Eng. & Tech. Journal , Vol.29, No. 14, 2011

Website Anti-Phishing Technique Using Randomized Dyadic Dilation DWT

Mohammed Gheni Alwan*, Khlood I. Abbas**
&Mohammed lbrahim Ahmed***

Received on: 2/2/2011
Accepted on: 21/7/2011

Abstract

Website phishing is a growing threat on the internet and its effect is devastating when it comes to phishing financial web sites, such as eBay, official bank accounts or other critical sources.

It is an easy way to fraud a web site source code and even a whole website can be totally downloaded using free software packages, clients do not know for sure if they talk to a genuine or fake websites by counting on what it looks like and its components.

This paper is proposed to use randomized dyadic dilation wavelet to enhance the authenticate designated website to the visitor and in the same time authenticate the visitor to the website. This technique mainly designed to defeat Man-In-Middle attack which is the main threat source for website phishing due to the difficulties of getting the transformed image reconstructed easily.

تقنية مكافحة تزوير المواقع الالكترونية باستخدام التحويلات الموجية المتقطعة المعتمدة على الازاحة العشوائية

الخلاصة

ان عملية تزوير وتقليد الموقع الالكتروني اصبحت من الاخطار المتنامية على الانترنيت ولها من التأثيرات المدمرة خاصة عندما يكون هذا التأثير على مواقع الكترونية خاصة بالامور المالية مثل البنوك ومواقع الدفع الالكتروني وغيرها من المواقع الحكومية الحساسة.

انه من اليسير ان تتم عملية تزوير موقع الكتروني واستخدام الشفرة الخاصة به كونها من النوع المفتوح والمتاح الى الجميع، بالاضافة انه من الامكان تنزيل موقع كامل بكل برامجياته باستخدام حقائب برمجية مجانية موجودة على الانترنيت.

هذا البحث المقدم يقترح استخدام اسلوب المعالجة الصورية باستخدام طريقة التحويل الصوري المعتمد على الازاحة العشوائية لتحسين عملية تحقيق هوية الزائر الى الموقع الالكتروني وفي نفس الوقت تحقيق هوية الموقع الالكتروني الى الزائر، ان الطريقة المقترحة مصممة خصيصا لافشال المهاجمة المعتمدة على اسلوب الجلوس في الوسط والتي تعتبر الخطر الاكبر لتزوير المواقع الالكترونية.

1- Introduction

In the field of computer security, phishing is the criminally fraudulent process of attempting to acquire sensitive information such as usernames, passwords and credit card details By masquerading as a trustworthy entity in an electronic communication. Communications purporting to be from popular social web sites, auction sites, online payment processors or IT

* Computer Science Department ,University of Technology
** Science College, university Al-mustansiriyah /Baghdad
***Ministry of Higher Education and Scientific Research / Baghdad

https://doi.org/10.30684/etj.29.14.3
2412-0758/University of Technology-Iraq, Baghdad, Iraq

administrators are commonly used to lure the unsuspecting public. [1]

The term was coined in the 1996 timeframe by hackers who were stealing America Online (AOL) accounts by scamming passwords from unsuspecting AOL users. The popularized first mention on the Internet of phishing was made in alt.2600 hacker newsgroup in January 1996, however the term may have been used even earlier in the popular hacker newsletter “2600”.[2,3]

By 1996, hacked accounts were called "phish", and by 1997 phish were actively being traded between hackers as a form of electronic currency. There are instances whereby Phishes would routinely trade 10 working AOL phish for a piece of hacking software or warez (stolen copyrighted applications and games). The earliest media citation referring to phishing wasn’t made until March 1997.[2] Phishing is typically carried out by e-mail or instant messaging, and it often directs users to enter details at a fake website whose look and feel are almost identical to the legitimate one. Phishing is an example of social engineering techniques used to fool users,] and exploits the poor usability of current web security technologies. Attempts to deal with the growing number of reported phishing incidents include legislation, user training, public awareness, and technical security measures.[3]

Over time, the definition of what constitutes a phishing attack has blurred and expanded. The term Phishing covers not only obtaining user account details, but now includes access to all personal and financial data. What originally entailed tricking users into replying to emails for passwords and credit card details, has now expanded into fake websites, installation of Trojan horse key-loggers and screencaptures, and man-in-the-middle data proxies – delivered through any electronic communication channel.[2,3]

The proposed scheme in this paper is an enhancement to website authentication in order to prevent Man-in-Middle attacker from hijacking sensitive information passed by clients to fake websites.

