Project Number: HNH-HH07 52

WEB BROWSER SECURITY

AND USER AWARENESS

An Interactive Qualifying Project Report

submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

by

Alexander Jansky

Wojciech Krajewski

Aaron Maestre

Aaron Maestre

Date: April 29, 2003

Approved:

Huong Higgins

Professor Huong Higgins, Advisor

Abstract

The goal of this project was to explore security features encountered while browsing the Internet, and to test people's knowledge of them. We chose this subject due to security issues of online commerce. Our research yielded that encryption based on the Secure Sockets Layer (SSL) was most widely used. A survey, administered to college students, found that only half were well informed. This leaves a large population to be educated in order to feel comfortable engaging in online commerce.

Acknowledgements

We would first like to thank our advisor Professor Huong N. Higgins for making this Interactive Qualifying Project possible. She provided us with a great deal of comments, suggestions and ideas on ways to improve this project.

We would also like to thank David Yakovich, the Web Developer at the Town of Manchester in Connecticut for his time and suggestions relating to this report. He gave us excellent insight into the field of Internet Security.

We would like to extend our thanks to everyone who assisted in this project and made it possible.

Authorship

These are the tasks which each member of the group preformed.

Written by All:

- Introduction
- Design of Questions for Survey
- Conclusions and Recommendations

Alexander Jansky and Aaron Maestre:

- Survey
- Survey Data Analysis

Alexander Jansky:

- Background
- Security Alternatives to SSL

Wojciech Krajewski:

- Methodology Introduction
- Encryption Technologies
- Interview
- Survey Webpage Creation & Data Collection

Aaron Maestre:

- HTTP vs. HTTPS
- Certificate Authorities

TABLE OF CONTENTS

Table of Figures

1 INTRODUCTION

1.1 Goal

The goal of this project is to explore Internet security features which the average individual encounters while browsing the Internet, and also to discover how knowledgeable people are concerning these security features.

1.2 Purpose

The reason we chose this topic was that the Internet has become widely used, making personal security an issue of concern. Security has been an issue since the introduction of the first popular web browsers. The first web browsers made it possible for the Internet to be easily accessible by the average person. Today, the Internet is used to purchase a vast array of items from clothing to real estate. Banks have also recognized the convenience of using the web and now provide online banking services. Since the Internet has become a popular tool for conducting commerce, we decided to focus on the Secure Socket Layer (SSL), which is the most important security feature that is used by web browsers. With the increasing number of opportunities for users to transmit personal information there is also an increased chance of a third party intercepting the information and committing fraud. Keeping critical data such as credit card and social security numbers private will be an individual's primary concern while transmitting personal information. As a result, SSL is very important to anyone making Internet transactions because it is the security feature that encrypts their personal information as it is transmitted.

1.3 Procedure

To effectively test people's knowledge of online security, we needed to research and become well informed of the technology available. Once the technology was studied in detail, we determined which security features are most prominent. This was done through reading and analyzing books, articles, and other sources from the Worcester Polytechnic Institute Gordon Library and online. We set up an interview with an Internet security expert to give us more insight into what we were researching. With that background, we conducted a web based survey that allowed college students to rate their understanding of Internet security features, and also allowed us to test their knowledge. The research, interview, and survey allowed us to determine which technologies are prominent and whether they are recognized by individuals.

1.4 Results

Through our research we discovered that the Secure Socket Layer (SSL) is the most widely used and popular method of securing transactions on the Internet. Most information transferred over the Internet is not encrypted in any way. SSL allows the encryption of information so that it cannot be easily intercepted by a third party. From the survey, we were able to conclude that seventy-six percent of the respondents used online purchasing and banking to some extent. We found that, of this percentage, half the respondents were not knowledgeable of security features. Another finding was that people who made online purchases tended to be more knowledgeable concerning online security.

1.5 Significance

From our results, we found that a very large percentage of people make online purchases, yet only half of these individuals are aware of the technology involved in keeping their personal information safe. This leaves a large portion of the population susceptible to fraud. Certain measures need to be taken to encourage the majority of the population to educate themselves about Internet security when engaging in online commerce. By making information relating to Internet security readily available, companies can instill comfort in consumers when making online purchases.

The following sections of this report will include a Background section which will discuss the historical context of the Internet and the particular security features pertaining to web browsers. After that, the Methodology section will cover the steps we have taken to research web browser security and relate our research to society. In the Results, we will go into greater detail about the data collected during the survey. In the Conclusion we will then summarize our findings, speculate on future developments, and give advice on how people can become better informed of the conclusions we drew from our research.

2 BACKGROUND

2.1 Historical Background

People now use the Internet everyday in a vast amount of ways. The Internet has made its way into our finances, our education, our commercial uses, and has become an incredible tool for finding information. To understand how the Internet has become such a powerful tool it is important to take a look at the Internet's short history.

2.1.1 Brief History of the Internet

In the early 1970's the U.S. Defense Department's Defense Advanced Research Projects Agency (DARPA) initiated a research program on how to link packet networks. A packet network is a network that transfers information by using packets. Basically, data is broken down into smaller pieces called packets, which are sent out and reassembled at their destination. The original intent of the research project was to design protocols that allowed networked computers to communicate with many other packet networked computers. This project was dubbed by DARPA the "Internetting project", and the networks that were constructed as a result of this project came to be called the Internet (www.isoc.org). Due to the fact that the beginnings of the Internet were military related, security was always under consideration. During this time the basic architecture of the Internet was formed. The two basic protocols, which are still used today, were first developed in the DARPA project. These protocols were the Internet Protocol (IP), and on top of that is the Transmission Control Protocol (TCP) (www.isoc.org). DARPA's Internet known as the DARPANET soon was to be replaced by newer better networks. In the early 1980's, networks such as the BITNET (Because It's Time Network), and the

National Science Foundation's CSNET (Computer Science Network) came into use (www.zakon.org). Since these networks were for non-military use, the Internet came into more public use. They took the emphasis off of the DARPANET, and allowed for more room on improvement.

The evolution of the Internet became guided by other means instead of just the government. Industry and the academic community became strong forces in the Internet development, although they both worked in strong collaboration with government agencies. In 1983 the Internet Activities Board (IAB) was created to guide this evolution. This Board itself now has two sections to it. These are the Internet Engineering Task Force (IETF), and the Internet Research Task Force (IRTF) (www.isoc.org). By 1990 the DARPANET finally faded out. The next biggest breakthrough in the Internet evolution took place in 1991. Tim-Berners Lee released the development of the World Wide Web (WWW) (www.zakon.org). This made the Internet available and appealing to the average person. With this wide spread use, unsuspecting computer-illiterate people became easy targets for the everyday thief.

2.1.2 Historical Context of Internet Security

Luckily the original architects of the Internet and the World Wide Web had already taken into consideration the security threat. These problems were first addressed by Netscape Communications Corporation in 1993, when the Secure Socket Layer (SSL) was designed and introduced. SSL was first implemented in general web use when Netscape Communications released Mosaic Browser Version 1.0 in November 1993. Mosaic became the first popular web browser. SSL Version 2.0 was later released with Netscape Navigator. The development of SSL was not unilateral because many contributions were made to SSL development from the web community. This is due to the fact that Netscape encouraged this development even though Netscape did own a patent on the SSL protocol. One example of this is when the Microsoft Corporation developed the Private Communication Technology (PCT), which enhanced the SSL protocol. Starting in May 1996, the Internet Engineering Task Force took on the job of developing SSL. This same task force is responsible for the creation of other standard Internet protocols such as TCP. SSL also later became renamed to Transport Layer Protocol (TLS). This was done by the task force to show that they were not harboring favoritism towards one company. Despite the name change, TLS is basically identical to SSL. TLS is just a slightly newer version. Today, support for these protocols is built into most web browsers (Thomas, Sec 1.2). Along with the development of SSL came the development of certificate authorities which put in place a system that allows the user to verify the identity of the server they are sending information to. SSL with the use of certificates has become the current standard for securing information in transit over the web.

