Kendal, Simon (2013) Selected Computing Research Papers Volume 2 June 2013. Selected Computing Research Papers . University of Sunderland, Sunderland.

Downloaded from: http://sure.sunderland.ac.uk/id/eprint/9587/

Usage guidelines

Please refer to the usage guidelines at http://sure.sunderland.ac.uk/policies.html or alternatively contact sure@sunderland.ac.uk.

Selected Computing Research Papers

Volume 2

June 2013

Dr. S. Kendal (editor)

Published by
the
University of Sunderland

The publisher endeavours to ensure that all its materials are free from bias or discrimination on grounds of religious or political belief, gender, race or physical ability.

This material is copyright of the University of Sunderland and infringement of copyright laws will result in legal proceedings.

© University of Sunderland 2013

Authors of papers enclosed here are required to acknowledge all copyright material but if any have been inadvertently overlooked, the University of Sunderland Press will be pleased to make the necessary arrangements at the first opportunity.

Edited, typeset and printed by
Dr. S Kendal
University of Sunderland
David Goldman Informatics Centre
St Peters Campus
Sunderland
SR6 0DD

Tel: +44 191 515 2756
Fax: +44 191 515 2781

Contents Page

An Evaluation of Current Innovations for Solving Hard Disk Drive Vibration Problems

Isiaq Adeola

Abstract

Vibration is a very huge factor affecting hard disk drives (HDD) due to the rotational movement of the spinning disk during usage. This paper describes an evaluation of current innovations to mitigating HDD vibration problem and also takes note of the methodology used to back up the claims made. Methods are then compared so as to recommend most effective solution to solving HDD vibration problem from both internal and external stand point.

1. Introduction

Vibration is a very big disadvantage in hard disk drives (HDD) due to the rotational movement of the platters (spinning disk). HDD are even subjected to more rigorous, harsh and more intense conditions in present days, users demand for even faster and high performance HDD have also lead developers into pushing the boundaries of tackling the vibration issues with the HDD. Kirpekar and Bogy (2008) explained that the need for faster data transfer rates has led to an increase in rotational speed of the spinning disk in HDD, however it isn't the dynamic stability of the HDD that is the only concern, vital and delicate parts of most modern day HDD such as the Read-Write head are positioned only inches away from the rotating disks by means of air-bearing sliders which are detrimental to the accurate positioning of the read–write head on data tracks. Shen (2000) proposed developing alternative substrate materials for the disks such as ceramic with extremely high stiffness to reduce the vibration amplitude and also increase resonance to higher frequency. Kameya et. al. (2012) also proposed the use of laminated disks consisting of a viscoelastic polymer core sandwiched between two aluminum disks, during rotation the viscoelastic polymer undergoes a distributed form of deformation because the stiffness of the polymer is much less than that of the aluminum disks. Further research has also being done towards vibration problems such as Jintanawan (2009) explaining hard disk vibration energy being converted to acoustic noise and the thorough understanding of such noise generating mechanism can improve in the design for better HDD. Sibielak (2012) also claimed using piezoelectric elements which results in high resolution rotation causing no friction.

In this paper I will critically analyze various academic journals focusing on hard disk vibrational problems, methods that have been proposed towards mitigating the vibration issues and also experiments used to back up their claims and how their experiments are being evaluated in order to achieve the most effective solution.

2. Evaluating Researches On Vibration Problem

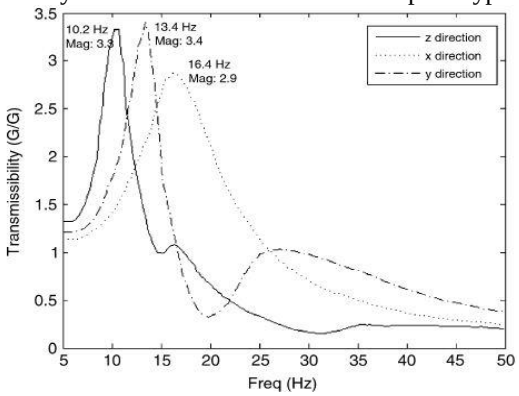
The rotation of the spinning disk in a HDD is the root cause of the internal vibration problems of HDD. Horowitz et. al. (2007) proposed two innovations to be implemented in the near future in order to increase the storage density of HDD, the first is the use of high bandwidth dual-stage actuator servo system so as to improve the precision and track-following capability of the read/write head of the HDD and secondly The second is the instrumentation of the HDD suspensions with vibration strain gage sensors, so as to improve airflow induced suspension vibration suppression in the HDD. Majority of the researches on HDD vibration problems are focused on internal vibrational problems and little

research is being made towards external vibration problems which are even more unpredictable because of the evolution of technology. Appliances are now being made portable and taken into extremely rugged environments, external forces can also be a big problem to most of the best HDD vibration problems out there. Yap et.al. (2006) explained the importance of an external vibration isolation system of an HDD as not being stressed enough because HDD are being used under even more rugged circumstance in this new age, their procedures for testing vibration and their means of evaluation where documented and presented. Hoque et.al. (2010) claimed that passive techniques such as passive damping and isolation are mostly used in many applications but are extremely limited because the isolation elements need to be soft enough to absorb vibration and stiff enough to prevent on-board generated direct disturbances. Choi and Yoon (2011) also claimed that applying herringbone groove pattern to the dampers of an HDD can reduce the axial vibration of the disk rotating at high speed.

Yap et.al. (2006) also claimed that the best standard of vibration isolation should be between 10Hz to 20Hz and their prototype fulfills all the necessary criteria to be confirmed successful. This research paper is focusing on external HDD vibration isolation which also included the military standard 810E method. They initially experimented on a HDD without vibration isolation in order to get an acceptable limit, the purpose of this experiment to ensure the acceleration response of the vibration isolation prototype is lower than the acceptable limits. Two experiments were conducted in this research which will be analyzed below to see what kind of experiments was conducted and how they were evaluated to back up their claims.

Experiment1 where Yap et.al. (2006) developed an external vibration isolation prototype and conducted an experiment known as the vibration testing and evaluation criteria, in this case the HDD was tested using random-on-random vibration signals according to military standard 810E. Vibration signals are generated to the controller, by this way the controller is connected to an electrodynamics shaker to physically produce the random-on-random vibration to the prototype analysis was made while this experiments was undertaken and data was taken as results and then graphed and was also evaluated in two different ways so as to confirm consistency. Firstly Yap et.al. (2006) evaluated the first experiments by streaming movie data (audio and video) from the hard disk drive during the vibration test, the evaluation is said to be successful if there are no pauses in the movie while being played continuously during the vibration test. Secondly Yap et.al. (2006) also evaluated the first experiment by recording real time data to the HDD during vibration test with constant transfer rate, this evaluation is said to be successful if the desired data recording rate can be maintained without loss.

The diagram below is a graph showing the measured transfer of vibration from the vibration profile to the vibration isolation prototype in x, y and z direction. Yap et.al. (2006) claimed that the natural frequencies and vibration isolation are within the range predicted to give satisfactory vibration isolation to the prototype.



Measured transmissibilities of isolator system in x, y and z directions (Yap et.al. 2006)

Experiment2 was a computer simulation, Yap et.al. (2006) developed a computer simulation model for the HDD vibration isolation model using SimMechanics which takes input in form of vibration time history data generated from the MIL-STD 810E PSD profiles and predicts the output response.

Going through this methodology, it is observed that two different means of evaluating the experiment was adopted, this shows consistency in their result, the experiment was also controlled which eliminates other unrequired factors such as unbiased input. They also conducted an experiment using an HDD without a vibration isola-

tion system incorporated to set a benchmark. Progress could easily be recorded using this benchmark result. The methodology was also properly documented in order for repeatability by other researchers. Experiment was also conducted repeatedly with different parameters so as to double-check their result wasn't random. This confirms the reliability of the experiment.

Looking at experiment2, computer simulation is also a controlled form of experiment, and even if it's entirely different from experiment1 in form of application, the results obtained from both experiments are almost alike. The prototype developed has not yet been tested in a real life scenario putting the prototype under real life vibration effects would have been suggested as a second experiment. Experiment2 was also compared in details with the experiment1 in order to show the consistency of the prototype but computer simulation was just entering parameters produced from expriment1 into the simulator. As much as the result values were similar, so many factors could affect the input of data into a simulator such as how legit are the values being inputted into the simulator?

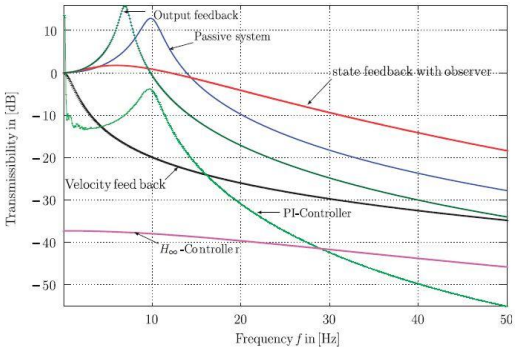
Looking at the methodology used in the research it is proven that external HDD vibration problems is as important as internal vibration problem because as much as a HDD is well designed, the external factors such as where the HDD is being used should also be considered and not enough research is being made towards that section. A prototype was developed and was only tested in a lab environment using a particular kind of HDD (3.5 inches) which is a basic computer desktop HDD, they did not make provision of testing the prototype with other kinds of HDD that are mostly subjected to external vibration forces, doing so it is only proper to agree that this methodology is good for products using the 3.5 inches HDD, on the other hand, most computers using the 3.5 inches HDD are desktop computers and are almost not affected by external vibration which defeats the purpose of the claims made.

Kerber et. al. (2007) considered other variables such as spring stiffness, damping, moments of inertia and the vertical position of the center of mass so as to eliminate other irrelevant factors affecting the external vibration of the HDD also comparing different controller design and their results. Looking through the research, the experiment was setup was using a HP-Paragon measurement system which is computer simulated experiment to drive specific actuators with a white noise input, the frequency is then measured using an acceleration measuring system and then built in amplifiers in the vibration isolation system is used to amplify the excitation signal, this response is also measured using commercial available accelerometer. This experiment could only produce two parameters which are geometric parameters and the mass of the upper plate depicted below.

Parameter	Value
m	19.3 kg
L_1	0.348 m
L_2	0.3 m
L_3	0.55 m
$L_4 + L_5$	0.19 m

Mass and geometry parameters (Kerber et. al. 2007)

Other parameters were derived mathematically such as rotational motion, moment of inertia and many more e.g. for getting damping parameter d_4 the transfer function $G = zp_3/F_{p1.}$ A before and after comparison was also made so as to be ensure progress was being made towards improving the benchmark of vibration isolation. Kerber et. al. (2007) also compared other controllers in order to come up with the best system for vibration isolation, below is a graph depicting their comparison. Kerber et. al. (2007) further explained that in order to be able to compare performance of other kinds of controllers, their acceleration transmissibility curves would have to be determined. The system is then excited in a z direction and the vertical response is computed. Using a general attenuation of -40dB at 0.1Hz, it was determined that the $H\infty$-controller outperformed other controllers at low frequency, considering the velocity feedback controller, it depicts that there is reduction in movement of the upper plate mostly at low frequency. Also looking at high frequency, the PI controller is seen to perform better which obtains a reduction of 20dB compared to the passive system. In terms of measurement noise, the PI controller is most affected and the $H\infty$-controller is least affected by it.

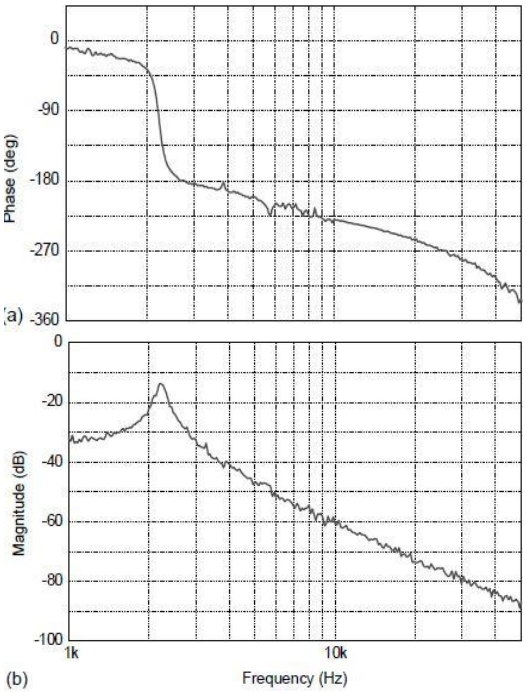


Acceleration transmissibility for vertical excitation and response (Kerber et. al. 2007)

Kerber et. al. (2007) claimed according to their experiments carried out and results shown they have reach the conclusion of the H_∞-controller giving the best performance thus using it in developing their own vibration isolation system which is known as a six degree-of-freedom model.