2- Phishing a Website

The fraudulent web site that supports the phishing email is designed to mirror the legitimate web site it is purporting to be. The fraudsters use multiple methods to do this, including using genuine looking images and text, disguising the URL in the address bar or removing the address bar altogether. The purpose of the web site is to trick consumers into thinking they are at the company's genuine web site, and giving their personal information to the trusted company they think they are dealing with. [2,3]

Once a victim visits the phishing website the deception is not over. Some phishing scams use JavaScript commands in order to alter the address bar. This is done either by placing a picture of a legitimate URL over the address bar, or by closing the original address bar and opening a new one with the legitimate URL.[2]

An attacker can even use flaws in a trusted website's own scripts against the victim. These types of attacks (known as cross-site scripting attack) are particularly problematic, because they direct the user to sign in at their bank or service's own web page, where everything from the web address to the security certificates appears correct. In reality, the link to the website is crafted to carry out the attack, making it very difficult to spot without specialist knowledge. Just such a flaw was used in 2006 against PayPal. [2,4]

A Universal Man-in-the-middle (MITM) Phishing Kit, discovered in 2007, provides a simple-to-use interface that allows a phisher to convincingly reproduce websites and capture log-in details entered at the fake site as it can seen in figure1, the attacker (criminal) is publishing the fake web site to the users. [2,3]

For man-in-the-middle attacks to be successful, the attacker must be able to direct the customer to their proxy server instead of the real server. This may be carried out through a number of methods:[1,5]

- Transparent Proxies
- DNS Cache Poisoning
- URL Obfuscation
- Browser Proxy Configuration
- Phishing eBay Website Scheme

The following is a real case taken as a demonstration on phishing a website, this hit was really conducted on the targeted famous site 'www.eBay.com', the attack was detected and contained with a minimum collateral damages. [5]

Two critical pieces of information were targeted in this scheme: the authentication credentials (i.e., username and password) and the user’s credit card information. Figure2 shows the critical steps of the scheme from beginning to end.[5]

To build the fraudulent web site, the attacker(Man-in-Middle) simply sends requests to eBay for the HTML markup and images needed to render critical pages of the eBay site. Because the Web works by having clients (such as Mozilla or Internet Explorer) download HTML from the server and then display the results to the user, there is no way for eBay to stop users from downloading its source. In fact, easy replicability of content from . Instructing the eBay site to send a copy of the source is as simple as having the attacker point his browser to http://www.ebay.com.

Each of the thirteen steps identified in figure2 supports one of three goals needed for the thief to achieve his objective. Those goals are creation of the fraudulent eBay site, directing users to the fraudulent site, and then operating the fraudulent site such that users never suspect what has happened.[5]

3- AugmentedPasswords Logins

The Bank of America's website is one of several that ask users to select a personal image, and display this user-selected image with any forms that request a password. Users of the bank's online services are instructed to enter a password only when they see the image they selected. However, a recent study suggests few users refrain from entering their password when images are absent. In addition, this feature (like other forms of two-factor authentication) is susceptible to other attacks, such as those suffered by Scandinavian bank Nordea in late 2005, and Citibank in 2006.[4]

A similar system, in which an automatically-generated "Identity Cue" consisting of a colored word within a colored box is displayed to each website user, is in use at other financial institutions, as in figure(3).[4]

Security skins are a related technique that involves overlaying a user-selected image onto the login form as a visual cue that the form is legitimate. Unlike the website-based image schemes, however, the image itself is shared only between the user and the browser, and not between the user and the website. The scheme also relies on a mutual authentication protocol, which makes it less vulnerable to attacks that affect user-only authentication schemes.[6,4]