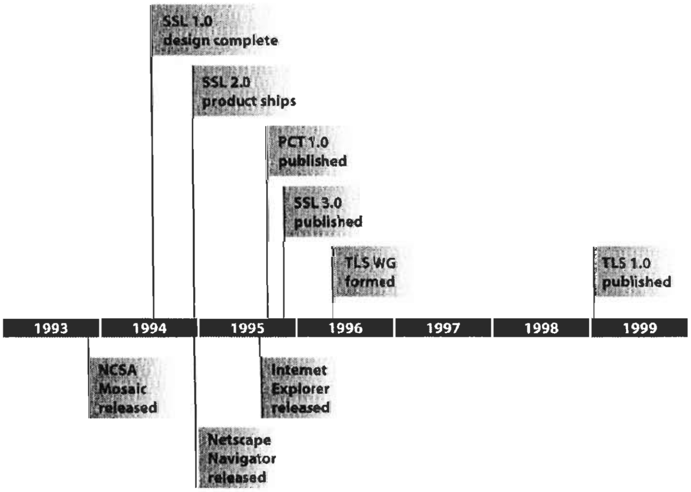

Figure 2.1: SSL Development (Thomas, Sec. 1.2)

3 METHODOLOGY

The previous section was intended to provide us with a basic understanding of Internet communications. To be able to judge different Internet security technologies we needed to have a much more thorough understanding of each topic introduced in the background. The following sections of this report will explore, in detail, the topics of HTTP, authentication methods, encryption, SSL, and more. Once each subject is studied in detail, we will determine which technologies are most prominent. To help us assess the technology, we conducted an interview with an expert in the field. The goal of this project however is not only to gain an understanding of the technology, but also to discover how often people encounter the technology and whether they have a good understanding of it. To achieve this part of the goal, we decided to survey a population of college students to determine their knowledge of available technology.

3.1 Procedure

3.1.1 Research

The background research done in the early part of this project introduced us to important technologies involving Internet security and their development. We needed to explore all of these topics in great detail before coming to any conclusions. In order to do this we first researched the companies which developed the original technologies. Netscape Communications Corporation for instance was the original creator of SSL, while GeoTrust and VeriSign are the major SSL certificate distributors. We were able to find a great deal of sources from both the WPI Gordon Library and reputable Internet websites.

3.1.2 Interview

The research provided us with a great deal of material to work with; nevertheless, we decided to verify the information with someone in the field. We were able to conduct an interview with David Yakovich, the Web Developer for the town of Manchester in Connecticut. David confirmed our research and was able to provide us with several of his thoughts on the future developments of Internet commerce, which will be included in our final assertion.

3.1.3 Survey

Once a detailed understanding of Internet security features was established, we were able to complete the second part of our goal. To find how aware people are of Internet security features, we conducted a web based survey. We asked WPI and other college students to participate in the survey. It was designed in a way that allowed us to collect information representative of how much people knew about Internet security features, and how much they believed they knew. The information gathered in the survey can be used to find correlations between people's purchasing habits and their knowledge. This can then be used to make recommendations to banks and companies on how to make consumers feel more comfortable and what information they should provide on their websites.

The sections below will give a thorough explanation of the Internet security features available. Then, the interview will be described in detail, followed by a discussion of what we expect to find from the survey.

3.2 HTTP vs. HTTPS

As a computer user browses the Internet a lot of activity takes place in the background. Great amounts of information are transferred between the user's web browser and the site's web server. When a user attempts to access a site, the web server responds by sending the Hypertext Markup Language (HTML) that makes up that site. The user's web browser receives this information and determines if there are images, applets or text to load. It then displays the images as they are downloaded. This is all accomplished using the Hypertext Transfer Protocol (HTTP), the protocol that web browsers use to transfer information over the Internet. A protocol is a set of rules that governs a process, in this case the process of transferring HTML from a location on the Internet (Heaton, Chapter 3).

3.2.1 What HTTP Does

HTTP moves web content such as plain text, pictures, movies, and audio files quickly and reliably from web servers to web browsers on user’s computers all over the world. It uses reliable data-transmission protocols so as to guarantee that data can not be scrambled or damaged while it is transmitted. Web servers provide the data when it is requested by web browsers. Browsers send request messages to servers when data is needed. Servers respond with response messages. These are the only types of HTTP messages (Gourley, Chapter 1.1).

3.2.2 The Need for Authentication

When the early Internet protocols such as HTTP were designed security was not a major concern. The Internet was not intended to be the huge global network that it is today. As the Internet is used more and more in the daily lives of people around the world, security becomes much more of a concern. People perform private transactions and access private information via the Internet. Due to the ease of performing these transactions, there needs to be some assurance that people's private information is safe. Not all information is intended for the general public. People need to feel assured that unauthorized users can not view their private information or publish documents on their websites without their consent (Heaton, Chapter 3).

Web servers decide who can access what type of information by requiring authentication. HTTP utilizes a challenge/response structure to authenticate users. When a web application receives an HTTP request message, the server can respond with an authentication challenge. This tells the user to prove who they are by providing a username and password. If the information is incorrect the server can challenge the user again or display an error. If the information is correct the server grants access. This process is illustrated in the figure below (Gourley, Chapter 12.1).

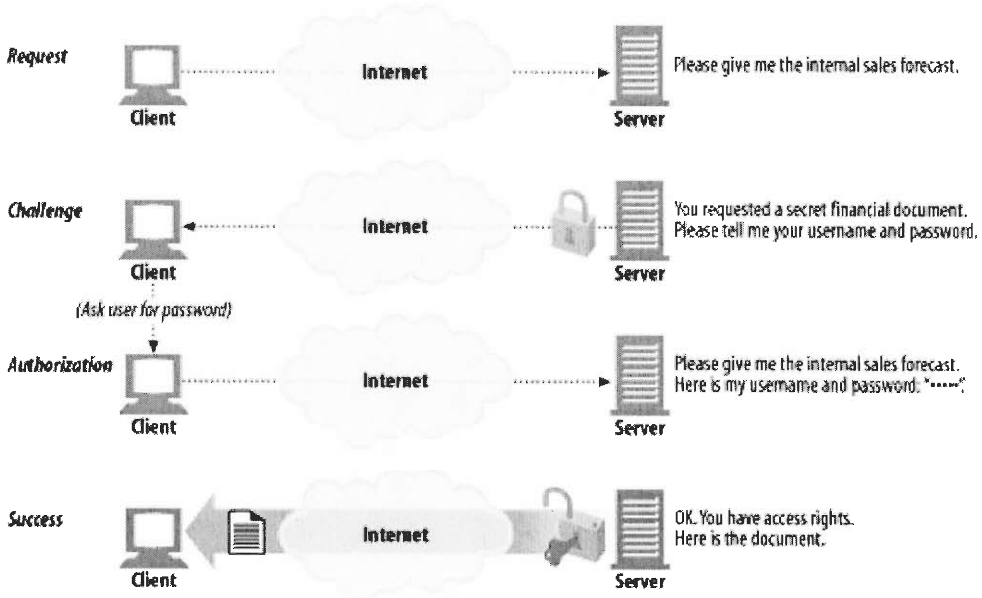
Request
Client
Internet
Server
Please give me the internal sales forecast.
Challenge
Client
Internet
Server
You requested a secret financial document.
Please tell me your username and password.
(Ask user for password)
Authorization
Client
Internet
Server
Please give me the internal sales forecast.
Here is my username and password: "•••••"
Success
Client
Internet
Server
OK. You have access rights.
Here is the document.

Figure 3.1: Simplified Challenge/Response Authentication (Gourley, Sec. 12.1)

3.2.3 Basic Authentication

Basic authentication is the most popular HTTP authentication protocol. In basic authentication, a server can refuse a transaction and challenge a user to provide a valid username and password. When the user's browser receives the challenge it opens a dialog box asking for a username and password. When the user enters the information it is sent back to the server. HTTP basic authentication encodes the username and password using base-64 encoding. Base-64 encoding was developed to temporarily convert text into a portable alphabet for communication. It could then be decoded without the fear that it would become corrupted during transmission (Gourley, Chapter 12.2).