Looking at the claim made in the research and methodology, putting other systems used by previous researchers into consideration is convincing such as the PI-Controller where Haung et. al. (2003) proposed that the PI-Controller can achieve a steady-state decoupling by adjusting the tuning parameters of alpha and beta in the controller matrices. Testing various systems and comparing them under different variables and being able to pick which system suit their aim is also positive, what they didn't do was test their own system with the other systems that was previously used. Though the experiment carried out are repeatable, this experiment is majorly mathematical and does not have any input of real life factors included, a real HDD was not even used in any of the experiments so as much as the model worked inside a computer simulator it cannot be proven the it will perform the same way on a real HDD let alone in the real world.

Mou et. al. (2003) proposed the use of a single crystal silicon microstructure for positioning of magnetic heads for HDD compared to the dual stage actuation schemes in terms of material stability and lesser thermal coefficient of expansion which was backed up by experiment known as the finite element simulation experiment. Mou et. al (2003) described their actuator as an electrically isolated microstructure with aspect ratio of 20:1 directly processed from a single crystal silicon substrate, by fabricating their own micro actuator and comparing it in the same simulation test as the most widely used dual stage actuation schemes knows as the piggy back micro electro mechanical system. A dynamic testing of micro actuator bonded with slider experiment was also conducted so as to test a real life HDD.



Frequency response of micro actuator bonded with slider (Mou et. al. 2003)

Mou et. al. (2003) explained the graph above (a) and (b) as the frequency response from driving the voltage to displacement which shows that the first in-plane rotational mode had a resonance frequency of 1.91 kHz, and the first uncontrollable out-of-plane had a resonance frequency of 16.6 kHz which indicates that the micro actuator is feasible to operate as a fine actuator for high density HDD.

The above research is solely tackling internal HDD vibration by designing their own micro-actuator prototype for an HDD which in their result show better dynamic performance when compared with the regular micro-actuator however the experiment fails to consider other rotating parts such as the disk platter of the HDD since the micro-actuator is not the only rotating

component of an HDD additional testing is suggesting including other HDD factors before their claims can be justified. Although further testing might still be required, any kind of HDD manufacturer should adopt this methodology due to it mitigating the micro-actuator vibration section of an HDD thus reducing one factor that could be a vibration problem for HDD in general.

3. Comparisons

After analyzing the research methodologies, some methods can complement each other for instance the method used by Yap et.al. (2006) did not base their prototype on any previous research being made towards vibration isolation instead used a standard that is being used by HDD manufacturers, developing a prototype using blueprint of previous successful journals would be a good suggestion, comparing this to the method applied by Kerber et. al. (2007) who only conducted a computer simulation experiment which does not provide enough proof that the system would actually work on a real HDD, conducting experiments on an actual HDD helps back-up the claims proposed. On the other hand both journals focused on external vibration isolation. Yap et.al. (2006) showed that their methodology supports a low-cost vibration isolation system can be designed to allow HDD operate well when subjected to strong vibration input due to the experiments and component compared to Kerber et. al. (2007) who's methodology is on the high cost side of the budget due to the fact that multiple processes where experimented on using standard criteria to get the best process before developing their own based on that process. Recommendations can include applying the first research's methodology for a medium sized organization and majorly averagely skilled personnel since it is direct and easier to apply and also affordable in terms of cost. Also a mixture of the methods applied in first and second research for a larger organization and professionally skilled personnel due to the complexity of the methodology however recommending the methodology used in the second research as a stand-alone method isn't advisable simply because the only experiment carried out by Kerber et. al. (2007) was a computer simulation and can easily be affected by biased factors. It has no guarantee to be effective with real life factors. On the other hand comparing the two research analyzed with the third research, it is seen that Mou et. al. (2003) designed an internal part of an HDD that is prone to vibration and was tested both in a simulator and on a real HDD. They also compared their design with the most commonly used design adopted by other researchers. Their conclusion also involved the proposed solution. This methodology could easily be merged with any of the other two methods to get a far better vibration isolation system.

4. Conclusions

Comparing the research methodologies applied above in solving both internal and external HDD vibration problems, it is convincing that both problems are as important as the other, although without a good internal vibration isolation model an external vibration isolation model isn't of much importance however a really good internal vibration isolation system could function without the input of an external vibration isolation system simply because the external HDD vibration solutions require building some kind of device supporting the HDD compared with the solution which is incorporated inside the HDD itself which is why more research is being focused to the internal solutions. Recommendation can include inculcating both methods into each other in other to get the best system, a mixture of a method where you build a vibration resistant component of an HDD which is an internal methodology and some kind of support where the methodology is direct and easy to apply with consistent evaluation which is an external methodology is by far a better method of mitigating HDD vibration. This can be applied by anybody in general interested in solving vibration problems for HDD.

References

Choi M. and Yoon C.H., 2011, 'Design of Herringbone Groove Disk Damper for Effective Suppression of Axial Vibration of Disk in Hard Disk Drive', *IEEE TRANSACTIONS ON MAGNETICS*, Vol.47, No.7.

Horowitz R., Li Y., Oldham K., Kon S. and Huang X., 2007, 'Dual-stage servo systems and

vibration compensation in computer hard disk drives', *Control Engineering Practice*, Vol.15, pages 291-305.

Hoque E., Mizuno T., Ishino Y. and Takasaki M., 2010, 'A six-axis hybrid vibration isolation system using active zero-power control supported by passive weight support mechanism', *Journal of Sound and Vibration*, Vol.329, Pages 3417-3430.

Huang X., Elliott S.J. and Brennan M.J., 2003, 'Active isolation of a flexible structure from base vibration', *Journal of Sound and Vibration*, Vol.263 pages 357-376.

Jintanawan T., Sillapapinij A. and Ajavakom N., 2009, 'Effects of Tolerance Design on Suppression of Electromagnetic-Induced Acoustic Noises and Vibration Transmission in Hard Disk Drive Spindle Motors', *IEEE TRANSACTIONS ON MAGNETICS*, Vol.45, No.11.

Kameya T., Matsuda Y., Yamaguchi H., Egami Y. and Niimi T., 2012, 'Pressure-sensitive paint measurement on co-rotating disks in a hard disk drive', *Optics and Lasers in Engineering*, Vol.50, pages 82-86.

Kerbera F., Hurlebaus S., Beadle B.M. and Stobener U., 2007, 'Control concepts for an active vibration isolation system', *Mechanical Systems and Signal Processing*, Vol.21, pages 3042-3059.

Kirpekar S. and Bogy D.B., 2008, 'Computing the aero elastic disk vibrations in a hard disk drive', *Journal of Fluids and Structures*, Vol.24, pages 75-95.

Mou J.Q., Lu Y., Yang J.P. and Li Q.H., 2003, 'Single crystal silicon rotary micro actuator for hard disk drive', *Sensors and Actuators*, Vol.108, pages 59-63.

Shen I.Y., 2000, 'Recent vibration issues in computer hard disk drives', *Journal of Magnetism and Magnetic Materials*, Vol.209, pages 6-9.

Sibielak M., 2012, 'Optimal controller for vibration isolation system with controlled hydraulic damper by piezoelectricstack', *Mechanical Systems and Signal Processing*, Volume and Pages Unknown.

Yap F.F., Vahdati N. and Harmoko H., 2006, 'Design and analysis of vibration isolation systems for hard disk drives', *Journal of Magnetism and Magnetic Materials*, Vol. 303, pages 52-56.

A Critical Evaluation of the Current User Interface Systems Used By the Blind and Visually Impaired

Amneet Ahluwalia

Abstract

This paper evaluates some current research methods on computer interaction systems that are being used by the visually impaired and blind users. Though there are many user interfaces provided for such users, it has been analysed that the use of auditory interfaces are the most efficient. An evaluation of a few user interfaces has been discussed and some future work has been recommended in this paper.

1 Introduction

"Throughout the history of human-computer interface development, one aspect has remained constant: output from computers has been almost entirely visual" (Edwards 1989). The visual interface has always been a barrier for the blind and visually impaired, when it comes to using the computers. There has always been a need for easy user interfaces that can assist the blind and visually impaired users in performing complex tasks on computer systems.

A research by Moll et. al. (2012) suggests a system that combined sound and tactile devices in order to receive auditory and haptic feedbacks. Another research by Bellik (1997) recommends an interface that uses speech along with multimodal text editor. The Braille system is the most commonly used method to input data in the computer systems, using specialised Braille devices such as Braille keyboards. However, this method cannot be termed as a perfect method as the recognition and generation of Braille documents is still a problem. Tai et. al. (2010) have suggested belief propagation to recognise Braille documents so that these documents can be preserved and are made easily available to large number of visually impaired people. Hentzschel and Blenkhorn (1995) propose the use of twin shadows approach for embossed Braille characters in an optical reading system. Jeong (2008) produces a prototype of an online Braille generator that is used by touch technology. Many other research projects have aimed at providing accessibility options to the visually impaired and blind users such as (Dixon, 2011), (Simsek et. al., 2010) and (Hellar, 2002).

This paper will produce a discussion of the various interface methods that have been proposed to assist the visually impaired and blind users. The purpose of this paper is to evaluate these interfaces in order to determine which methods are sound and have been evaluated well enough with proving results of the claims made.

2 Braille Input Method

Huang et. al. (2004) in their research, have presented a user interface system of the desktop environment for visually impaired people. According to their proposed method, the current Braille input technology will be modified by integrating their presented interface into the kernel of the operating system. They think that his would be a better idea rather than developing an add-on user interface which would then only accept input from the applications that will meet the specifications of that interface. The user interface would enable the visually impaired users to have their input text displayed on the screen of the computer as printed text and on tactile Braille reading device. They can also hear the input with the help of voice from an audio device. This new interface has been implemented on Linux operating system, using a Chinese/Braille translator called JMCe. A test experiment was also performed to evaluate the efficiency of the presented interface. This test was performed on a machine using Pentium II-

233 processor and a 64M RAM using a file which contained 170 lines of text, each line including a combination of English & Chinese characters that were being created. When the mouse cursor was randomly moved in the text file 150 times, the average time that it took to export the Braille code to the Braille output buffer after detecting the current position of the cursor was 11 ms, which is a very short time interval for the visually impaired people to use the computer systems easily and fluently. The research has been concluded as a successful attempt to assist the visually impaired users to read the screen text which further helps them in internet browsing. This new user interface was also implemented in April 2002 and since then has been utilized in the Blind Information Centre, Tamkang University.

However, there is only one experiment performed to evaluate the new designed interface. This experiment only evaluates the time taken by the system to export the Braille code to the Braille output buffer. Though this is a very good experiment and has provided excellent results even though it was tested on a slow computer system (Pentium II, 64M RAM). Keeping in mind the faster computer systems being used today with heavy processors and large RAM sizes, it is clearly evident that this system would work even faster on them as it works along with the kernel of the operating system. There was no experiment performed to evaluate the tactile device and the output given by it, which is one of the major drawbacks of this research project. The tactile reading device plays an important role in this interface and its working needs to be tested in order to claim that the users can touch and identify the text on screen through the tactile Braille reading device. Also, no experiment has been performed to evaluate the working of the audio device that provides with a voice feedback of the input text. The audio device needs to be tested before it can be put into use. As this new interface is using various output methods, it is very important that the working of each element is tested and evaluated. Though the authors have mentioned that their presented interface is being utilized since 2002 in the Blind Information Centre, Tamkang University, we cannot ignore the fact that this research paper was only published in 2004. A time span of 2 years is not a very long period to testify for such an integrated system. Also, it has not been mentioned if there were any complaints made regarding the working of this interface. Therefore it can be concluded by saying that the designed interface does look very impressive, but a further detailed evaluation could have been more helpful to make a choice.

3 Homer II

Homer II is an interface that has been developed for the visually impaired and the blind people by Pavesic et. al. (2003). This is a text-to-speech interface system which is driven by voice for reading texts pertaining to the language of Slovenia. The system has been designed in a way that it can be integrated into the applications based on the Internet which would enable the users to have a remote access to all the available written information in the language given. Homer II is a spoken language interface where the system's voice control facilitates the entire working of the system which leaves out the use of the other input interfaces such as the keyboard and the mouse. It is a database of the text corpora which has been set up at the source centre of information to provide users with the text-to-speech function. Homer II has been divided into four modules where the first module works at enabling the internet communication and retrieves the text and changes it into a standardised form on a local disc. The second module of the system works on managing the dialogues that take place with the users and performing an access to the database of the text. The third module is output module that works at converting text to speech and the fourth module is the speech recognition module which is speaker-independent or uses input from the keyboard and keeps running parallel with all the other modules.