Still another technique relies on a dynamic grid of images that is different for each login attempt. The user must identify the pictures that fit their pre-chosen categories (such as dogs, cars and flowers). Only after they have correctly identified the pictures that fit their categories are they allowed to enter their alphanumeric password to complete the login. Unlike the static images used on the Bank of America website, a dynamic image-based authentication method creates a one-time passcode for the login, requires active participation from the user, and is very difficult for a phishing website to correctly replicate because it would need to display a different grid of randomly generated images that includes the user's secret categories.[5]

4- Wavelet Transform

A wave is usually defined as an oscillating function of time or space. Wavelet is a "small wave" which has its energy concentrated in time to give a tool for the analysis of transient, non-stationary, or time-varying phenomena. Wavelet transform is a transform that provides the time-frequency representation which is capable of providing the time and frequency information simultaneously [7]

$$[W_\psi f](a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \psi\left(\frac{x-b}{a}\right) f(x)\,dx \quad \text{- (1)}$$

The wavelet coefficients are expressed by the following eq.

$$c_{jk} = [W_\psi f]\left(2^{-j}, k2^{-j}\right) \text{---- (2)}$$

Here, a = 2^{-j} is called the binary dilation or dyadic dilation and (i.e., stretching or scaling factor) b = k2^{-j} is the binary or dyadic position (i.e., shifting factor)

Simply, it can be said that wavelet transform for a certain image is a correlation and similarity matching with a small wave scaled and shifted along the image, thus the following equation is satisfied: [7]

$$C(stretching, shifting) = \sum x_i \psi(stretching, shifting) \quad \text{---------} (3)$$

Wavelets come in various shapes and sizes as in figure (4). By stretching and shifting ("dilating and translating") the wavelet we can "match" it to the hidden event and thus discover its frequency and location in time. Stretching or shifting by powers of 2 is often referred to as "dyadic". For example, dyadic dilation means stretching (or shrinking) by factors of 2 (e.g. 2, 4, 8, 16 etc). [7] With scale varying along dyadic sequence (2^j), j ϵ Z, the support of wavelet base $\Psi_j(x)$ (will increase rapidly.[8] Shifting or sliding is referred as translating in wavelet terminology, with their ability to stretch and shift, wavelets are very adaptable, even new wavelets can be constructed to find more new details. [7]

The generalized equation (equ.3) for CWT is a shortcut that

$$[W_\psi f](a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x-b}{a}\right)} f(x)\,dx$$

shows that the correlation coefficients depend on both the stretching and the shifting of the wavelet, Ψ, to match the signal (X_n here) as we have just seen. The equation shows that when the "dilated and translated" wavelet matches the signal the summation will produce a large correlation value.[7,5]

Randomized Dyadic Dilation DWT

This paper dose not propose an image processing algorithm to get enhanced capturing of more detailed information from the mother wave (i.e., original image called mother wave). It is well known that wavelets are so many and each is dedicated to collect different feature from the mother wave, what is important here is the ability to scale (dyadic dilation) and shift (dyadic position), which is clearly announced in (equ.3). [7]

This paper is proposing a random dyadic dilation when executing wavelet transformation to the original image, the resultant of this approach could produce worse in term of transformation resolution (i.e., many details could be slipped a way) which makes the construction of the original image harder and harder.

$$C_{ij}\ (K, shifting) = \sum X_n\ \Psi\ (K, shifting) \quad \text{---------} (4)$$

Where K: random variable (i.e., power of 2)

C_{ij}:**Transformation co-efficient**

This transformation represented by equ.4 brings new criteria regarding the construction back of the image, where it is not possible to reconstruct the original image without knowing the pattern, through which, K is changing its values.

This paper will use pseudo-random generator to generate semi-random values for K and use these values in wavelet compress-coding certain images. These images are not easily re-constructed without having K generator and it tends to be a very sophisticated and complex issue if the period of the pseudo-random generator is bigger than image being wavelet compressed and coded.