Although convenient to use, basic authentication is not secure. The information being sent is encoded plain text which could be intercepted by a third party. It is meant to be used in a friendly environment to prevent malicious users from obtaining access. Ideally basic authentication should be used in combination with some encryption technology such as Secure Sockets Layer (SSL). Without the use of encryption technology basic authentication has many security flaws (Gourley, Chapter 12.2).

Base-64 encoding allows the transmission of the information without having to worry about accidental interception of passwords. However the information can be decoded relatively easy by reversing the encoding process. For this reason it is important that base-64 encoding only be used in friendly environments. Basic authentication is sometimes used by fake servers. A user can be led to believe that they are connecting to a certain server by basic authentication when they are really connecting to a fake server. Then when they enter their username and password the server can store the information and respond with an error. The server then has the information and can use it at any time, possibly for malicious use (Gourley, Chapter 12.3).

Despite the security risks, basic authentication is useful in providing convenient access to many users in a friendly environment. It should not be used when privacy is absolutely necessary. If privacy is necessary basic authentication should be combined with encrypted data transmission such as SSL to conceal the username and password from malicious users. This is a common technique known as secure HTTP transactions (Gourley, Chapter 12.3).

3.2.4 HTTPS

The introduction of the Hypertext Transfer Protocol Secure (HTTPS) provided the necessary privacy. HTTPS is the most popular way of securing HTTP. It was developed by Netscape Communications Corporation and is supported by most browsers and servers. When using HTTPS all the HTTP request and response data is encrypted before being sent over a network. HTTPS works by providing a transport-level cryptographic security layer using the Secure Sockets Layer (SSL) (Gourley, Chapter 14.1).

The average user does not notice the HTTP or HTTPS specification in their browser. Most are unaware of browser security in general. The only clue that something is different is the lock symbol that appears in the status bar of secure pages and the https scheme. Other than that a secure document looks the same as an unsecured one. Although visually HTTP and HTTPS appear the same they are very different. Unlike HTTP, HTTPS is encrypted between the browser and the server. A complex series of validations ensure that the information being transmitted can not be intercepted by a third party. HTTPS also assures a user that they are connected to the website they expect. It achieves this by telling the browser to warn a user if it is unable to authenticate that the site the user is trying to communicate with is the correct one (Heaton, Chapter 3).

Once the HTTPS packets are received and decrypted, the elements of the protocol are the same as those of the HTTP protocol. HTTPS is security-oriented, meaning that it protects information as it is transmitted between the web server and the web browser. Using HTTPS ensures a web server that the information can not be intercepted by a third party (Heaton, Chapter 3).

3.3 Encryption Technologies

For web commerce to be successful, information such as credit card numbers, people's names and addresses must be kept secret. As covered in the section on HTTP and HTTPS, most communication methods and authentication methods can be listened to by malicious parties. To stop foreign parties from obtaining passwords, credit card numbers and other private information, it must be encrypted. There are several methods in which information can be encrypted. The simplest type, called symmetric encryption, uses a secret lookup table that both parties can refer to. Another method, which is much more powerful, is called asymmetric encryption and uses public and private keys. There are also methods that combine the use of symmetric and asymmetric encryption.

Symmetric encryption, also called secret key cryptography, is a scheme which has been used since the first use of languages. Kings and leaders have always needed to communicate with their armies without anyone else intercepting their communications. In symmetric encryption both parties must know two things, how the data is encrypted, meaning the algorithm, and a reference key which is held by both the person encrypting the data and the recipient. They both have the same exact key, and it is used to both encrypt and decrypt the contents of a message. A key can be something as simple as an instruction to shift every letter in a text one or more letters up or down in the alphabet. For instance the letter 'a' would become a 'b,' and 'b' would become 'c,' and so on. Other schemes can be used where each letter is paired with another letter of the alphabet. Then to encrypt a message we look up the letter we need, and replace it with the letter it is paired with. In short, all types of encryption methods which use any sort of substitution or permutation algorithm fall into the category of symmetric encryption. Over the years symmetric encryption methods have become very complex and secure. In addition, to make the encrypted data even harder to decipher, we can simply increase the length of the key, which dramatically increases the complexity of the encrypted data. Another benefit of such a method of cryptography is that it is rather simple to perform. If done by a computer for instance, only a small number of computations are needed to encrypt a piece of data (Singh, pgs 3-14).

There is however one great disadvantage to symmetric encryption. Both parties must know the secret key and what algorithm is being used before they can communicate securely. This means that the two parties must meet at some point in a secure location to decide on an algorithm and key prior to their secure communications. If the key is ever passed along a non secure path and someone else intercepts it, the message from either party can be easily decrypted.

We can clearly see that this method of encrypting data would be rather impractical over the Internet because of the difficulty of initially establishing a secure connection to exchange the key and algorithm that is to be used. It would be close to impossible for every customer who ever buys a book from Borders.com, for instance, to first go to the store and exchange a key and algorithm which they will use when making purchases over the Internet with Borders.com. This would have to be done with every store a customer would ever want to buy anything from. Such a scheme would make Internet commerce awfully inefficient. For this exact reason, symmetric encryption is not the best suited for establishing secure connections across the Internet.

Asymmetric encryption is another method used for encrypting data, but it has several advantages which make it very well suited for Internet use. Another, more familiar, name used for asymmetric encryption is public key cryptography. This name describes what happens in this encryption scheme quite well. There are two keys involved in transmitting a message from a sender to a recipient when using this method. One key is called the public key, and the other is called the private key. Both keys are generated by one person and they are generated together, as a pair. The public key is made public, even published in a newspaper for everyone to see. The private key however is kept secret. When someone wishes to send a private message to you, they use your public key to encrypt it with an algorithm such as the Rivest-Shamir-Adleman, or RSA as it is commonly referred to. Even though the public key used to encrypt the message is available to everyone, it is computationally infeasible for anyone to decrypt the message with only the public key. However, the person holding the private key can easily decrypt the message. The concept seems almost magical, but it really has a very good mathematical backing. A very simple way to look at how this works is by taking two numbers, for instance 368 multiplied with 246. This gives us the answer 90528. Now if we try to do this in reverse, if we have the number 90528, how do we find the original two numbers that were multiplied together? They could have been 6 multiplied with

15088 or 8 multiplied with 11316. Finding exactly which two numbers were originally multiplied is almost impossible. These two original numbers would be like the data encrypted with the public key. Without knowing the private key which gives us a method of going in reverse, it is very impractical and time consuming to decipher the data (Stalling pgs164-169).

We can see that this is a very powerful tool used on the Internet where individuals cannot possibly exchange secret keys with everyone else on the Internet in person. Using public-private key cryptography, two parties can each have a public key which they share with each other. When a message is sent from the first party to the second, the message is encrypted with the second parties' public key, which only the second party can decrypt. The reverse is done when the second party wants to send a message to the first party. This method has been proven to be very reliable and secure, however, nothing is perfect. The down side of public-private key cryptography is that it is very computationally intensive. A great deal of time is needed to encrypt and decrypt messages. If a server run by a bank or a business had to encrypt and decrypt every piece of data using public and private key cryptography, it would quickly become overrun with requests which it could not fill fast enough. Since symmetric encryption is many times less computationally intensive, but makes establishing a secure connection difficult, using the two methods together makes sense (Thomas, Section 2.2).

3.3.1 What is SSL and TLS

The Secure Socket Layer (SSL) and the Transport Layer Security (TLS) are standards which have been developed for secure communication between computer systems. The two terms, SSL and TLS are almost interchangeable. The difference between SSL and TLS is that SSL was first developed by the Netscape Corporation. Once the version number of this standard reached 3.0, it was submitted to become an Internet standard worldwide. The Internet Engineering Task Force (IETF), which decides on standards that are used on the Internet then gave SSL version 3.0 the title TLS, and has since then been developing it in a more open approach accepting input from outside parties on how to improve the standard. Even though TLS is now the official name for this standard, the name SSL is still more commonly used (Stalling, pg 444).