A preliminary offline evaluation of the fourth module has been done to check how accurate the recognition of the command that is spoken is. A database of clean speech has been used of 6 testing speakers and 20 training speakers. The test results show an error rate of recognition which is as low as 2%. Further testing has been done to consider the main factors such as the number of syllables that are present within one word, chosen speaking rate and the word's position in a phrase, the phrase initial, and the

phrase final or nested inside a phrase. The experiment has been performed under the norms of the ITU-T Recommendation P.85 and the method takes into consideration the attitude as well as the performance of people using it. The attitudes of the users are evaluated by using multiple scales. All the users were given out response sheets and were told to fill in the different answer templates that were related to the application domain chosen and were based on the information they heard. Airline timetable information retrieval was chosen as the application domain and the results showed that more than 90% of responses were filled in correctly. And the ones that were mistaken were mainly the names of foreign airports which may be unknown to the users or may not be easy to spell. The second part of testing was aimed at comparing the several features that described the synthetic voice attribute to those that described the attributes of natural speech which was distorted with various levels of Gaussian noise. This experiment took place according to the norms of the ITU-T Recommendation P.81 which describe a technique to compare synthetics speech with natural distorted speech. The results of this test showed that the synthetic speech scored a mean opinion score with a SNR ratio of 5 and 10dB and that was between distorted natural speeches.

The performed tests have evaluated the speech synthesis. All the tests have been performed under the specified regulations of the Telecommunication Standardization Sector (ITU-T). The tests have yielded great results. Since each test has been performed under the International Standards any possibility of having biased results is ruled out. Every function of the new interface has been tested individually and the results are very positive in stating that the new designed system is very efficient in providing help to the visually impaired users as well as the blind users. The first test used to evaluate the accuracy of the recognition of spoken command has been performed with a variety of voices, which is a very good method as there can be a difference in the voice of different users. It is very important to consider the different voice tones of different people as it may affect the system's performance when it is being used. But Homer II has yielded very good results in recognizing the voice commands with a very low error rate. The second test has also been performed with various users and has achieved a very good result. The test results from the third test claim that the voice is clear enough to be used and understood by the users. Therefore all the tests are clearly stating in their results that the new designed interface is justifying all the claims made by the authors and that this method is very useful and ready to use for the visually impaired and blind users.

4 Spatial and Non Spatial Interfaces

A research produced by Sodnik et. al. (2011) is focussed on the use of synchronized spatial sounds in an audio interface that has been designed for the visually impaired and blind computer users. There are two audio interfaces proposed in their research in non-spatial and spatial sound conditions, representing a simple word processing application's hierarchical menu structure, which would enable all basic operations of the application such as creating a text document or editing it. In the interface, each menu item is represented by a spoken command and it provides a function of selecting between the various available menu items at each menu level. The spoken commands of each menu or its sub-menu have been categorized into horizontal and vertical spatial configurations. Both these configurations can be used with or without the use spatial sounds; hence four different audio conditions have been formed: Horizontal Interfaces (H1-non spatial and H3-spatial) and Vertical Interfaces (V1-non spatial and V3-spatial). The menu commands in H1 and H3 are positioned on the horizontal plane of a virtual ring which surrounds the user's head. The sound commands in V1 and V3 are associated vertically in front of the user. The interaction to this interface is based on a QWERTZ keyboard where the navigation for all the four interfaces is based on the arrow keys.

Test experiments have been performed to evaluate how efficient are multiple spatial sounds as compared to the non spatial sounds, when they are used for hierarchical menu navigation. The experiment was performed by 8 visually impaired users and 4 blind users, and because it was an evaluation of an auditory interface, it was made sure that all the participants had normal hearing capabilities.

TS1, TS2, TS3 and TS4 were the four sets of tasks designed for this experiment. Each set consisted of ten different tasks with almost the same difficulty levels to have a comparison between H1, H3, V1, and V3. Each set was allocated a completion time of 5 minutes. The users were given clear audio instructions for each task and were also made to undertake a warm up exercise of a set of five tasks. All the 12 participants were put into four groups consisting of three users each, to perform the experimental tasks. The evaluation of the four interfaces resulted in the measurements of the total time taken to complete the tasks by each interface and navigation performance of each interface. The time taken by tasks to complete was automatically measured by the application and the results of the average taken by each interface is shown in Figure1.

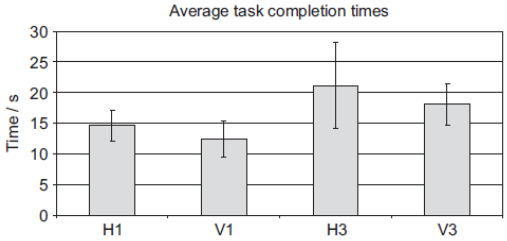


Figure1. The average task completion times of all tasks with four interfaces (Jaka Sodnik et. al. 2011).

The results clearly showed that H1 and V1 are faster than H3 and V3 which means that the hypothesis, made about spatial interfaces being faster than the non spatial interfaces, has not been confirmed. To evaluate the performance of navigation, the participants were supposed to turn the virtual ring left and right, in the horizontal interface and the users were supposed to move up and down the menu for selecting the required command, in the vertical interface. This navigation was measured automatically. The final sum of all the participants' actions was calculated (which means selections + movements), and the results shown in Figure2 clearly state that the interaction with spatial interfaces is more efficient than that with the non spatial ones.

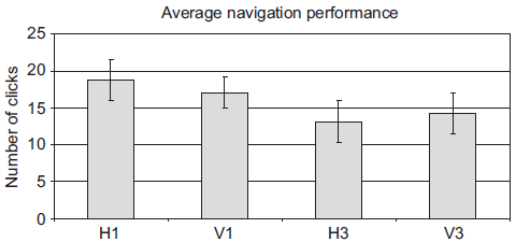


Figure2. The average navigation performance (the sum of key presses) of all tasks with the four interfaces (Jaka Sodnik et. al. 2011).

This research project undertakes a very good testing strategy to evaluate the hypothesis made by the authors. The user experience tests have been performed in an effectively planned manner. A variety of users (12 participants) were made to participate in the test plan. It is always good to have a number of users to test an interface, rather than having just 1 or 2 users, because the testing then gives a clear result picture of how the interface is going to perform when it is used by different type of people. The same thing has been kept in mind by the authors to evaluate their method. It was also made sure that all the participants had normal hearing and all the participants were divided into groups, so that there is no bias on the basis of knowledge of an individual. Performing in a group means that the results will be an outcome of collaborative knowledge. The times were recorded by the application itself and not by an individual therefore the times are accurate as per the performance of the participants. However, the results of task completion time test have proven that the hypothesis of the authors is incorrect. But this does not mean that it is a failed research. It gives us some new knowledge that the use spatial sounds in an interface could slow down the performance of the users, as after a while, spatial sounds become stressful. However, it is also proven by the navigation performance of the users that spatial interfaces are more efficient than non spatial interface. Therefore it can be concluded by saying that spatial interfaces are better than the non-spatial interfaces, if they are used for short interval tasks. For day to day use of computers, it is easier for the visually impaired and blind users to use non spatial interfaces.

5 Conclusions

The user interface is the most important part of interaction between computers and visually impaired and blind users. The interface that is the easiest to use is the most helpful one. After evaluating a few interfaces, it is clearly evident that the use of audio devices and tactile devices helps in making the use of computers easy. It can be concluded by saying that auditory feedback is the most important aspect of designing a user interface for the blind and visually impaired people. However, while developing a new method, it is very important to keep in mind that the use of spatial sounds has been proven to be more efficient but stressful at the same time. An interface that would use Braille input method and provide visual display, tactile feedback, voice command recognition and would enable the user to switch between spatial and non spatial sounds, based on their preference, would be very helpful for the users to do all the tasks efficiently and to use computers easily without any stress in their day to day life. Such a user interface could be recommended for the future.

References

Bellik Y., 1997, 'Multimodal text editor interface including speech for the blind', *Speech Communication*, Vol. 23, Pages 319-332.

Dixon J. M., 2011, 'Braille: The Challenge for the future', *Journal of Visual Impairment and Blindness*, Vol. 105, Pages 742-744.

Edwards A. D. N., 1989, 'Soundtrack: An auditory interface for blind users', *Human-Computer Interaction*, Vol. 4, Pages 45-66.

Heller M. A., 2002, 'Tactile picture perception in sighted and blind people', *Behavioural Brain Research*, Vol. 135, Pages 65-68.

Hentzschel T.W. and Blenkhorn P., 1995, 'An optical reading system for embossed Braille characters using a twin shadows approach', *Journal of Microcomputer Applications*, Vol. 18, Pages 341-354.

Huang J., Tung M, Wang K. M., Chang K, 2004, 'A user interface for the visual-impairment', *Displays*, Vol. 25, Pages 151-157.

Jeong W., 2008, 'Touchable Online Braille Generator', *Information Technology and Libraries*, Vol. 27, Pages 48-52.

Moll J., Huang Y. Y., Sallnas E. and Sundblad Y., 2012, 'Auditory feedback in haptic collaborative interfaces', *International Journal of Human-Computer Studies*, Vol. 70, Pages 257-270.

Pavesic N., Gros J., Dobrisek S. and Mihelic F., 2003, 'Homer II- Man-machine interface to internet for blind and visually impaired people', *Computer Communications*, Vol. 26, Pages 438-443.

Simsek O., Altun E. and Ates A., 2010, 'Developing ICT skills of visually impaired learners', *Procedia Social and Behavioral Sciences*, Vol. 2, Pages 4655-4661.

Sodnik J, Jakus G. And Tomazic S., 2011, 'Multiple spatial sounds in hierarchical menu navigation for visually impaired computer users', *International Journal of Human-Computer Studies*, Vol. 69, Pages 100-112.

Tai Z., Cheng S., Verma P. and Zhai Y., 2010, 'Braille document recognition using Belief Propagation', *Journal of Visual Communication and Image Representation*, Vol. 21, Pages 722-730.

Current Research Aimed At Improving Bot Detection In Massive Multiplayer Online Games

Jamie Burnip

Abstract

Advancing internet speeds and processing power has made the dream of a second life in a medieval world has become a reality, or rather a virtual one. Gamers put a lot of time and effort into making their online avatar the best they can be, but there are some out there who choose to take the easy route and have a program do all the work while they take all the glory. This paper looks into the current research that is aimed at stopping this specific cheat, each method will be critically evaluated to determine if they are a suitable solution for preventing MMORPG servers becoming predominantly populated by automated characters.

1 Introduction

MMORPGs (Massive Multiplayer Online Role Play Game) are massively popular these days and it's all down to escaping real life to do something that isn't possible like slay a dragon. It's not uncommon for people who put a lot of time into their game character to be proud of their achievements.

Even in a game where the hard work pays off, some people still choose to cheat and the most common type of cheat in MMORPGs are game bot. Game bots are programs that will play a character automatically, without the need to rest or necessary break like a human player.

"The bot can reap those rewards very efficiently 24 hours a day, without fatigue or boredom" (Mitterhofer et al. 2009).

This is a major problem for the games development because if legit players begin to feel hard done by, they may leave the game. Since most MMORPGs are subscription based, this would lead to a massive revenue loss if it got out of hand.

"The cheaters destroy the game balance by rapidly depleting in-game contents and resources. Honest human gamers may thus feel deprived, lose interest, and eventually leave the game" (Kang et al. 2012)

This paper looks at research that has been done to develop new ways to stop people from using bot programs and critically evaluates their application to a real game system.

2 Presentation and Evaluation

In order to complete a critical evaluation on current research that is aimed at detecting bots in MMORPGs an extensive literature review was done. 10 papers in total were found although only 4 of them were evaluated. The work of Bakkes et al. (2012), Bethea et al. (2011), Prasetya & Wu (2010) and Pao et al. (2012) were not specified enough for use in MMORPGs, in fact they were tailored for first person shooters. Their application to MMORPGs would be too impractical.

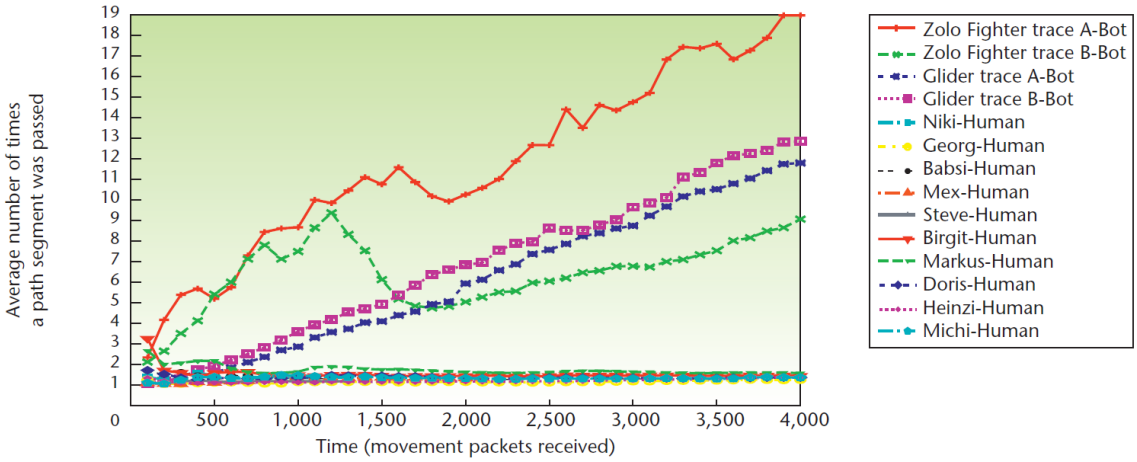
The work by Gianvecchio et al. (2009) was focused on human observation; this was way out the scope for this paper. This papers scope is concentrated on a more automated approach to bot detection, therefore it was not evaluated.