LFSR is selected to be the random generator with a maximum period (2^n-1) equal or bigger than (logN) where N is the size of the image being transformed and (n is the length of the register).

5- The proposed Mutual Authentication Based RDWT

The proposed authentication scheme by this paper is relying very much on the manipulating of the independent variable K presented in (equ.4), which representing shifting of certain wavelet to find correlation coefficients with the mother signal (i.e., the original image).

By stretching and shifting the wavelet numerous times we get numerous correlations. If the mother signal (i.e., image) has some interesting events embedded, the transformation will get the best correlation when the stretched wavelet is similar in frequency to the event and is shifted to line up with it in time as in figure (5). Wavelet transformation algorithm selects K to have values to be factors of 2 to be the right dyadic dilation for better transformation.[7]

In this proposed authentication scheme, the resolution of the image reconstructed from wavelet coefficients is not the important issue here due to the concentration is upon the security not the performance of the DWT transformation, which is the case in image applications.

The following flowchart shows this operation in detailed session. This session should be at the trusted level between client and the server, due to the infromation at this level should be confedintial and not exposed to the public because it is a personal informaion and should be kept a way of others.

6- Conclusions

1- Website phishing is a crucial factor to bring confidence to customer before conducting any transaction over the internet. The traditional methods of security (login, password) now became obsolete toward website phishing.

2- Website phishing can impose more dangerous actions in local by conducting ARP spoofing and direct DNS to attacker machine.

3- Visual interactions with clients are more reliable, in security point of view, than the traditional login techniques.

4- Using visual profile for the client and deploy it later for authentication purposes is not enough without embedding techniques emerging from cryptography.

7- References

[1]Koon Yaw Tan,"phishing and Spamming via IM ", 2006.

[2]http://www.fraudwatchinternational.com/phishing-fraud/phishing-web-site-methods.html

[3]http://www.technicalinfo.net/papers/Phishing.html

[4]http://en.wikipedia.org/wiki/phishing

[5]Matt Curtin, "Anatomy of Online Fraud", interhack,2003

[6]http://www.wordspy.com/words/phishing.asp

[7]D.Lee Fugal," Conceptual Wavelets in Digital Signal Processing", Space and Signals Technologies LLC,2009.

[8]Lei Zhang and Paul Bao, “Edge detection by scale multiplication in wavelet domain”,2002