SSL is a standard that defines how to use the features available in symmetric and asymmetric encryption. The standard describes how to establish a secure connection between two computer systems, how to allow them to verify who they are connected to, and then how to transmit data over a secured connection while verifying its integrity, and finally how to end the connection so that an intruder cannot pretend to be one of the parties. When accessing a secure website on the Internet, the SSL standard is used to encrypt HTTP connections, and make them secure. This is where the term HTTPS comes from.

To establish a secure connection, a client, such as the web browser someone might use, first contacts the web server. The client first sends a message called a "ClientHello." This is an unencrypted message which tells the server that it wants to establish a secure connection. It also tells the server which versions of SSL or TLS it supports, different types of compression algorithms that the client supports and some other information. The server then replies with a response to which version of SSL the two will use and what compression method to use. The server also sends the client its public key for which it holds a private key. Using the public key received, the client can now send the server a message containing the exact algorithm which will be used in their communications. Asymmetric encryption is only used to transmit this single message, saving a great deal of server and client resources. Now that both the server and client have the same secret key, they can use symmetric encryption to continue with the connection. This method of using both symmetric and asymmetric cryptography allows the two systems to establish a secure connection, and then switch to a secret key method which makes the transmission of data much faster and simpler for both sides (Thomas, Section 3.3).

Even though this is a very effective way of establishing a secure connection between two parties, and is very close to impossible to crack, there is of course one problem. Just because the data being transmitting is as securely as it can be, there is no way to tell weather the connection has been established with the intended individual, or an imposter. Since public keys can be freely distributed, and there is only one private key with which encrypted data can be decrypt, private keys can be used to identify an individual or server. As long as a server never changes its private key, then it is the only place in the world that can decrypt any data sent to it encrypted with its public key. By keeping track of public keys, and knowing who the real owner of the key is, we can distinguish between the real server and an intruder posing to be the server. For this exact reason, several organizations have been set up to keep track of public keys. Servers then can buy certificates from these companies which prove that they are who they really are.

3.3.2 Certificates

When most people think of security, their first thoughts are of encryption, and scrambling messages in such a way that only the intended recipient can understand them.

There is one very important aspect which is rarely remembered. How do we know we are sending our encrypted information to the right party? No matter how good an encryption algorithm is, if we authenticate ourselves with an intruder, then anything we send them will be encrypted specifically for them, meaning they will able to read it with absolutely no problem. For secure communications to be possible there must be a way of preventing this type of attack. Again, the use of public-private key cryptography can be used to facilitate the situation (Thomas, Section 3.5).

As mentioned in the previous section, with public-private key cryptography, we have a public and private key that are matched with each other. There can only be one private key for one public key. The public key is made available to the general population, and the private key is used to decrypt the message which is received encrypted with the public key. The problem with just making keys public is that anyone can say they are someone else. For example, if user A wants to send user B a secured message. They find a key which was made by user C, who claims to be user B. When user A sends a message encrypted with the public key of user C to user B, user B cannot read the message, while user C who is an intruder can listen in to the communications between the two and read the message. This could obviously be a huge security flaw. The most obvious solution is to include a fourth party, let's call them user D, who is known by everyone. Everyone introduces themselves to user D, their identification is checked, and their public key is saved by user D. Now when user A wants to send a message to user B, they ask user D what user B's public key is. Now that there is a known party everyone can turn to for public keys, user C cannot pretend to be anyone else (Stalling, pgs 182-189).

The method described above would work very well in an environment where there were only a few members to keep track of. However, when dealing with the Internet for instance, there are millions of servers and even more users. If each user had to ask a known party for the public key of every server, then the party referred to before as D would have millions of requests, and could not fulfill them. For this reason certificates have been developed. A certificate would most closely resemble a driver's license for a computer. It is given out by a known party (D in the example), known as a Certificate Authority. The certificate contains some information just like a driver's license. There is an expiration date, the name of the server who owns the certificate, who issued the certificate, and of course the public key of the server, which could be compared to the picture on a driver's license. In addition, a certificate is then encrypted by a Certificate Authority to generate a type of signature, proving that the certificate was made by the certificate authority. Since every user knows the Certificate Authority's public key, they can make sure the certificate is valid. In this fashion, every time a server and user establish a connection, the user can verify the server's true identity by asking for its certificate. As long as the certificate is valid and was signed by a trusted Certificate Authority, the server can be trusted to be who they say they really are. In this fashion there is no reason to connect to the Certificate Authority every time to verity a server's true identity. There are a number of Certificate authorities available on the Internet. Two of the main ones are VeriSign and GeoTrust (Thomas, Section 2.3).

3.4 Certificate Authorities

3.4.1 VeriSign

VeriSign is a company that sells certificates to large businesses. They offer businesses a service called Secure Site which is usually purchased for a one or two year period. It comes with an Authentication Service, which authenticates the business, and utilizes 128-bit SSL encryption for every SSL session. This is done by using 128-bit Global Server IDs, a form of digital identification.

VeriSign's 128-bit Global Server IDs verify a website visitor's identity and allow them to access information online. The IDs allow a website to carry out authenticated, strongly encrypted on-line commerce. VeriSign's authentication service assures customers that it is safe to submit credit card numbers and other personal information by proving that they are doing business directly with the website owner rather than some imposter's site. It also assures them that the information they are sending cannot be intercepted by anyone else.

The VeriSign Secure Site Seal, included with the Secure Site service, is displayed on a website symbolizing security and trust, allowing customers to provide credit card numbers and other private information while remaining secure. When the seal is posted on a website's home page, privacy policy page, or transaction pages, it is connected to the website owner's SSL Certificate. When visitors click on the seal, a pop-up screen appears displaying information about the SSL Certificate, authenticating that the transactions with the site are encrypted by SSL. This allows visitors to verify the website's identity.

VeriSign Payflow Pro, included with Secure Site, is an additional service which is available to website owners to handle some of the additional work associated with payment processing. Website owners are able to use a virtual terminal to handle credit card orders that are received by telephone, e-mail and fax. The virtual terminal is a secure database that stores order information as it is entered manually. Payflow Pro is ideal for businesses that require high website performance and the ability to modify all of the processes associated with online transactions (www.verisign.com).

3.4.2 GeoTrust

GeoTrust utilizes a True Site method of authentication mainly for company websites. True Site also uses 128-bit SSL encryption. It provides a way for customers to view a company's authenticated information. This can be useful for small companies who don't have a "brand name" because it lets customers know that the website is legitimate. True Site places a "smart icon" on a web page which checks to make sure it is valid by contacting a trusted third party. As the web page loads, the icon requires that the trusted third party load the image as it confirms the identity of the site owner. If the identity is confirmed the icon will load. If the identity is not confirmed the icon does not load.

A company can apply for True Site authentication through GeoTrust's website. Once the application is complete, the company's website is authenticated by an independent third party. The company's information is entered into the GeoTrust database and the site owner is then provided with the True Site "smart icon."

True Site authentication has many benefits. It confirms a company's identity without requiring a visitor to click an icon or do anything. When the smart icon appears it does not display the company's name in terms of a confusing URL. It displays the legal, recognizable company name. The smart icon, since it is loaded remotely from GeoTrust, is not an image that can be copied. Also, an entire web page can not be copied to a new site because the new site name will result in a failed lookup in GeoTrust's database. Lastly the smart icon can not be recreated because it contains a watermark that is very difficult to reproduce (www.geotrust.com).

3.5 Security Alternatives to SSL

There are many ways to add security features to web transmissions. Although SSL is the most popular, inserting a Secure Socket Layer in between the HTTP and TCP protocols is not the only way to secure Internet connections. It is possible to add security features directly to the software of the different protocol layers of the Internet architecture. Security can be added to HTTP and IP directly as an alternative to adding the Secure Socket Layer.