Thawonmas et al. (2008) proposed an analytical approach based on the behaviour of the character in question, unfortunately this was very similar to the work of Park et al. (2010).

The latter was chosen instead because the results proved to be a more viable research paper.

2.1 Waypoint Analysis

Mitterhofer et al. (2009) proposed that by tracking the waypoints of a character from the server-side, they could accurately expose bots without the exploits found in client-side solutions.



Graph 1(Mitterhofer et al. 2009)

By using this the authors claim that their system will be invisible to the player that doesn't impact the legitimate players at all. The information they require for this is easily acquired because the game client is regularly communicating with the server, updating it on the new coordinates of the character. By using the coordinates sent to the server, they reconstruct the route of a character in the game by connecting them up. This is done using a line simplification algorithm that will ignore waypoint dots that build up quickly – for example, if the character does a lot of movement in a short radius.

Mitterhofer el al. (2009) used a controlled lab experiment to test their method on a private game server against two different bot programs, the first bot being the popular 'Glider' bot and the second being the 'ZoloFighter' bot. The authors had 7 people play the game at a LAN party with 3 other human players join them over the internet. The experiment ran for 4 hours while they had the human players concentrate on farming other than questing, this is so that their game style would be mirrored that of a bots.

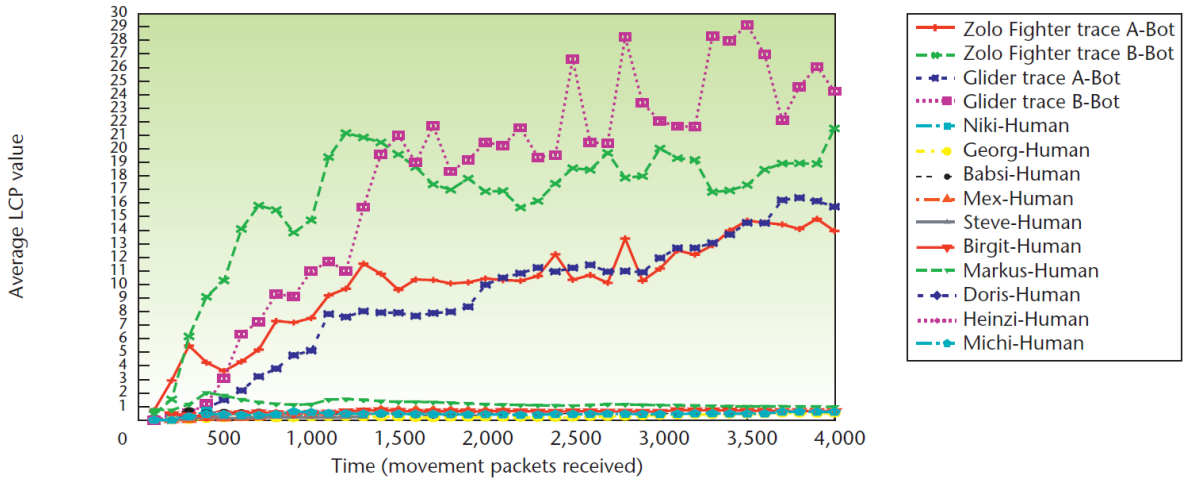
Graph 1 shows the difference of playing styles between humans and bots. The graph highlights the frequency of time spent in a specific radius of both bots and humans The authors explain that the 'dip' in the ZoloFighter b-bot is due to a human regaining control at 1,200 packets and then setting the bot away on a new path at 1,800 packets.

"The goal was to purposely disrupt the bot path and evaluate the influence on our detection mechanism" (Mitterhofer et al. 2009).

Graph 2 shows the length of time these waypoints or paths were repeated (shown in graph 1), the graph shows that human players never moved in repeating patterns, which kept them low. The graphs accuracy is based on whether they can detect these repetitions in waypoints so to differentiate between human players and bots they set the threshold for the y axis to 5 on both graphs.

The system that Mitterhofer et al. (2009) has developed is big step up when it comes to analysing character waypoints to detect if it is being controlled by a bot; the experiments results pretty much speak for themselves for that. However, it is not without its flaws as Mitterhofer et al. (2009) explains that if knowledge of the system became free on the internet, bot programmers could attempt countermeasures.

These could include a bot that was specifically made to avoid this type of detection. This would require more random movements which are considered unsafe. The reason for this is because it will cause stupid looking behaviour that could be obvious to human players and even has the character get stuck on a rock or fall off a cliff. This is acknowledged by Mitterhofer et al. (2009) and only further backs up their claims.



Graph 2(Mitterhofer et al. 2009)

If bots were to use longer paths they would lower the risk of being detected as it would take a while to come back to a familiar waypoint. However this is not likely as creating a large path may result in leading the character into an area too dangerous or out of a point of interest; for example if they were killing certain creatures for a particular reason. The authors admit that if bots were to use longer paths they run the risk of wasting valuable 'farming' time and if they were to use paths from the internet it is likely that it will be known by the games bot detection system too.

As it currently stands, the claims made by the authors are valid, so long as bot programmers don't develop a bot that can navigate the game world without the use of a path or coordinates.

2.2 Behaviour Analysis

Certain features make up a characters behaviour, Park et al. (2010) uses some of these (map changes, counterturn, rest states, killing time, experience point and stay in town) and collects data built from them to distinguish a human player from a bot.
Park et al. (2010) claim that behaviour patterns between human players and bots are easy to differentiate; they classify this with Table 3.

Table 1 outlines in short how a bot and a human player will demonstration these patterns. By using that model Park et al. (2010) compared data between 2 human players and a character controlled by a program called 'OpenKore'.

"This Game Bots is the most famous Game Bots in Ragnarok Online" (Park et al. 2010).

Behavior Features	Behavior Pattern	
	Human Player	Game Bot
Map Changes	Various of Map changes	Uniformity of Map changes
Counterturn	Occasionally occur	Frequently occur
Rest States	Various terms of Rest States	Uniformity terms of Rest States
Killing Time	Various time to spend Enemy kill	Uniformity time to spend Enemy kill
Experience Point	Various amount of Experience Point Gain per time	Uniformity amount of Experience Point Gain per time
Stay in Town	Various time of Stay in Town	Uniformity time of Stay in Town

Table 3(Park et al. 2010)

The results of the data collection are shown in multiple tables, depending on what behaviour feature they belong to. Table 4 is showing the results of how much time is spent in a town, clearly marking the human players and the bot.

Behavior Features	Game Bot	Human Player A	Human Player B
Mean (second)	0:00:41	0:04:32	0:14:27
Standard deviation (second)	0:13:32	1:59:07	9:03:29
Coefficient of variation (second)	0.82	1.09	1.57

Table 4(Park et al. 2010)

Park et al. (2010) claimed that the time spent in a town will be different in human players and bots, bots actions in a town are set the same, specific actions that should lead back out to the open world as quickly as possible. "They act only necessity behavior like store items, withdraw items, sell goods, buy items, etc. But in human player case, they do various activities even if they are useless" (Park et al. 2010). The results for the other behaviour features shared this trend, albeit in a different format.

The method proposed by Park et al. (2010) has a strong base, their claims are met with sizable results but they only seem to take certain things into account. Table 4 shows the results of time spent in what looks like a relatively small town, this is due to the speed in which the bot enters the town, visits the vendor and leaves town. This may be fine in games like Ragnarok Online, but if they were to test this in World of Warcraft where the towns range from a small one like shown in table 4, to grand city's that take a few minutes to navigate. Other features used such as experience points will also vary as these games offer out a resting bonus to players who have been offline. These bonuses grant an experience point multiplier to compensate the player during their offline state allowing for 'catch-up' time.

2.3 Clustered Network Analysis

An analysis of a clustered network is a method proposed by Lee et al. (2011); in which authors claim that finding the bank character for a bot nest is a more practical way of stopping bots all together. In this case, the clustered network is the trade logs on the server, according to Lee et al. (2011) due to the vast amount of money and items traded between bot characters. The detection rate is quite high, but this is currently an untapped method.

Since the bank character is in the center of this, as shown in figure 5, it opens a network of not only 'farming' bots, but also players who have been buying game money, this is also against the games terms of use.

Although they can only identify a bank character based on the trade logs of a pre-detected bot. The authors leave the pre-detecting of the 'farming' characters to such systems as developed by Mitterhofer et al. (2009) Thawonmas et al. (2008) whom have both been mention in this research. Lee et al. (2011) state that if a character has had one or more trades with pre-detected bot, then they must be a bank character or at least involved in the bot nest. After this a character is flagged as a suspected bank character and the analysing of their trade network begins.

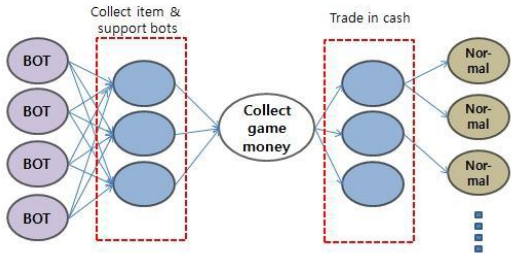


Figure 5(Lee et al. 2011)

Figure 6 represents a trade graph of detected bots and bank characters in the game Aion, the game tested by the authors. It very closely resembles the sample graph that the authors make at the start of their research. With some miner spin off trades which are to be expected from normal characters, yet still stands as proof of how a bot nest works and the authors claims. The high density of trades shows how regular bots trade to bank characters that then trade onto broker characters and human players.

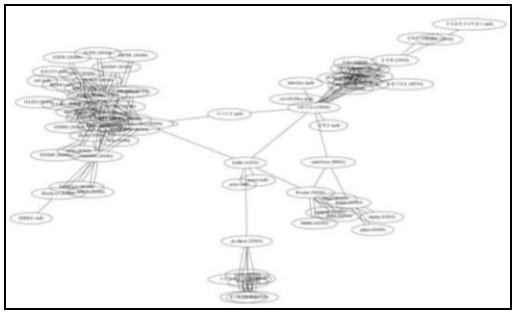


Figure 6(Lee et al. 2011)

The authors applied their system 10 times with the log data taken between the dates of 22/10/2010 and 30/12/2010, and they summarised the results into a table (table 7). The table shows the number of bots found and the number of bank characters involved with them, as well as the amount of game money traded for each per application.

	Number of accounts		Game money (unit :million)	
	BOT	Bank	BOT	Bank
1st	25,695	1383	737,000	1,600,000
2nd	6,322	446	10,200	333,400
3rd	1,217	119	21,600	162,700
4th	974	106	14,000	28,100
5th	3,763	171	348,700	81,100
6th	1,031	140	15,500	97,900
7th	1,053	137	12,700	167,400
8th	467	71	13,400	12,300
9th	13,742	512	807,800	870,000
10th	704	72	54,400	27,300
total	54,968	3,157	2,035,300	3,380,200

Table 7(Lee et al. 2011)

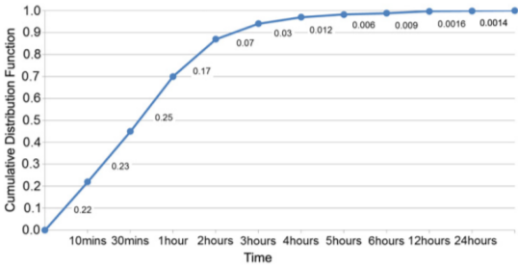
According to the authors the current system recovers on average 37 million kinas (currency in Aion) per bot character. Their system recovered 1,070 million kinas per bot character, stating that this has resulted in a massive hit to the economic flow of the bot nest because they managed to detect both the bank and broker characters.

The results of this experiment are definitely measurable to their claim. Not only exposing the bot but the entire of network of bots who exploit the game, that is why one day this may be the system that pushes bots to disappear completely if it were not for one crucial flaw. The system that Lee et al. (2011) have created uses bots that have been detected by other detection systems making it only as effective as the pattern detection system. This causes the system to be quite poor, as long as the 'farming' bots are working around the detection system. The authors system could also have bad implications on legit players that may get mistaken by the current system for a bot.
If this happens it will produce a false and meaningless trade networks ending in wrongful banning or suspension.

2.4 Party-Play Log Analysis

Kang et al. (2012) have proposed that analysing the log data of the in-game parties, based on the fact that sometimes an activity in the game requires a player to group up with other players, but that they wouldn't be in a party for an extended period of time. The authors claim that human players would normally group up for a few hours at a time. Whereas a bot group would have one hunt and/or harvest whilst the other one acts as a bodyguard "parties are to be disbanded once their goals are met, bot parties whose goal is to accrue as many resources as possible are expected to last indefinitely" (Kang et al. 2012).

The authors collected log data for seven days and found that there were 63,092 parties made on a single server. This data is represented in graph 8 and shows that only 0.3% of all parties last longer than 12 hours. To which the authors have judged to be bot parties.



