitors these messages to ensure their validity.

In a very straightforward implementation of the first design, the user comes to the Internet cafe with her PDA. She connects it to the untrusted computer's USB port, and connects to her proxy using her PDA's SSL-capable web browser. She can then do all her interaction with her proxy through her trusted PDA.

However, all of these solutions are dissatisfying because the user is barely making any use of the untrusted computer's comfortable screen. The Internet cafe's computer is just being used as a network access point. Meanwhile, the user has to study the stock market through the tiny screen of her hand-held device.

With monitoring devices, the user gets information directly from the untrusted computer. Some extra information is also sent through the computer, that the user need not concern herself with, but that the device uses to verify the authenticity of the information the user is getting. If the device detects tampering, or if for some technical reason the device is unable to authenticate the image (too much noise, the connection from the untrusted computer to the device is bad, etc.), it warns the user. Monitoring approaches are more convenient as the user fully uses the untrusted computer's interface. However, a number of difficulties arise.

First, it is important to realize that the device must be authenticating the information that the user is getting. It would not be acceptable for the device to receive information about what the user is seeing on the screen through a USB link, as a malicious computer could send one thing through the USB port, and something completely different to the screen. This means that to authenticate screen content, the device must be equipped with a camera.

Even if the device and the user are both getting their data from the same source (we will consider a screen), caution is still required because of noise. Indeed, it is impossible for the device to reconstruct what is being displayed on the screen down to the exact RGB components of each pixel. Variations in screen brightness, camera noise and reflections off the screen all contribute to imprecision in what the device can reconstruct. If a rogue message is displayed in grey on a slightly lighter grey, the user will be able to read it, while the device might see it as uniform grey. It is because of this difficulty that our implementations use black and white (no grey) images.

Thus we see that if the device does not perceive as much information as the user, then the user must be aware of that limitation, and be able to tell when an image might be ambiguous to the device. For example, if the device can only distinguish between black and white, then a shade of grey should reveal to the user that tampering has taken place, even if the device does not detect that tampering.

Progress: Currently under development are two implementations of camera-based authentication, which will be referred to as the pixel mapping and the optical character recognition (OCR) methods. In both cases the user is expected to carry a camera-equipped device that monitors the screen of the untrusted computer she is using. The visual processing involved in extracting on-screen information can be costly in computation resources. The first method we propose tries to minimize this cost, while the second one uses a high bandwidth network connection to move the computation to the proxy.

In the Pixel Mapping method, the camera-equipped device is assumed not to move relative to the screen during the authenticated session. An initial calibration phase is used to construct a mapping between screen pixels and camera pixels. The mapping is then used by the device to exactly reconstruct the screen content. A small area at the bottom of the screen is used to transmit a nonce, a one-time password and a message authentication code (MAC).

In the OCR method it is assumed that the user's device is equipped with a camera and an infrared link to the untrusted computer. This link is used to exchange data with the proxy, via the untrusted computer, to take advantage of the large amount of computation available at the proxy. The device takes a picture of the screen, sends the picture to the proxy, and then the proxy uses OCR to read the text in the picture of the screen.

Future: We have been most concerned with authenticating communication in which the user is receiving visual information. Our protocol could be applied just as well to audio information. Though this is probably not very useful to the average user, it would certainly benefit the visually impaired.

The camera-based system does not provide any privacy, as queries and responses are transmitted in the clear. A limited amount of privacy could be added by allowing the user to point at areas of the screen. Selections made in this way would be visible to the device but not to the untrusted computer. The possibilities of such a system are yet to be explored.

Research Support: This research was supported by Acer Inc., Delta Electronics Inc., HP Corp., NTT Inc., Nokia Research Center, and Philips Research under the MIT Project Oxygen partnership, and by DARPA through the Office of Naval Research under contract number N66001-99-2-891702.

References:

[1] M. Naor and B. Pinkas, "Visual authentication and identification," in *CRYPTO*, 1997, pp. 322–336.

[2] M. Abadi, M. Burrows, C. Kaufman, and B. W. Lampson, "Authentication and delegation with smart-cards," in *Theoretical Aspects of Computer Software*, 1991, pp. 326–345.