One of the first alternatives is an extension of the Hypertext Transfer Protocol; this has become known as S-HTTP (Secure HTTP). Unlike SSL, S-HTTP does not use one single encryption system. There is not a standard for this technology. However, it can support the public-private key system of secure web transmission. The major difference between SSL and S-HTTP is that S-HTTP allows the server to authenticate the user, while with SSL only the server can be verified. For this reason the S-HTTP feature is used instead of SSL in a case where the user has to be verified, such as a bank transaction (SearchSecurity.com). The S-HTTP standard has been published by the Internet Engineering Task Force as an experimental protocol. The reason is due to the fact that it is rare to come across a web server that supports S-HTTP. There are two reasons that S-HTTP is not as popular as SSL. The first is that this security feature is only available for HTTP; it cannot be used by other applications such as Net News Transfer Protocol (NNTP), or File Transfer Protocol (FTP). The second disadvantage is that it ties the security feature tightly to the HTTP software. Every time HTTP is upgraded S-HTTP must be taken into consideration (Thomas, Sec 1.3.2). In contrast SSL is extremely flexible, and does not have to taken into consideration when updates are released.

Security features can be added to core networking protocols also. This approach is done with IP Security (IPSEC). IPSEC has some of the same advantages as SSL, such as it being independent of the application protocol, so just about any applications can use it. Most operating applications are even completely unaware of the use of IPSEC (Thomas, Sec. 1.3.3). Although IPSEC is slightly more powerful than SSL concerning encryption, the complexity of IPSEC makes it less desirable than SSL (www.openbsd.org). This complexity causes delays in development and deployment of this security feature. Another noted disadvantage is that this feature greatly isolates security features from the operating application such as HTTP. SSL at least provides isolation from the operating application, but still allows communication between the two (Thomas, Sec. 1.3.3). Whether IPSEC will become more widely utilized over the Internet remains to be seen. Currently IPSEC's use on the web is miniscule compared to SSL.

There is another way to add security features to web operating applications. This method is called a parallel security protocol. The most popular example of this is the Kerberos software. Kerberos was developed by the Massachusetts Institute of

Technology, to provide an authentication and access type of security. Kerberos can be used by programs such as Telnet to verify a user's identity securely (Thomas, Sec. 1.3.4). Kerberos has become an advantageous alternative to firewalls in securing a network. Kerberos uses a technology method called secret-key cryptography. It is different from SSL in that it allows verification of the user by the server or client. Kerberos uses very strong encryption technology (web.mit.edu). The main drawback is that this software does not have the access to the hard data that is being transmitted while in use. Some other software must be used to provide this service (Thomas, Sec. 1.3.4). Kerberos is not an all encompassing solution.

3.6 Summary

Although there are alternatives, SSL is by far the most widely used and most popular solution for securing the transmission of personal information over the Internet. The reason for this is that SSL was introduced with the first popular web browser, Mosaic. Since SSL worked so well and is very flexible it was easy to develop further. SSL can be utilized in a number of different ways as was discussed, and doesn't interfere in any way with the normal flow of information over the internet. SSL simply encrypts the information so that it can't be read, and uses certificates to verify the identity of servers.

3.7 Interview

To make sure that this report was not missing anything major, nor misrepresenting any facts, an interview was conducted with a professional in the field. After contacting a number of individuals, we were able to conduct an interview with the

Web Developer from the Town of Manchester in Connecticut. David Yakovich's main task at the Town of Manchester is to maintain the town's homepage. He is also responsible for making sure the web servers are properly configured, secured, and have software patches applied as they become available.

Dave believes that with the currently available technology, SSL is the most reliable security feature. Any business that wants to give customers the opportunity to purchase items directly online must have an SSL certificate that is registered with a Certificate Authority. The process is not very difficult, and it takes as little as forty-eight hours to be certified with a Certificate Authority. However, the process is not as simple as buying a certificate and associating it with a website. A software package must first be run on the server that is to host the certificate, which runs a security audit. If the server has major security flaws which could potentially allow an intruder to gain access to the private key, the Certificate Authority will not issue the certificate. In general, the entire process is very straight forward, and there is no reason for a business not to invest in a registered SSL certificate.

Even though Dave believes SSL is a great piece of technology, he does not believe Internet commerce is perfectly safe or ever will be. He believes the simplest way to improve SSL is to increase the length of the encryption key that is currently used. However, SSL is not the only, nor is it the major, reason for him not to feel completely secure about online purchases. When dealing with the Internet, SSL is usually the strongest link in the chain. There is a great deal of software involved in passing web pages from servers to clients. Vulnerabilities can be found in any one of the applications used to serve content, which can be a great security risk. The only way to ensure that this does not happen is to keep servers up to date with software patches made available by software manufacturers. This is a system administrator's job, and unfortunately consumers have no way of telling if the administrator of a website is doing his or her job. Again, the issue of trust comes into perspective. To be safe on the Internet, we have to be able to trust websites with our information.

Dave believes that no matter how much is done to improve security; it is our job as consumers to make sure we are careful when we give out personal information. When making purchases online, everyone should look for reviews of the website they are buying from before making online purchases. No matter how sophisticated the security features on a website are, we must still be able to trust the individuals who maintain the website and process orders.

3.8 Survey

Our research has allowed us to understand the Internet security features available for Internet use. To bridge the connection between this technology and society, we decided to design a survey. The survey will allow us to see how often the average person encounters security features when using the Internet and how much they know about them.

From the survey we must also find out how often individuals encounter security features. This will allow us to judge whether or not the person has a good reason to know about them. If someone never makes online purchases nor uses online banking, this is a good reason for them not to know much about Internet security.

Due to location restrictions, we will only be able to survey college students in the Worcester Massachusetts area. Unfortunately, this will somewhat skew our results because college students have grown up using the Internet, and have a better understanding of it then other population groups. Also, due to the fact that there will be subjects from Worcester Polytechnic Institute, we expect a large number of respondents to be males. We also expect the numbers to be skewed due to WPI respondents having a more technical background. We understand that these results will not be representative of the overall population. We are surveying a very small group, and the demographics of this group will yield results that are slightly skewed from the general population.

4 RESULTS

Question 1

Have you ever made any purchases online or use online banking?

The survey that was conducted received one hundred twenty-seven responses. Of these responses thirty-one were discarded due to the fact that the participants indicated that they never use online purchasing and banking. We felt that including the twenty-four percent of those who do not use online purchasing and banking would skew the results. Concentrating on the remaining seventy-six percent of the population would still yield accurate results pertaining to this project.

4.1 Statistical Analysis

4.1.1 Question 2

How often do you make online purchases?

We felt this question was necessary to show that Internet commerce is used frequently by the people we surveyed. It turned out that of all the people that make online purchases, forty-five percent only make those purchases some of the time. The rest either never make purchases online or always make purchases. This distribution forms a bell curve as can be seen from Figure 4.1. The bell curve distribution makes sense due to the reality that although online shopping is convenient, it has not completely taken the place of conventional shopping. By this reasoning it would make sense that most people would make purchases using the Internet half of the time. One could extrapolate that the reason for there not being more people using the Internet for shopping could be because people don't feel completely confident in their safety.

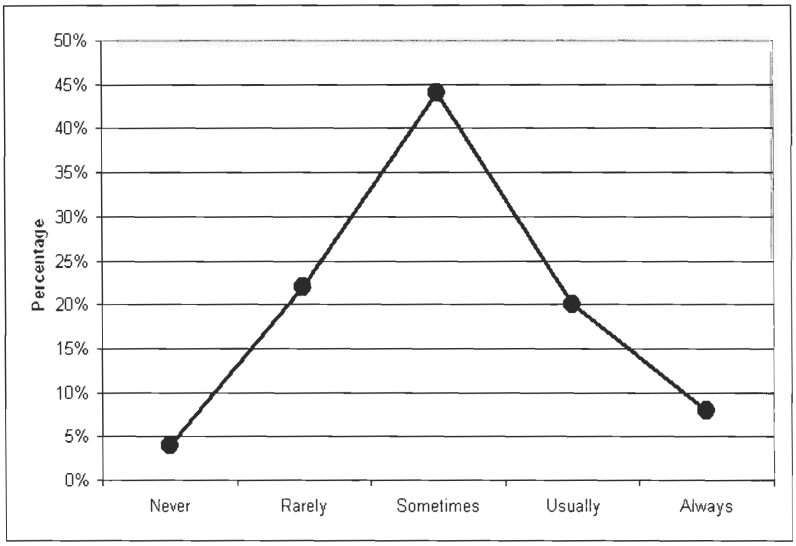

Figure 4.1: How often do you make online purchases?