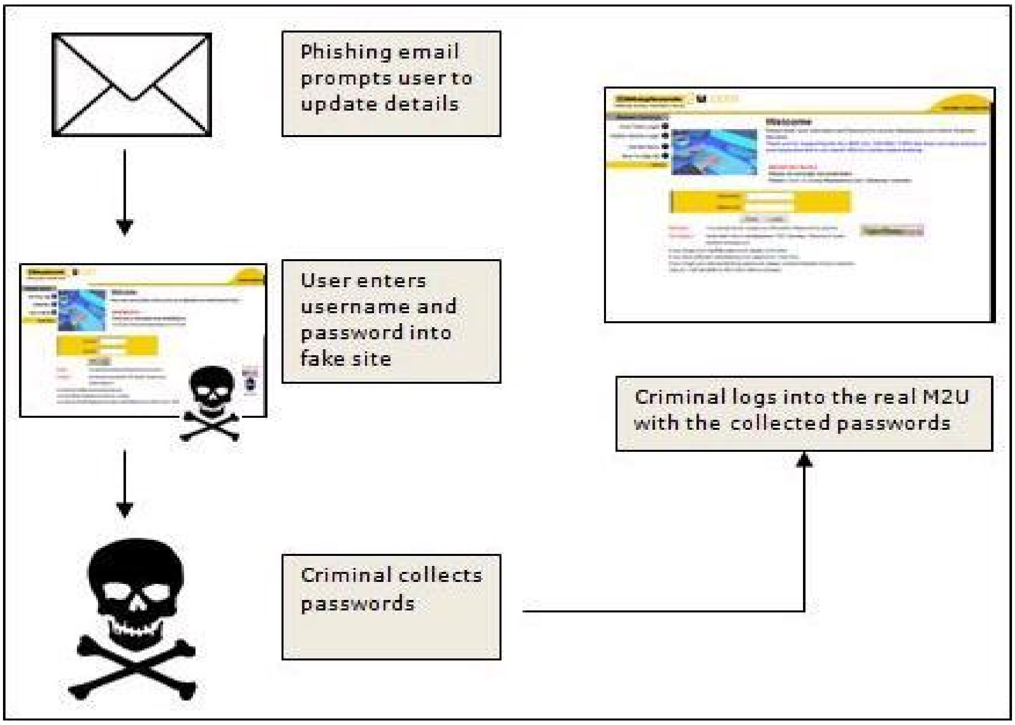
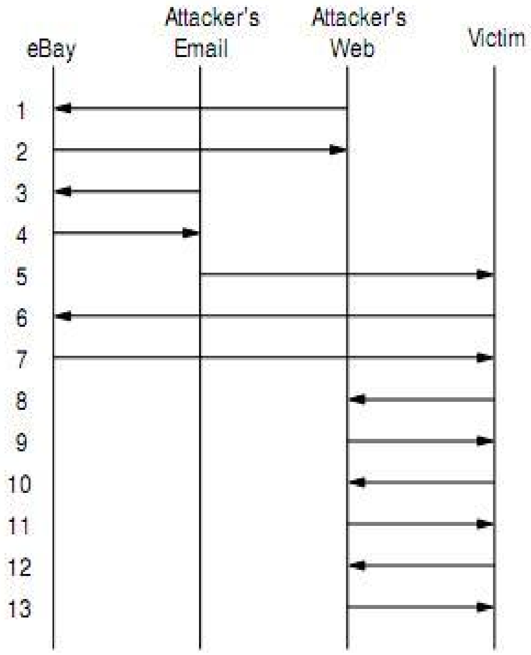

Phishing email prompts user to update details
User enters username and password into fake site
Criminal collects passwords
Criminal logs into the real M2U with the collected passwords

Figure 1: fraudulent website phishing process



Attacker sends request to eBay for source eBay web server gives source Attacker pulls additional source from eBay web server gives source ‘‘Please resubmit credit card’’ Mail client requests real eBay images eBay web site delivers images as requested Victim clicks on a link, thinking it is to eBay Attacker’s web site display, looks like eBay User puts in username and password Attacker accepts password, asks for credit card User uploads credit card information Attacker thanks victim for the ‘‘update’’