Graph 8(Kang et al. 2012)

The authors then analysed how many characters made up the parties they collected from the log and this was put into multiple pie charts shown in figure 9. The stats in figure 9 also stand to further verify their claims that bot parties are made up of two characters. "This is highly consistent with our hypothesis that bots form parties in groups of two so that one can protect the other while it is harvesting or hunting for items"(Kang et al. 2012).

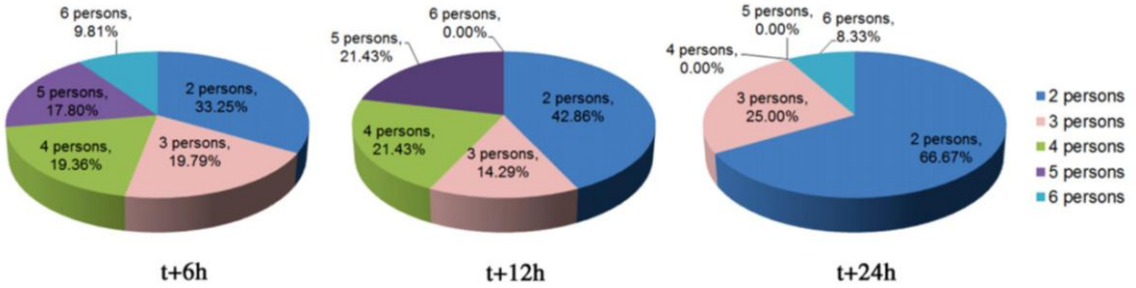


Figure 9(Kang et al. 2012)

The data collected also shows that the ratio of getting experience and collecting items rises dramaticly for parties that have been flagged as bot parties compaired to parties containing human players, by 20% in most cases. "This implies that game bots focus on getting game money and items while human players do not"(Kang et al. 2012). This is also shown in quest completion, which was expected as bots do not normally care about completing quests because it's considered a waste of time. They also used other in-game features to analyse behaviour similar to the work of Park et al. (2010).

Their experiment was concluded by taking the data and flagged bot parties and running them by a rule. The rule they used is shown in figure 10, where if a party passes them boundaries, they must be a bot and are immediately banned from the game.

Bot detection rule.

Rule-base
Getting experience log >= 34% and getting race point log <= 1.69% and sitting log <= top 10 and using item log <= 1.19% and quest completion log <= 0.16% and start volplane log >= top 34 and party member = 2 and party duration >= 600 s

Figure 10(Kang et al. 2012)

The method developed by the authors managed to identify 49 bots out of 52,377 party-play users by appling the rule shown in Figure 10. Although out of the 49 identified, only 47 were banned from the game to which the authors claim, this makes their method 95.92% accurate.

The venn diagram in figure 11 shows the results compared to that of the games current internal monitoring system. Out of the 49 bots that were identified, 21 of them weren't picked up on at all by the games current internal monitoring system and two were picked up but didn't get banned.

Proposed detection method	Bot detection log	Banned account list
49	26	47

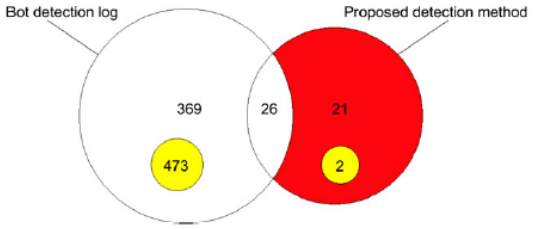


Figure 11(Kang et al. 2012)

The authors claimed at the beginning of their paper that through the use of party log data, they could identify bot characters who are working as a team. Although this concept has proven in the results to be valid due to 21 bots that the games current internal monitoring system didn't pick up on, it requires that the bots are in the same party. This means that any bot that is working alone is slipping under the radar, this is proven by the 369 bots that it didn't pick up on. There also remains the 2 un-accounted for bots in their research. The authors also claim that after their experiment, there method has a proven accuracy of 95.92% which is somewhat true, the reason it is not 100% is due to the two un-accounted for bot. Although the margin of detected bots is small in comparison to the games current internal monitoring system, it still increasing the detection rate. The authors have indeed achnoleged this but stress that it cuts the risk of false-negative detection erros.

3 Conclusions

The research that has been reviewed in this paper all have decent claims with results to match.

Although some might not be able to function without prior data from the system, their ability

to analyse even the finest of movement or behaviour poses great threat to bots.

If the research proposed by Lee et al. (2011) were to be linked onto the work of Mitterhofer et al. (2008), they would have a system that detects the farmer bots who are out 'grinding' but then use their trade logs to find and entire nest of bots on the server. This would be possible only if system by Mitterhofer et al (2008) passes data that identifies the character along to the system proposed by Lee et al. (2011). From there the log data needed is already saved by the server so it is easily accessed.

Alternatively the behavioural approach that has been proposed by Park et al. (2010) could be used as opposed to waypoint analysis. This would eliminate the possibility of bot remaining undetected because they don't follow a pre-set path. Although due to the method itself, it would take longer to react to a bot and could run the risk of false detections.

Analysing party-play logs was proposed by Kang et al. (2012) and this has shown to be effective at exposing bots; it can only do so if the bots are grouped together in a party. It is for this reason that it would be an unlikely choice for game developers to adopt this system over the other systems. If a system is already exposing bots that are both in and out of parties, it's unnecessary to focus on that factor. Although it has been shown to more accurately detect the ones who are partied up, it would only be applicable if the games system had the power or resources to run that along with another one.

References

Bakkes, S.C.J, Spronck, P.H.M, Lankveld, G.v. (2012). 'Player behavioural modelling for video games', *Entertainment Computing.* 3 (3), p71-79.

Bethea, D, Cochran, R.A. and Reiter, M.K. (2011). 'Server-side Verification of Client Behavior in Online Games', *Transactions on Information and System Security*. 14 (4), p1-27.

Gianvecchio, S, Wu, Z and Xie, Mengjun and Wang, H. (2009). 'Battle of Botcraft: fighting bots in online games with human observational proofs', *Proceedings of the 16th ACM conference on Computer and communications security*. p256-268.

Kang, A.R, Woo, J, Park, J, Kim, H.K. (2012). 'Online game bot detection based on party-play log analysis', *Computers & Mathematics with Applications*.

Lee, E, Lee, J, Kim, J. (2011). 'Detecting the bank character in MMORPGs by analysis of a clustered network', *The 3rd International Conference on Internet*, p507-511.

Mitterhofer, S, Kruegel, C, Kirda, E, Platzer, C. (2009). 'Server-Side Bot Detection in Massively Multiplayer Online Games', *The IEEE Computer and Reliability Societies.* 09, p29-36

Pao, H, Fadlil, J, Lin, H, Chen, K. (2012). 'Trajectory analysis for user verification and recognition', *Knowledge-Based Systems*. 34, 81-90.

Park, S.H, Lee, J.H, Jung, H.W and Bang, S.W. (2010). 'Game behavior pattern modeling for game bots detection in MMORPG', *Proceedings of the 4th International Conference on Uniquitous Information Management and Communication*. 33, p1-5.

Prasetya, K and Wu, Z.D. (2010). 'Artificial neural network for bot detection system in MMOGs', *Proceedings of the 9th Annual Workshop on Network and Systems Support for Games.* (16), p1-2.

Thawonmas, R and Kashifuji, Y and Chen, K.T. (2008). 'Detection of MMORPG bots based on behavior analysis', *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*. p91-94.

Evaluation Of Methods For Improving Network Security Against SIP Based DoS Attacks On VoIP Network Infrastructures

David Carney

Abstract

This paper investigates current proposed methods that have been recommended to protect VoIP network infrastructures from (Session Initiation Protocol) SIP based (Denial of Service) DoS attacks. As with any service or application that uses the internet VoIP is exposed to all of the known attack threats in the IP environment, but also has to contend with additional attacks that are specific to VoIP. Several of the protocols used in association with VoIP have known vulnerabilities. This paper specifically looks at 2 proposed methods, used to detect and prevent DoS attacks aimed at SIP on VoIP infrastructures. The focus is on evaluating different research in this area that has been published. All experiments and results carried out by the research authors are considered, and the paper concludes by looking at the strengths and weaknesses of each of the proposed methods in order to reach conclusions as to which offers the most secure method of security.

1 Introduction

As VoIP is becoming more and more widely used in everyday life for both personal and commercial use it has opened up new avenues of attack for hackers with malicious intent (Edelson 2005). VoIP uses an IP-based infrastructure which means that it will inherit all of the currently known attack threats in the IP environment such as spoofing, man-in-the-middle attacks and DoS attacks. VoIP also has to contend with additional threats which are specific to the Session Initiation Protocol (SIP), including registration/call hijacking, setting up and tearing down sessions, impersonating a SIP server and SPIT (Spam over internet) (Bradbury 2007). SIP is not the only protocol used by VoIP (although it is currently the most dominant). It may also use other technologies such as H.323, MGCP and MEGACO, and it is the fact that these services are based on standardised open technologies, reachable through the internet, that make it vulnerable to attack (Elhert et. al. 2008). This research paper will be specifically investigating SIP based DoS attacks on VoIP network infrastructures. It will look at the most up to date research, from various sources that have been carried out in this area, and the different methods that have been used to detect and mitigate these types of DoS attacks. Each method will be considered and evaluated, in order to conclude which offers the most secure method of security.

2 Background

Previous to using VoIP, the common method of communication was to use the Public Switched Telephone Network (PSTN) and this is regarded as more secure than VoIP, as it utilizes closed network architecture. It is for this reason that PSTN suffers from minimal security flaws, as an attack is very difficult unless the attacker has physical access to the network (Geneiatakis et. al. 2009). However, due to the reduction in cost it is likely that VoIP will eventually outstrip PSTN, although at the moment it is hard to see how VoIP could be used exclusively as it still has failings in certain areas. VoIP is dependent on power from the wall, so if the power goes, then so does broadband and the network. Unlike PSTN, calls using VoIP cannot be guaranteed (Bodhani, 2011).

This is one of the reasons why the two technologies are currently used in combination, but the downside of this is that the previously secure PSTN can now become vulnerable to VoIP attacks. The challenge for security researchers would be to find a way of making VoIP as secure as PSTN, but as mentioned earlier there are many security flaws associated with VoIP, mak-

ing it very difficult for a single method of security to cover all angles of attack. The subject of SIP based DoS attacks to be discussed in this paper is just one small area of VoIP security.
Researchers in this area have proposed several different mechanisms for protection against these types of attack. Elhert et. al. (2009) presents a survey paper showing 20 of the most recently published research works, which address SIP related DoS problems. However, according to Hsien-Ming H et. al. (2011) several of the intrusion detection systems and prevention mechanisms discussed in this survey paper are now regarded as unsuitable for use on today's internet.

3 Two layer architecture defence

Ehlert et. al. (2008) has proposed a two layer security architecture to prevent SIP based Dos attacks on VoIP infrastructures. They claim that their method is designed to contend with flooding attacks, malformed message attacks and irresolvable DNS attacks.

3.1 Method

The first line of defence is to be a Bastion Host (firewall) which is placed at the entry point of the VoIP network. This Bastion host includes an Intrusion Detection System (IDS) which is capable of filtering malicious messages and also updating the firewall according to these malformed packets. In this case the IDS implemented is Snort-IDS and has been extended so that it is capable of use on a VoIP network. The Bastion host is also able to prevent DoS flooding attacks.

The second line of defence is protecting the SIP proxy server, as this is the most likely point on the network an attacker will attempt to render inoperable using a DoS attack. A Deep Packet Inspection Module (DPI) is implemented on the proxy, which scans incoming packets to determine if they are malformed. A method of storing incoming DNS requests is deployed using an enhanced caching daemon, and this will prevent lockdown of the server if it receives too many unresolvable DNS requests.

In order to be able to react immediately, any possible attack is logged and sent to the operator console. This provides the admin with instant notification of an attack.

3.2 Experiments and results

To test the proposed method, the author uses different testbed setups to take the actual measurements from the different test attacks (stated in section 3). Although, the testbed and defence architecture do share some of the same components including the SIP proxy and the attacking tool.

Figures 1 and 2 present the results of high volumes of SIP REGISTER and SIP INVITE packets being generated to target the firewall (Bastion host) internal and external interfaces.