4.1.2 Question 3

How often do you use online banking?

This question was also asked to determine if the people we surveyed used this Internet feature. In contrast only about thirty percent of the people surveyed do not use online banking at all. The remaining percentage is distributed fairly evenly among people who use banking rarely, sometimes, usually and always. This distribution can be seen in Figure 4.2. The distribution makes sense because more personal information is stored when doing online banking. Therefore more people would be hesitant in taking the risk that their information could be stolen.

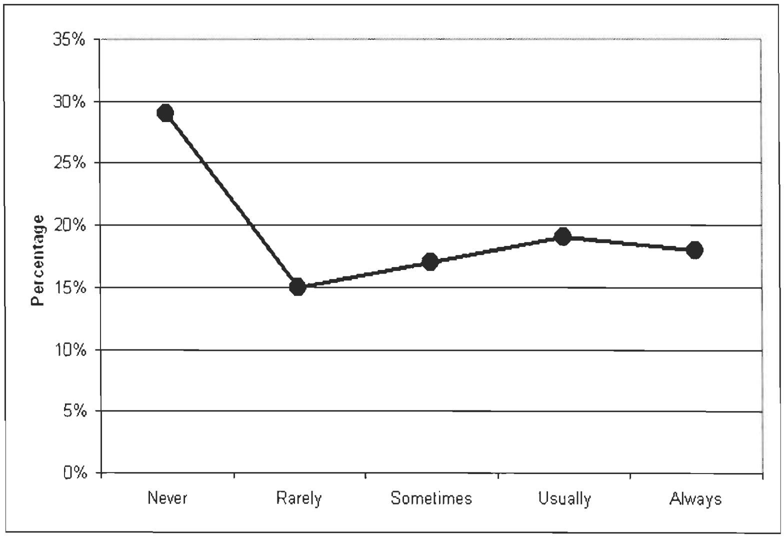

Figure 4.2: How often do you bank online?

Although people don't use online banking services often, they do make purchases online a substantial amount of the time. Due to this large use of the Internet most of the people we surveyed should be aware of the conclusion that is drawn from our research. They should be relatively aware that SSL and certificates are the features that predominantly protect their personal information. The rest of the survey will determine if the respondents actually do know this.

4.1.3 Questions 4-5

How often do you check for: the lock in the status bar/https in the website address?

These two questions in the survey addressed the issue of discovering how many people look for the actual visual signs that their network connection has been secured when they are transmitting personal information over the Internet. This was extremely important to us because it is the first part of the survey that tests how aware people are of security features. It would make sense to conclude that if they look for the two visuals in these questions, they would certainly have some sort of basic understanding of the Internet security features that our research showed people should be knowledgeable of. These two visuals were the lock symbol in the status bar of the browser and looking for HTTPS in the URL. The results yielded that over half of the participants (60%) look for the lock usually and always (Figure 4.3). Forty-two percent of the people surveyed always look for HTTPS in the URL (Figure 4.4). Although more people always look for HTTPS than the lock, more people look for the lock overall.

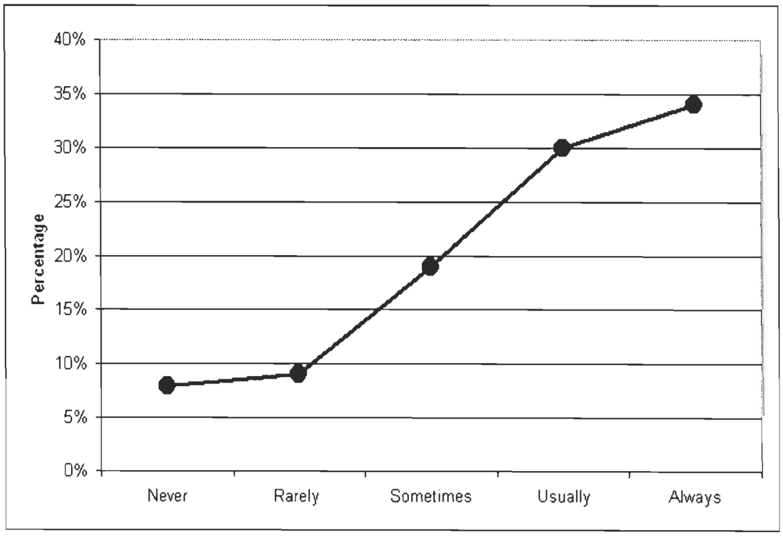

Figure 4.3: How often do you look for a lock symbol?

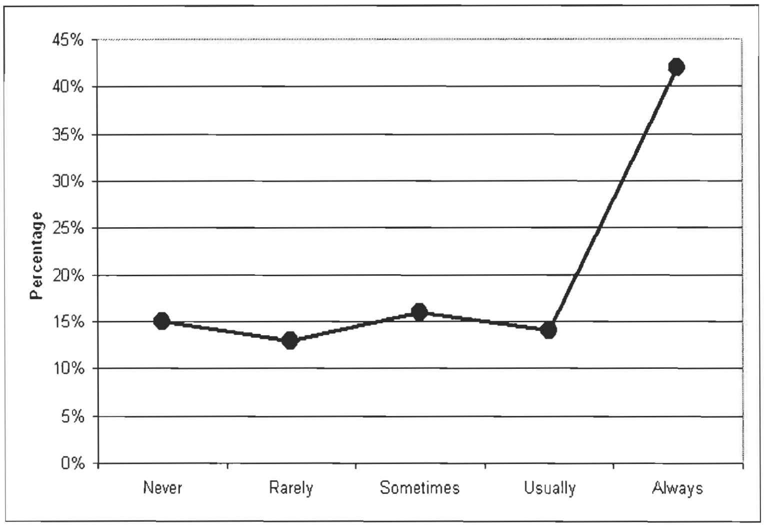

Figure 4.4: How often do you look for HTTPS?

From these results it can be seen that there is still a substantial amount of people that do not check for these two visual indicators. Luckily they are not the predominant group of people surveyed. The majority does seem to check for these indicators.

4.1.4 Questions 6-9

Recognizing Internet Security Terms:
SSL, Public/Private Key Cryptography, Certificate Authorities, CA's Names

The survey also set out to determine people's familiarity with certain Internet security features. These questions were necessary because they provided another way to determine people's knowledge concerning the most important topics we researched. The results yielded that forty-one percent know what the term SSL is while the remaining sixty percent is fairly evenly distributed among people who have some understanding to no understanding of the term. Certificate Authorities and Public/Private Cryptography were also well recognized terms. The actual names of Certificate Authorities, however, were only somewhat familiar to the majority of people (Figures 4.5 through 4.8).

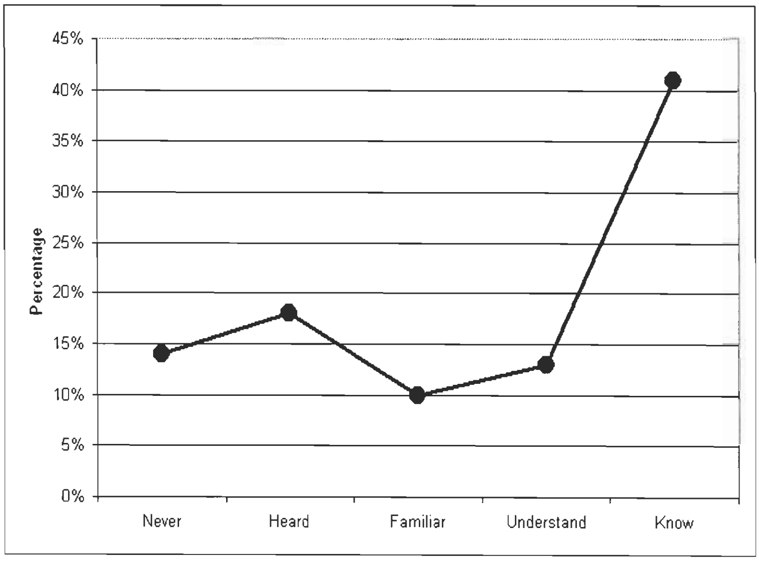

Figure 4.5: How familiar are you with SSL?