Figure 2: Interaction Diagram Showing Scheme

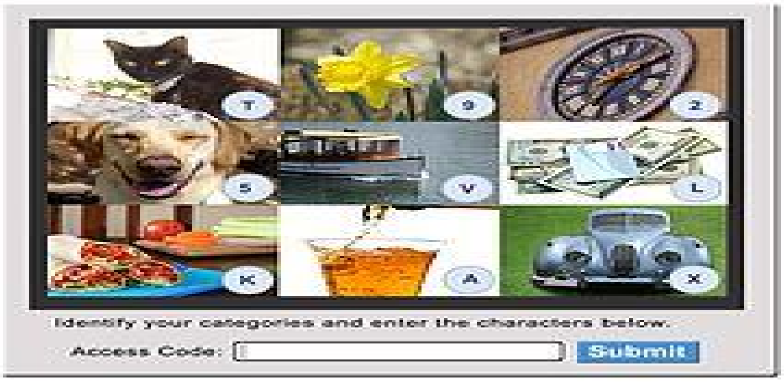


Figure 3: Dynamic image-based authentication for anti-phishing

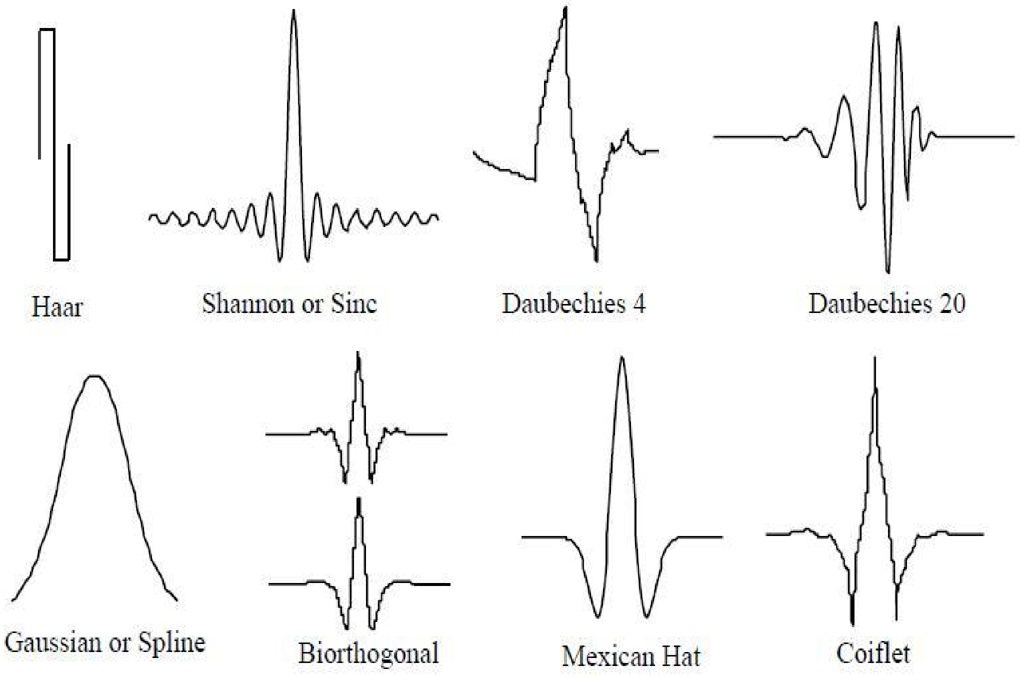


Figure 4: Different Types of Wavelet Signals

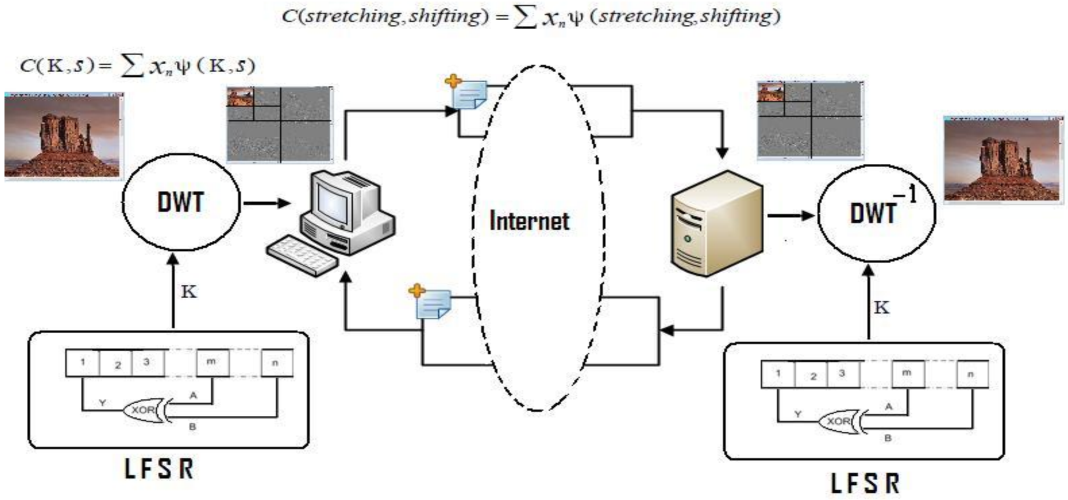

$C(stretching, shifting) = \sum x_n \psi\,(stretching, shifting)$
$C(K, s) = \sum x_n \psi\,(K, s)$
DWT
K
Internet
DWT^{-1}
K
1
2
3
m
n
Y
XOR
A
B
LFSR
1
2
3
m
n
Y
XOR
A
B
LFSR