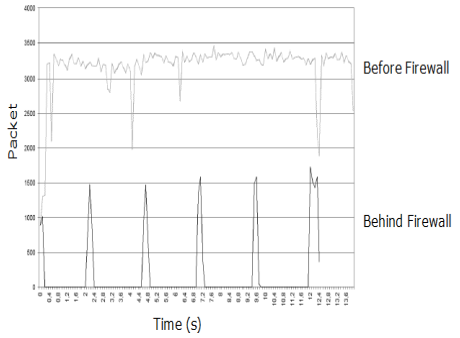


Figure 1 Defence against SIP REGISTER message flooding (Elhert et. al. 2008)

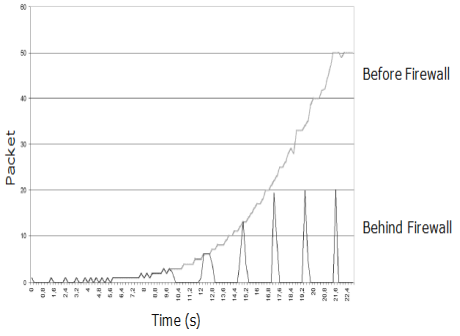


Figure 2 Flooding by increasing the rate of SIP INVITE messages (Elhert et. al. 2008)

Figures 1 and 2 demonstrate the detection of high volumes of traffic on the external interface and less traffic on the internal interface demonstrating that the methods configured on the firewall are successful.

To test for malformed message attacks the SIP proxy was fed traffic that had been copied from a real SIP based VoIP network in order to simulate regular background traffic. The proposed method will inspect incoming traffic and raise an alarm if a SIP message is identified as malicious or malformed. The author presents the results of experiments in Figure 3 based on the scenarios shown in Table 1 and Table 2.

DESCRIPTIONS OF THE EMPLOYED REAL SCENARIOS

Scenario Number	Scenario-Description
Scenario 1	This scenario utilizes the PROTOS test to create malformed messages
Scenario 2	This scenario utilizes specific malformed messages that contain errors in the first line only
Scenario 3	This scenario utilizes specific malformed messages that contain errors in one header
Scenario 4	This scenario utilizes only real-life traffic

Table 1 (Elhert et. al. 2008)

MALICIOUS MESSAGE FALSE ALARM ALERTS

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Processed Mesg	42036	310000	40968	56308
False Positives	246	198	274	209
Probability	0,005852	0,000639	0,00668	0,003712

Table 2 (Elhert et. al. 2008)

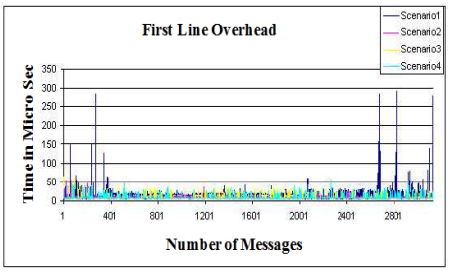


Figure 3 First Line inspection overhead for scenarios 1 to 4 (Elhert et. al. 2008)

To test for a DNS attack the SIP proxy is attacked by continuously sending messages that contain unresolvable domain names. The proposed method is tested with no DNS cache installed and then tested again with the proposed DNS caching solution. Results are shown in Figures 4 and 5.

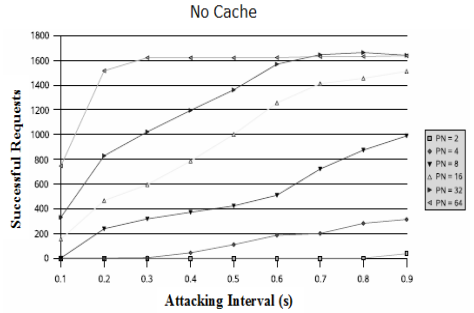


Figure 4 (Elhert et. al. 2008)

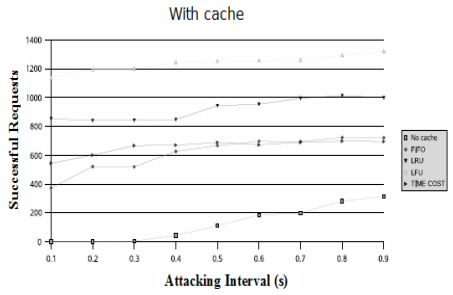


Figure 5 (Elhert et. al. 2008)

3.3 Conclusions

It should be noted that the author of this proposed method is not creating a new algorithm for attack detection/prevention, but is using a combination of already existing tools and then setting them up and configuring them to contend with attacks. The security tools used were Snort, Prelude framework and its database, SnortSam + IPtables/ IPpset firewall.

The results from the flooding experiments show that the methods deployed are successful in the detection of attacks using SIP INVITE and REGISTER messages. However, the method was not tested using other SIP messages: ACK, CANCEL, BYE and OPTIONS, so it is unknown how the proposed method would react if attacks were made using these SIP messages. The author does not show how flooding attacks will be prevented, with prevention being defined as actually stopping the attack whilst it is taking place. On detection of an attack, an alarm is sent to the operator console, but this would then then rely on further interaction from a system admin to stop the attack. It is not actually the proposed method that is preventing the attack, as the prevention relies on human interaction.

For the experiments testing for malformed messages, the author focuses on false alarms that are generated, and also the overhead on the system that is created by using the proposed method. Test results showed that false alarms were quite minimal, and the author did give an explanation as to the reason they would exist. It was also demonstrated that the detection tool used did not introduce a significant amount of overhead to the network.

However, clear results are not presented as to what happens when an actual malformed message is used to attack the system.

The DNS attack essentially is another type of flooding attack and the author's method of DNS caching (Figure 5) is successful compared with a system not using DNS caching (Figure 4). Their results also showed different levels of performance dependent on the caching strategy deployed, as their experiments were carried out using several different scenarios. The author has demonstrated a clear difference between the results of using the method of DNS caching solution as opposed to using no DNS cache.

However, a more effective measure of comparison would be to test their proposed method against another method that also used DNS caching rather than against one that didn't use any at all.

The author didn't consider message flow tampering described by Elhert et. al. (2009) as one of the three methods of SIP based DoS attacks. This is an area of SIP based DoS attacks the author has left unexplored.

The author's proposed two layer architecture method of attack prevention has presented results showing clearly how the proposed method is effective in detecting flooding attacks and DNS attacks. It may well be that it is also successful in detecting malformed message attacks, but conclusive evidence of this was not presented by their results. It is not made clear if the two layers of the architecture are to be used in conjunction with each other. No tests were carried that demonstrated how the second level of protection would perform if it was attacked by an intruder that had passed through the first level of protection. This would have demonstrated a real life situation. The experiments were direct attacks on each of the two levels as two separate entities.

4 Utilizing Bloom filters

Geneiatakis et. al. (2009) proposes the use of bloom filters to detect SIP based DoS flooding attacks.

4.1 Experiments and results

In order to show the effectiveness of the proposed method the architecture illustrated in Figure 6 is used to test the proposed method. The method is tested using different scenarios to illustrate its effectiveness under different conditions. These scenarios are shown in Table 3.

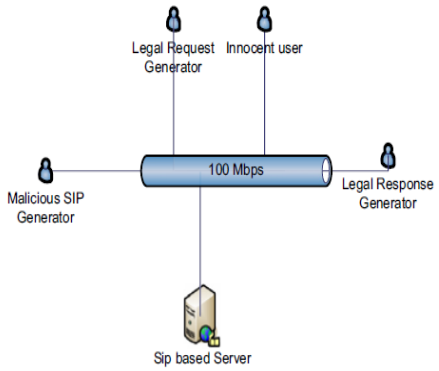


Figure 6 Attack architecture (Geneiatakis et. al. 2009)

Scenario name	Scenario description
Scenario 1 (S1)	In this scenario the "*legal request generator*" generates (serially) requests (at a pace of 1 req/μs), while the corresponding responses are generated by the "*legal response generator*".
Scenario 2 (S2)	In this scenario the malicious user generates requests (at a pace of 1 req/10 μs) that are addressed to an innocent user who tries to respond to all of them.
Scenario 3 (S3)	In this scenario the malicious user generates requests (at a pace of 1 req/10 μs) that are addressed through the proxy to clients belonging to non-existing domains.

Table 3 (Geneiatakis et. al. 2009)

The results of these scenarios are shown in Figures 7, 8 and 9.

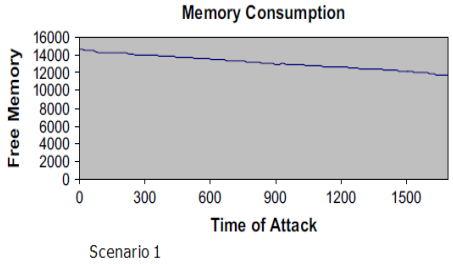


Figure 7 (Geneiatakis et. al. 2009)

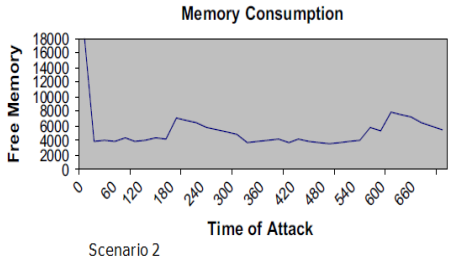


Figure 8 (Geneiatakis et. al. 2009)

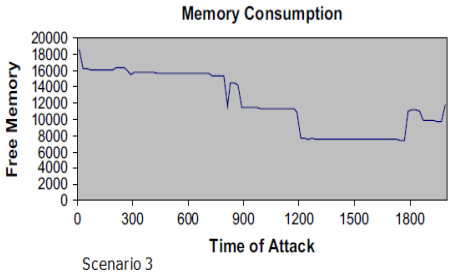


Figure 9 (Geneiatakis et. al. 2009)

4.2 Conclusions

The results presented by the author show a clear difference in the behaviour of the network memory between no attack being present (Figure 7), and when an attack is active (Figure 8 and 9). The author clearly demonstrates that the proposed method of using bloom filters is effective in detecting DoS attacks against SIP. It is also claimed that the detection mechanism is effective in preventing the attack. It does this by triggering a specific administrative task that has been pre-configured as a security policy. When an attack is detected it will cause traffic to be dropped and therefore prevent the attack. However, the author does not provide any evidence to show that the proposed prevention method is effective. No statistical or graphical evidence of this claim is shown.

The author provides detailed technical information on how their method works and has also considered the effects of a DoS attack on both a SIP proxy server (results shown in Figure 9) and also the end points or User Agents (results shown in Figure 8).

The author's method of detection is to display the effect of an attack on the system memory. Details are also provided of the affect these attacks have on the system CPU load. Graphical data is not provided but it is stated that the CPU load is 2-3 times larger (when under attack) when the proposed method is not in use as opposed to when it is. Details are not provided on how such an attack would affect the Bandwidth usage, which along with memory and CPU is the main resource required for a SIP server to operate (Sisalem 2006).

A distributed DoS (DDoS) attack is also considered. The author describes in technical detail how the proposed method is used to detect a DDoS attack, but no experiments were carried out, and therefore no data is provided to substantiate this claim.

5 Comparison

On the evidence provided by Ehlert et. al. (2008) and Geneiatakis et. al. (2009) it is difficult to conclude that one method is better than the other. Both methods have been tested under different conditions. To accurately conclude that one method is better than another, they would both need to be tested under exactly the same conditions. It is only then that a true comparison could be made.

However, it could be said that the method proposed by Geneiatakis et. al. (2009) offers a wider range of protection, as the author has shown evidence that their method takes into account security of the User Agents (UA's), while Ehlert et. al. (2008) does not. It is true that a UA is a less likely point of attack than a SIP server, but nevertheless it is still a point of attack. Geneiatakis et. al. (2009) has also considered DDoS attacks while Ehlert et. al. (2008) provides no evidence of research into this area.

Both authors have provided evidence that their proposed method is effective in detecting an

attack, and while both authors have made claims that their method is also capable of preventing an attack, neither has provided evidence to substantiate this.

Evidence was also provided that both detection methods do not create any significant overhead on a network when in use but this may not be the case with the prevention method. Any method of prevention would certainly require more processing power, but again neither author provided statistics to show this.

Neither of the authors provided evidence that their methods are scalable, or if they could be used in conjunction with different types of hardware and software.

6 Overall Conclusions

According to Palmieri and Fiore (2009) SIP is quite a difficult protocol to secure. There are many factors to take into account as DoS attacks are only one area of vulnerability. There are also security issues with the transport protocols that SIP uses such as TCP, TLS or IPsec (Rebahi et. al. 2011). Sip is in clear text format, so encryption also needs to be considered, but this paper is focussed on DoS attacks (Ehlert et. al. 2008).

Both of the methods described in this paper could be successfully deployed on a live network as statistical evidence shows that they are capable of detecting DoS attacks. The fact that evidence of prevention is not provided is not necessarily a drawback. Use of a prevention method would be dependent on the requirements and capabilities of the network. Some networks may not wish to deploy a prevention method because of the extra overhead that this will create.

It is also unclear as to how scalable either of the methods is, so a network that is expecting to increase in size may choose not to use either method. There are many other methods of detection and prevention available and Gold (2012) encourages the use of integrated security systems. However, this could create the issue that the different systems may be proprietary, so therefore not necessarily compatible, and also they may be open source which allows cyber-criminals access to the source code (Gold 2012).

The research carried out by Ehlert et. al. (2008) and Geneiatakis et. al. (2009) shows clearly that there is a necessity for further work in the area of protecting VoIP infrastructures from SIP based DoS attacks. The research looked at in the production of this paper has shown that, at present there are possible solutions for the detection of these attacks (although evidence indicates that this success is based on the condition of the network), but there is no conclusive method of prevention.