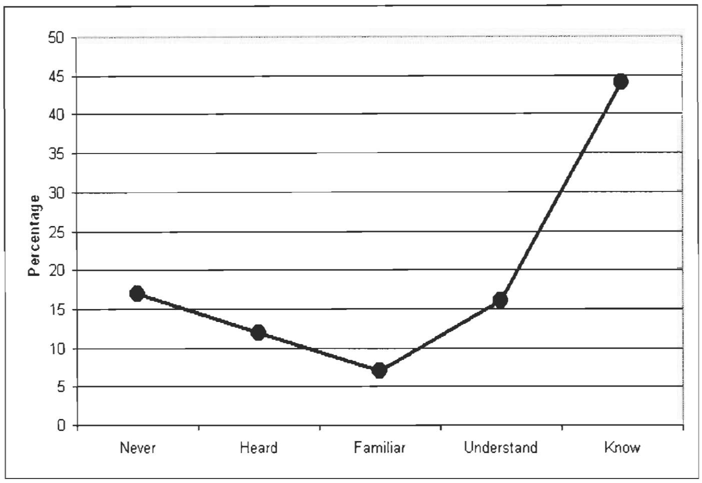

Figure 4.6: How familiar are you with Public/Private Key Cryptography?

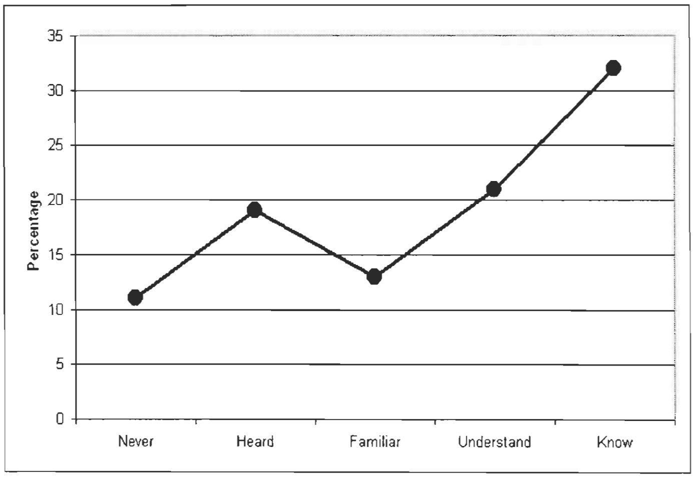

Figure 4.7: How familiar are you with Certificate Authority?

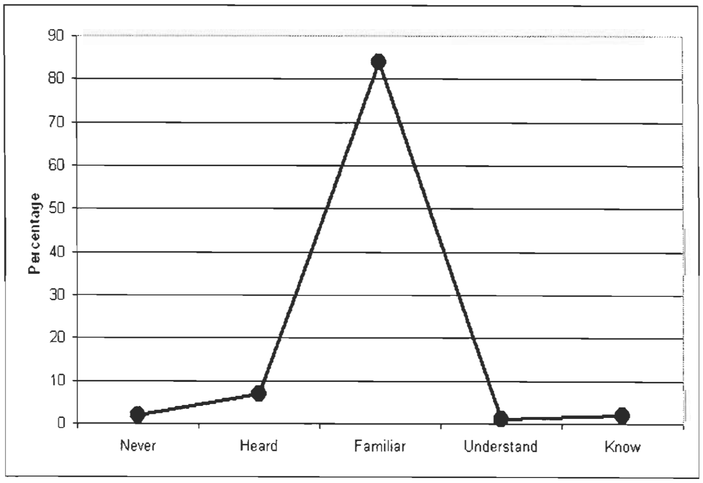

Figure 4.8: How familiar are you with VeriSign, GeoTrust, eTrust?

4.1.5 Questions 10-12

What length encryption keys are most often used by Internet browsers?
Can security services be added to the software of the networking protocol?
What does HTTPS stand for?

Self judgmental surveying yields results of what people think they know but it is also important to test their actual knowledge. Up until this point the survey has given us a general idea of people's knowledge. Questions 10-12 were the three questions that were necessary to see if the people that had detailed knowledge corresponded to the number of people that had general knowledge. The results were that two thirds of the people got the right answer for a question pertaining to encryption key length. This response was probably high because a large portion of those surveyed have technical backgrounds. Forty-six percent answered correctly to a question pertaining to HTTPS. This shows a fairly even distribution. Eighty percent properly answered a question relating to other security features besides SSL. The eighty percent is most likely too high due to the fact that this was a yes or no question. In other words the question was too easy, and the correct answer was easy to guess. The distributions for these three questions can be seen in Figures 4.9 through 4.11.

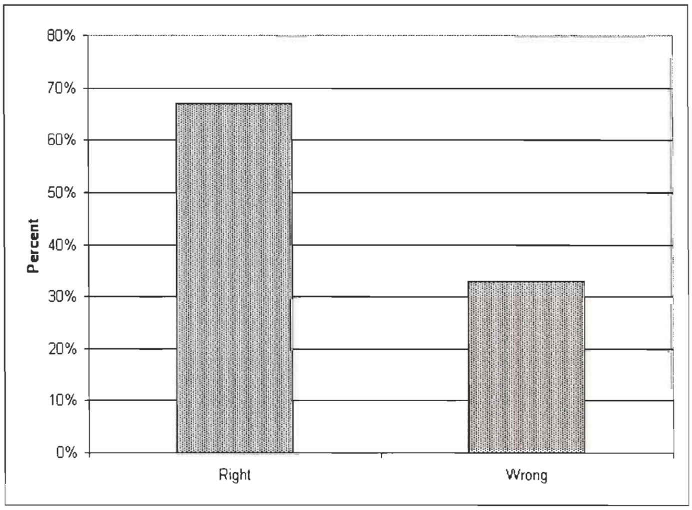

Figure 4.9: What length encryption keys are most often used?

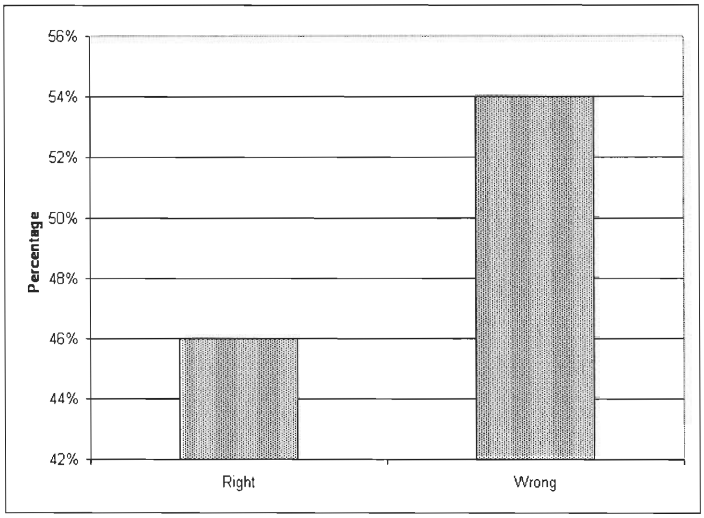

Figure 4.10: What does HTTPS stand for?

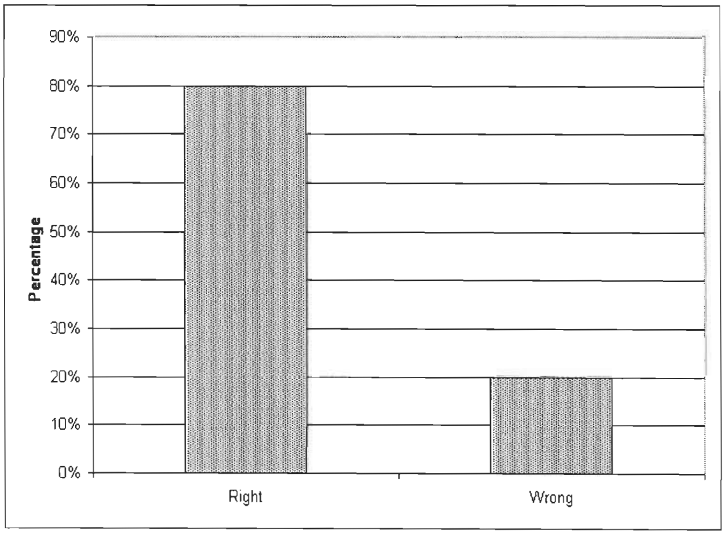

Figure 4.11: Can encryption be added to software?