Figure 5: Randomized Dyadic Dilation DWT based Authentication Scheme

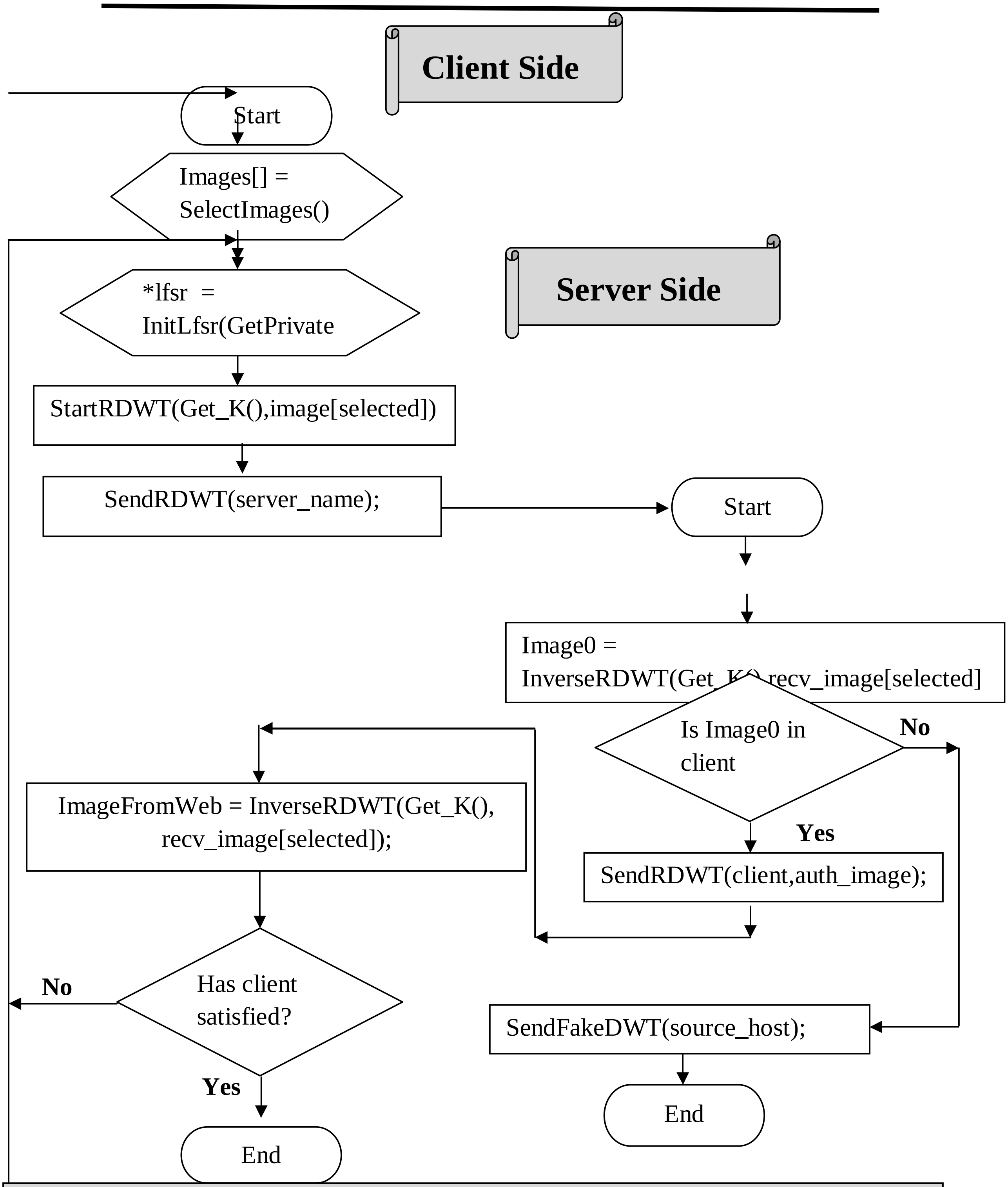


Figure 6: The Proposed Mutual Authentication Procedure based on RDWT