Responsibility for protection against SIP based DoS attacks is not limited to the local network. A major German communications service provider called Arcor claim that it should be compulsory for service providers to provide protection against DoS attacks by using Session Border Controllers (Bessis et. al. 2011). However, this would also increase the cost of protection, and may be outside of the budget of some networks.

VoIP will certainly be used more and more in the future as opposed to PSTN, and for this reason it will become a more attractive target for cyber-criminals. Geneiatakis et. al. (2008) states that, without doubt, in the very near future, attacks against VoIP services will become a very common phenomenon. Cyber-criminals will put more effort into attacking the potential weaknesses of SIP, so security methods will need to be constantly adapting and improving to stay ahead of the latest attack techniques.

References

Bessis T, Gurbani V and Rana A, 2011, 'Session Initiation Protocol firewall for the IP Multimedia Subsystem core', *Bell Labs Technical Journal*, Vol. 15 Issue 4, pages 169-187.

Bodhani A, 2011, '**VOIP** - voicing concerns', *Engineering & Technology*, Vol. 6 Issue 7, pages 76-79.

Bradbury D, 2007, 'The security challenges inherent in VoIP', *Computers & Security*, Vol. 26 Issue 7, pages 485-487.

Edelson E, 2005, 'VoIP: Voice over IP: security pitfalls', *Network Security*, Issue 2, pages 4 – 7.

Ehlert S, Zhang G, Geneiatakis D, Kambourakis G, Dagiuklas T, Markl J and Sisalem D, 2008, 'Two layer Denial of Service Prevention on SIP VoIP infrastructures', *Computer Communications*, Vol. 31 Issue 10, pages 2443-2456.

Ehlert S, Geneiatakis D and Magedanz T, 2009, 'Survey of network security systems to counter SIP-based denial-of-service attacks', *Computers & Security*, Vol. 29 Issue 2, pages 225-243.

Geneiatakis D, Vrakas N and Lambrinoudakis C, 2009, 'Utilizing bloom filters for detecting flooding attacks against SIP based services', *Computers & Security*, Vol. 28 Issue7, pages 578-591.

Gold S, 2012, 'Securing VoIP', *Network Security*, Vol. 2012 Issue 3, pages 14–17.

Hsien-Ming Hsu, Yeali S. Sun and Meng Chang Chen, 2011, 'Collaborative scheme for VoIP traceback', *Digital Investigation*, Vol. 7 Issue 3, pages 185-195.

Palmieri F and Fiore U, 2009, 'Providing true end-to-end security in converged voice over IP infrastructures', *Computers and Security*, Vol. 28 Issues 6, pages 433-449.

Rebahi Y, Nassar M, Magedanz T and Festor O, 2011, 'A survey on fraud and service misuse in voice over IP (VoIP) networks', *Next Generation Networks, Information Security Technical Report*, Vol. 16 Issue 1, pages 12-19.

Sisalem D, Kuthan J and Ehlert S, 2006, 'Denial of Service Attacks Targeting a SIP VoIP Infrastructure: Attack Scenarios and Prevention Mechanisms', *IEEE Network*, Vol. 20 Issue 5, pages 26-31.

An Evaluation of Current Database Encryption Security Research

Ohale Chidiebere

Abstract

Databases are vulnerable to theft of sensitive data because adversaries can exploit database server, application server to gain access to private data and because malicious or curious database administrator may capture and leak data. The aim of this paper is to analyze and critically evaluate some of the current research methods designed for protecting sensitive data using encryption and after this, documented evidence will be used to summarize findings and offer a proposed solution.

1 Introduction

Database encryption is one way the information or data of an enterprise can be protected against illegal access, because information or data are the major assets of any enterprise, so the security of the database of an enterprise is a serious challenge. Hence protecting the confidential and sensitive data stored in the database is very essential. It has been observed that encryption provides confidentiality of data in a database Iqra et. al. (2012) thus making the database secure from illegal access or threat at any level. Based on current research there is conjunctive keyword search over encrypted database where the server searches for all document containing each keyword using various techniques like the single keyword search which checks for the intersection set of all the document involved in the search and then returns the corresponding results to the user (Jin and Dong 2011). However Dan et. al. (2012) presented a method called BLESS which was used for object level encryption for securing objects based storage systems, which improved the overall performance of the database system. Thus a database encryption scheme was developed with the combination of conventional encryption and public key encryption which resisted attacks on encrypted data (Gang et. al. 2006). A new paradigm for database encryption was proposed by Lianzhong and Jingfen (2008) in which database encryption can be provided as a service to applications with seamless access to encrypted database. Masayuki et. al. (2011) proposed an efficient symmetric searchable encryption to achieve indistinguishability of indexes and trapdoors. Qing et. al. (2011) presented a Huffman table which was an encryption algorithm used in

protecting multimedia contents. Wu and Ming (2010) researched on hybrid cryptography encryption program where they merged IDEA and RSA which overcomes the problem of key management and database encryption.

In this paper we will evaluate different methods that have been proposed through series of current research to see how they have performed towards encryption of database and to find out the best method of securing sensitive data in databases using encryption.

2 CryptDB

Raluca et. al. (2012) carried out a research on CryptDB which was used to process query on encrypted database. CryptDB provides confidentiality for applications that use database management systems by using user defined functions to perform cryptographic operations in the Database Management Sytems. It handles two threats which are, an adversary who gains access to the database server and tries to learn confidential data and an adversary who gains complete control of the database server as well as the application server.

Six encryption methods were used namely Random(RND), Deterministic(DET), Order

Preserving Encryption(OPE), Homomorphic Encryption(HOM), Join and OPE-join, Word search(SEARCH).

Random (RND) protects the plaintext from being attacked and also rejects computations performed on ciphertext (Raluca et. al. 2012). In Deterministic(DET) equality checks are performed which means that the server are allowed to learn which encrypted values corresponds to the same data value which generates same ciphertext for the same plaintext. In the Order-Preserving Encryption (OPE) it allows the server to perform range of queries when given encrypted constants corresponding to the range. In Homomorphic (HOM) the server performs computations on encrypted data but with final results decrypted at the proxy. In Join and OPE-join the servers are allowed to determine repeating values between two attributes, while in Word Search the servers are allowed to search on encrypted text but are not allowed to decrypt the text.

Experiments were conducted to check the functionality, security and performance of CryptDB. The experiment of CryptDB on security and performance were performed on five different applications and one large trace, namely an open-source web forum application known as PhBB, a conference management system also known as HotCRP, the admission application from MIT EECS known as grad-apply, an e-medical record application that stores the medical data of patient known as OpenEMR, TPC-C, also known as industry-standard database and an SQl server located at MIT where traces of SQL queries where gathered. The query trace was achieved within a duration of about ten days where approximately 126 million queries over 1193 databases and 18,162 queries were recorded (Raluca et. al. 2012).

Application	Consider for enc.	Needs plaintext	Non-plaintext cols. with MinEnc: RND/SEARCH	DET	OPE
phpBB	23	0	21	1	1
HotCRP	22	0	19	1	2
grad-apply	103	0	95	6	2
OpenEMR	566	7	528	12	19
TPC-C	92	0	65	19	8
Trace from sql.mit.edu	128,840	1094	80,403	34,212	13,131
...with in-proxy processing	128,840	571	84,406	35,350	8513

Figure 1. This Diagram shows the results obtained from the Functional and Security experiments of CryptDB (Raluca et. al. 2012).

Experiment on performance was carried out were 2.4 GHz Intel Xeon E5620 4-core processors and 12GB of RAM were used to run MySQL 5.1.54 server, while a machine with eight 2.4 GHz AMD Opteron 8431 6-core processors and 64 GB of RAM were used to run the CryptDB proxy and the clients, both were connected via a shared Gigabit Ethernet network. The performance test was carried out on MySQl server and CryptDB proxy where queries per seconds was compared with the time interval from when the query is issued to when the results are achieved.

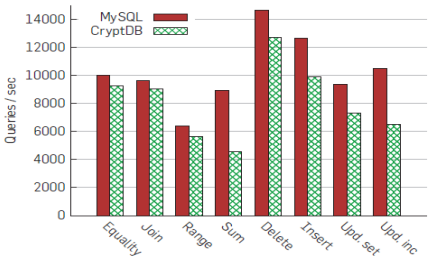


Figure 2. This diagram shows the performance of queries on MySQl server and CryptDB proxy (Raluca et. al. 2012).

In conclusion, CryptDB handled the threats of an adversary who gains access to the database server and tries to learn confidential data and an adversary who gains complete control of the database server as well as the application server by providing a strong level of confidentiality. It also achieved a modest performance overhead by handling a wide range of queries.

2.1 Evaluation of CryptDB and Conclusion

The authors of this research have demonstrated in the paper a successful and efficient approach for providing confidentiality for applications that uses the database management systems, this is because different levels of strong encryption methods were enforced into CryptDB thereby ensuring that confidential data were not lost, with the number of devised experiment the researchers conducted from different applications; they were able to gather enough information to reach a definite conclusion. From this analysis CryptDB will work in all database management system in terms of functionality, security and performance.

3 Transparent Data Encryption

Transparent data encryption is a technology that encrypts data on the network, storage device were the database is stored and on backup media. Transparent data encryption shields the table, column and tablespace in a database from being attacked.When the data of any organization is misused, then the operation of that organization will be terribly affected. This is because data is vulnerable to wide range of menace like Excessive Privilege Abuse, Legitimate Privilege Abuse, Weak authentication and Backup Data Exposure (Anwar and Riyazuddin 2011). Transparent Data Encryption will be used to reduce these threats. Transparent data encryption is used to prevent unauthorized access to sensitive data, manage users, facilitate privacy managements, encrypt and decrypt data and log files.

The following methods are used by transparent data encryption:

Authentication - allows access to the database by identifying the users accurately. This involves taking more information than the usual username and password.

Validation - Ensures that the identity of a sender is true, thereby proving that a column, table and tablespace have not been modified. Once validation is enforced the user will be very sure that the information is coming from a trusted source, which means that the information has not been modified
.

Data protection - ensures that the information is protected. This method is used to protect information contained in the column, table and tablespace.

Database Encryption Key (DEK) - located at the SQL layer used for encrypting data before it is written to disk and also decrypts the data before returning it to the application.

Transparent Data Encryption in the database in Microsoft Server 2008 is performed at the page level where the pages in the encrypted database are encrypted before they are written to disk and also decrypted while read into the memory. While the SQL server is being setup a key called the Service Master key is created which is used to encrypt the Database Master Key for the master Database. The Database Master Key of the master Database creates a certificate which encrypts the database encryption key in the user database (Anwar and Riyazuddin 2011).

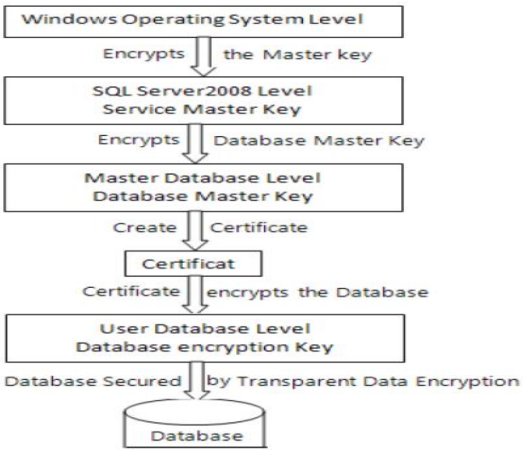


Figure 3. This diagram shows the steps of Transparent Data Encryption in Microsoft Server 2008 (Anwar and Riyazuddin 2011).

In conclusion, Data in transit were safeguarded by transparent data encryption. The transparent data encryption in Microsoft SQL server protected sensitive data on disk

drives as well as backup media from unauthorized access which reduced the impact of the backup media being stolen. The transparent data encryption in this research were successfully created, implemented and tested thoroughly, it was used on few computers and was found to be very effective.

3.1 Evaluation and Conclusion

It has not been confirmed that transparent data encryption achieved the conclusion reached by the authors of this paper because there were no experiments carried out on two of the methods mentioned which are authentication and validation. The experiment conducted was on data security, they used one database management system where as there are numerous database systems were transparent data encryption would have been tested on which they would have obtained reasonable results. This research is questionable because there is no concrete evidence to justify the results reached by the researchers. Although this does give the opportunity for the team or a new team to continue the research in the future.

4 Encryption-Based Multilevel Model for Database Management Systems

Ahmed et. al. (2012) carried out a research on Encryption-Based Multilevel Model for database management systems. The purpose was to merge encryption algorithm with multilevel relational model which will provide additional security layer on the multilevel security layer on the database to provide high level of security as well as solve the problems associated with the multilevel relational model. Two methods were used in this paper to solve the problems associated with the multilevel relational model, namely the multilevel database security method and the encryption-based multilevel method.