4.1.6 Question 13

Rate your understanding of Internet security features.

When asked to rate their knowledge pertaining to security features most respondents felt that they had adequate knowledge. This question served the purpose of determining how confident the participants were with their responses (Figure 4.12). Looking back at the other results it seems the people we surveyed were over confident concerning their knowledge. Over half the respondents said they had an adequate or better than adequate understanding of Internet security features, while it turned out our questions showed that only half the people actually do have a good understanding of the pertinent Internet security topics found in our research. This over confidence could be due to the technical background that a lot of respondents had, and that some of them just assumed that they knew what they were talking about.

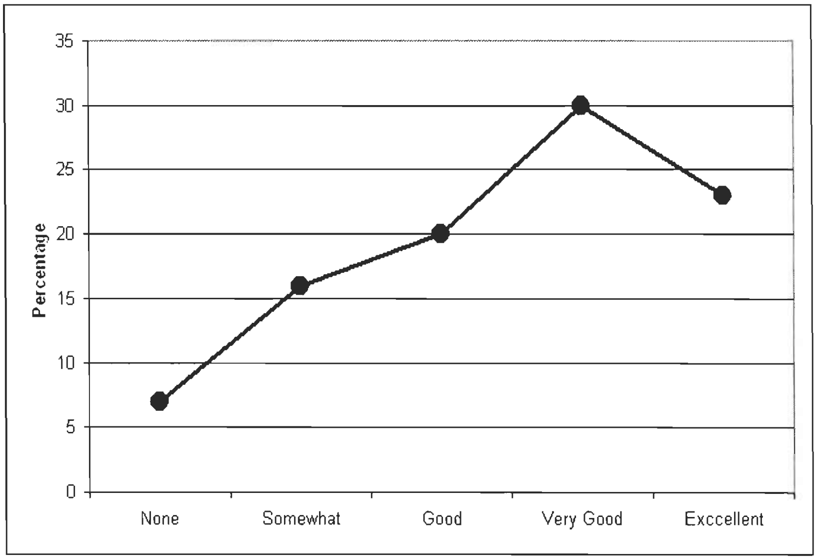

Figure 4.12: Rate your understanding of Internet security features?

Overall the results yielded that half the population surveyed were knowledgeable in the area of Internet security. It should be kept in mind that these results are slightly skewed due to the technical background of some of the respondents. Section 5.4 contains final conclusions based on these results.

5 CONCLUSION AND RECOMMENDATIONS

This project set out to gain an understanding of what internet security features the average person might encounter while browsing the internet, and also to discover how knowledgeable people are concerning these features.

5.1 Purpose and Methods

The internet has become popular and widely used, making security a concern for all users. Users should know that there are security features built into web browsers to keep personal information from being intercepted during transactions. It was vital to the project to first explore the details of these features. This was done by reading books and online articles discussing internet authentication and encryption. Once that was accomplished we administered a survey to over a hundred college students, to discover people's understanding of internet security features.

5.2 Research Findings

Through the research we found that the most widely used internet security feature is the Secure Socket Layer. SSL is very flexible and can be used not just for encrypting HTTP data, but many other internet based communications. The most important component of SSL is the use of Public-Private key cryptography, which allows users to establish secure connections to internet servers, as well as provide a way of verifying the identity of servers. There are other ways of encrypting internet communications, which not only allow for a way to identify servers, but also the users. This was discussed in section 3.5. However they are not very widely used at this time.

5.3 Interview Conclusions

In order to validate our research an expert was interviewed about internet security. He agreed that SSL is currently the best way to secure internet communication. However, there are always ways to improve, such as increasing the length of the encryption key used in SSL. Besides improving crypto graphical techniques, the best way for a user to stay safe is to become aware of the technology around them, and always be weary when giving out personal information.

5.4 Survey Results

From the results of the survey, it was found that half of our participants were knowledgeable about web browser security. This was a larger number than we expected to find. It was also found that individuals who made more online purchases were more aware of the technology securing their privacy. From this we believe that the more knowledgeable an individual is about the features available to them, the more comfortable they feel making online purchases. Therefore, educating consumers of the technology available to them could potentially increase internet banking and other commerce.

5.5 Significance and Recommendations

The conclusion that can be seen from half of the survey respondents being knowledgeable would be that there are potentially a larger number of people that can become educated and therefore feel comfortable engaging in online commerce. The simplest and fastest way for anyone to educate themselves and feel more comfortable when making online transactions, is to simply look for materials available on websites. Any legitimate business should have a link entitled "Privacy" or "Security" somewhere on their website. This can provide consumers with the precautions that companies take to ensure that their customers' information is both transmitted and held in a safe fashion.

Companies should also do their best to make this type of information readily available. Individuals who are careful about their finances and security should look for links relating to privacy and security on websites where they plan to make purchases. If web developers make information more readily available, then consumers will be more trusting towards those companies. These suggestions can create a much more secure and trusting community on the internet, improving the relationship between consumers and companies, allowing for economic growth.

5.6 Broad Discussion

Through our research we have found that, if caution is taken, using online commerce can be relatively secure. However, as David Yakovich mentioned, Internet communications will never be completely secure. What does it really mean to be safe? Can we ever feel truly safe transmitting personal information over the Internet? Will there come a time when a network can be considered completely secure? Most are hopeful that technology is progressing in that direction. It is probable that Internet commerce could reduce the number of small shops, but we do not believe stores as we know them today will ever become extinct. There will always be a need for people to touch and try on items before making purchases. Internet commerce has had a profound influence on society, but it cannot replace the enormous commercial infrastructure that we have today.

For online commerce to become an integral part of this infrastructure a majority of people have to be well informed of web browser security. If more people become educated we can expect that, with proper use, online shopping can become just as secure as making purchases in person.

BIBLIOGRAPHY

GeoTrust, Inc. 1 April 2003 <http://www.geotrust.com/index_flash.htm>

Gourley, David. *HTTP: The Definitive Guide.* Sebastopol: O'Reilly & Associates, Inc., 2002. Secs. 1.1-14.1. (Electronic Book)

Heaton, Jeff. *Programming Spiders, Bots, and Aggregators in Java.* Alameda: SYBEX Inc., 2002. Chapter 3. (Electronic Book)

Hobbes Internet Timeline v6.0, [On-Line Archive]. Available from http://www.zakon.org/robert/Internet/timeline/; Internet.

Internet Society, *A Brief History of the Internet and Related Networks*, [On-Line Archive]. Available from http://www.isoc.org/Internet/history/cerf.shtml; Internet.

Kerberos: The Network Authentication Protocol. [On-Line] Available at http://web.mit.edu/kerberos/www/#what_is; Internet.

Open BSD. *Using IPsec (Internet Protocol Security).* [On-Line Archive] Available at http://www.openbsd.org/faq/faq13.html#What; Internet.

SearchSecurity.com *S-HTTP* [On-Line] Available at http://searchsecurity.techtarget.com/sDefinition/0,,sid14_gci214580,00.html; Internet.

Singh, Simon. *The Code Book, The Evolution of Secrecy from Mary Queen of Scots to Quantum Cryptography.* 1999 Random House, Inc.

Stallings William. *Cryptography and Network Security Principles and Practice, Second Edition.* 1999, 1995 Prentice-Hall, Inc.

Thomas, Stephen A. *SSL and TSL Essentials.* New York: John Wiley and Sons, 2000. Secs. 1.2-1.3.2. (Electronic Book)

VeriSign, Inc. 1 April 2003 <http://www.verisign.com/>

Yakovich, David. Personal Interview. 7 April 2003.