In the method, whilst a level is created by the database administrator to be used in the multilevel database, a symmetric key is created by the database engine automatically that takes care of this level. The key is usually stored in the multilevel database which is used for the encryption and decryption of data element that are classified to the level associated to the symmetric key.

Their method was implemented in a database where they used the Data Manipulation Language which are INSERT, DELETE, UPDATE and SELECT to experiment.

They conducted a performance test with a CPU speed of 2.2GHz, 3GB of RAM, and 320GB of hard disk size. Microsoft SQL server 2008 R2 was used for the implementation of the proposed method. These are their performance report after their experiments.

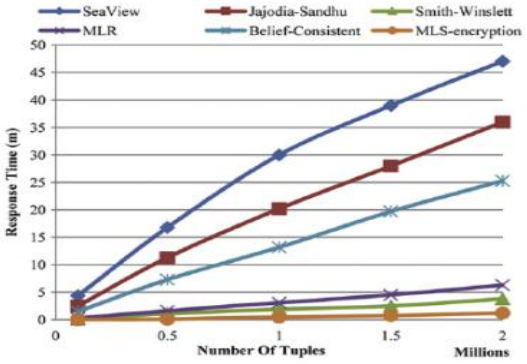


Figure 4. This diagram shows the performance of SELECT statement during the experiment (Ahmed et. al. 2012).

The number of selected rows was varied to 100,000, 500,000, 1,000,000, 1,500,000 and 2,000,000, fixed the number of column at 3, and fixed the number of security levels at 4 (Ahmed et. al. 2012).

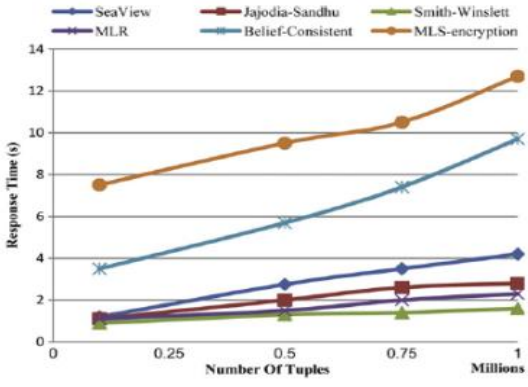


Figure 5. This diagram shows the performance of UPDATE statement during the experiment (Ahmed et. al. 2012).

The number of updated rows was varied to 100,000, 500,000, 750,000, and 1,000,000, fixed the number of column at 3, and fixed the number of security levels at 4 (Ahmed et. al. 2012).

Ahmed et. al. (2012) concluded that with encryption algorithm merged with multilevel relational model, the performance of retrieving data in the Data Manipulation Language improved.

4.1 Evaluation and Conclusion

Although providing an initial successful implementation of their method, the authors did claim that their method improved the performance of retrieving data but they deviated by not specifying the problems of the multilevel relational model and solving them as they proposed. Based on their conclusion there are many questions remaining as there method was only tested on one specific machine against one specific database. Unlike the previous research, Raluca et. al. (2012) were they used their system on numerous machines and so could confidently state that they had enough information to reach a definite conclusion.

5 Database Encryption using TSFS Algorithm

This is a research by Manivannan and Sujarani (2010) on database encryption techniques using Transposition, Substitution, Folding and Shifting (TSFS) algorithm with three keys for encrypting only sensitive data in database. The idea behind this is to block database administrators who are insiders that have DBA privileges to attack and compromise the database and steal sensitive data since the database systems are usually deployed deep inside the organisations network. In TSFS three keys are used for encrypting and decrypting information, the three keys are further expanded to into twelve sub keys via Key Expansion Technique which provides an efficient and more secured database. Four methods where used namely Transposition, Substitution, Folding, and Shifting (TSFS).

In transposition cipher, symbols are reordered, by changing the position of cipher text. Substitution ciphers deals with the replacement of symbols with another symbol. Folding is responsible for shuffling data from one position to another position. In Shifting each digit of the number is replaced by its position within its array element (Manivannan and Sujarani 2010).

Manivannan and Sujarani (2010) talked about the strength and security of TSFS algorithm. In the strength, the number of keys were more which increased the key combination that made guessing very difficult. It was also said that three plain texts which are numeric, character and alphanumeric have their corresponding ciphertexts; with this only the sensitive data in the database were encrypted.

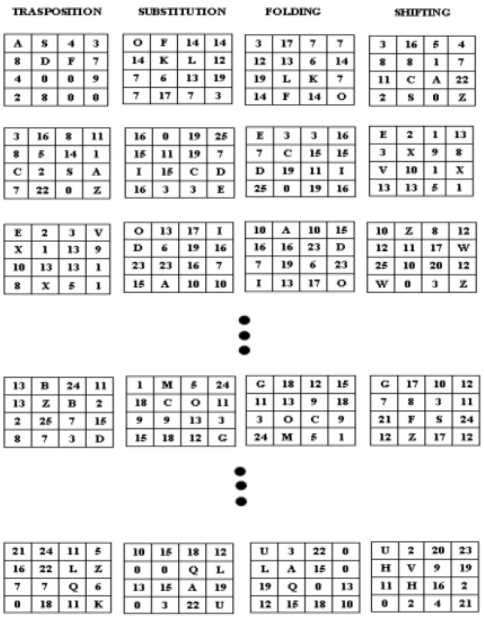


Figure 6. This diagram shows the strength of the TSFS algorithm (Manivannan and Sujarani 2010).

In the security analysis of TSFS algorithm, the data types of plaintext and ciphertexts are the same, so it is very difficult for an attacker who has access to sensitive data to read from the database, this is because the attacker will find it difficult to identify whether the information in the database is

encrypted or not encrypted (Manivannan and Sujarani 2010).

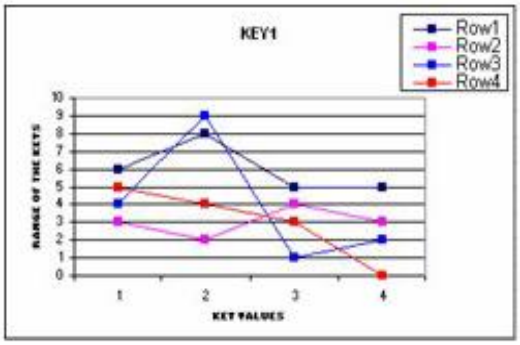


Figure 7. This diagram shows the security analysis of TSFS algorithm (Manivannan and Sujarani 2010)

Manivannan and Sujarani (2010) claimed that their algorithm was efficient because sensitive data residing in the database were encrypted. They also said that security risk can be eliminated if the storage of data is encrypted before they are stored into the database which will reduce the security issues found on databases using three cryptographic keys for the protection of data.

5.1 Evaluation and Conclusion

The solution presented by Manivannan and Sujarani (2010) to secure sensitive data in the database was successful but they did not present the experiments they carried out, what they did was explaining the outcome of the experiment within the text of the paper. They made obvious the efficiency of their algorithm by the rigorous cipher methods they implemented. Although the experiment was not published but the results they published backed up their claims. However based on the methods implemented in the TSFS algorithm, it is seen that TSFS can confidently secure sensitive data residing in the database with ease. They are also aware that the TSFS algorithm will not support symbols which open up an avenue for the enhancement of the algorithm in the future (Manivannan and Sujarani 2010).

6 Conclusions

In this paper current database encryption securities have been carefully evaluated. Although it may be impossible to completely wipe-out all threats posed by various adversary. The purpose of these solutions is to reduce the amount of attacks and threats to database systems. The research which has been looked at is very interesting in that they are looking at different methods of ensuring encryption on the database for the security of sensitive data.

Generally the current research looked at was of good standard, except for some as in the case of Transparent Data Encryption where they ignored carrying out experiment on two of the methods they proposed, they have also not got enough depth to it as repeatable testing was not performed, unlike the research by Raluca et. al. (2012) where they conducted various tests from different locations and different systems which backed up their claims. On a positive outlook for Transparent Data Encryption, if they could carry out experiments on the two methods as well as expand their mode of experiments into using different database management systems for testing they will be able to achieve concrete results to justify their claims.

The research carried out on CryptDb looks very promising, as the series of methods they implemented to reduce attacks to the database can be looked at as great improvement compared to Ahmed et. al. (2012) were two methods was used. Also what is really promising is the number of applications they used from different locations to carry out their experiment. This means that the testing done was of good standard and with good recorded results. The research by Raluca et. al. (2012) had a fair and unbiased experiment which had relevant results.

The research carried out by Ahmed et. al. (2012) seem to be a success, the noticeable problem is that there have been no definitive repeatable testing using different machines and different database management

systems, all that seem to have been done is they have tested their method on one machine with a specific database to reach their claims. With the evidence presented within this paper it is has been noticed that a lot more testing is needed for more concrete results to be obtained.

The research on TSFS algorithm was very interesting as their method will help to confidently protect sensitive data from adversary. The results they presented were of a high standard as they were able to illustrate that from the cipher method they implemented. If the results obtained can be replicated elsewhere on other systems it will prove very promising.

One thing still stands out with all the research, the need for more research and experiments is required to improve on some of the methods being proposed. One of the possible ways of achieving this could be to add different research together to see if two or more methods could get to a solution that is being looked for. Based on the methods, experiments, results and conclusion reached by Raluca et. al. (2012) and the results obtained by Manivannan and Sujarani (2010) it can be seen that if CryptDB is merged with TSFS algorithm it will yield a posible solution for securing sensitive data in the database since CryptDB has been applied in number of different applications from different locations and TSFS has been tested and tangible results obtained we can observe a better performance than Transparent Data Encryption and Encryption-based multilevel model for Database Management Systems with questionable results.

References

Ahmed I. Sallam, El-Sayed El-Rabaie and Osama S. Faragallah, 2012, 'Encryption-based multilevel model for DBMS' *Journal of Computers & Security*, Volume 31, Issue 4, Pages 437–446.

Anwar Pasha Abdul Gafoor Deshmukh and Riyazuddin Qureshi, 2011, 'Transparent Data Encryption- Solution for Security of Database Contents', *International Journal of Advanced Computer Science and Applications*, Volume 2, Number 3.

Dan Feng, Junjian Chen, Jingning Liu and Zhikun Wang, 2012, 'BLESS: Object level encryption security for object-based storage system', *Journal on Mathematical and Computer Modelling*, Volume 55, Issue 1-2, Pages 188–197.

Gang Chen, Ke Chen and Jinxiang Dong, 2006, 'A Database Encryption Scheme for Enhanced Security and Easy Sharing', *Proceedings of the 10th International Conference on Computer Supported Cooperative Work in Design*, Pages 1-6, ISBN: 1-4244-0164-X.

Iqra Basharat, Farooque Azam and Abdul Wahab Muzaffar, 2012, 'Database Security and Encryption: A Survey Study' *International Journal of Computer Applications*, Volume 47, Number 12.

Jin Wook Byun and Dong Hoon Lee, 2011, 'On a security model of conjunctive keyword search over encrypted relational database' *Journal of Systems and Software*, Volume 84, Issue 8, Pages 1364–1372.

Lianzhong Liu and Jingfen Gai, 2008, 'A New Lightweight Database Encryption Scheme Transparent to Applications', *International Conference on Industrial Informatics*, Pages 135-140, ISBN: 978-1-4244-2170-1.

Manivannan D. and Sujarani R., 2010**, '**Light Weight and Secure Database Encryption Using TSFS Algorithm**'** *International Conference on Computing, Communication and Networking Technologies*, Pages 1-7, ISBN: 978-1-4244-6591-0.

Masayuki Yoshino, Ken Naganuma and Hisayoshi Satoh, 2011, 'Symmetric Searchable Encryption for Database Applications', *Proceedings of the 2011 14th International Conference on Network-Based Information Systems*, Pages 657-662, ISBN: 978-0-7695-4458-8.

Qing Zhou, Kwok-wo Wong, Xiaofeng Liao and Yue Hu, 2011, 'On the security of

multiple Huffman table based encryption', *Journal of Visual Communication and Image Representation*, Volume 22, Issue 1, Pages 85–92.

Raluca Ada Popa, Catherine M.S. Redfield, Nickolai Zeldovich and Hari Balakrishnan, 2012, 'CryptDB: Processing Queries on an Encrypted Database', *Magazine Communications of the ACM*, Volume 55, Issue 9, Pages 103-111.

Wu Xing-hui and Ming Xiu-jun, 2010, 'Research of the Database Encryption Technique Based on Hybrid Cryptography', *International Symposium on Computational Intelligence and Design*, Pages 68 – 71, ISBN: 978-1-4244-8094-4

A Critical Appreciation of Current SQL Injection Detection Methods

Lee David Glynn

Abstract

SQL Injection attacks have become one of the most common threats that affect a variety of applications written for the World Wide Web. This paper is focused on the research being conducted, which proposes a variety of different methods for detecting these SQL injection attacks. Many of these come with their restrictions, and comparisons have been made to demonstrate the most competent method.

1 Introduction

There is no doubt that for businesses today the internet has become a necessity, whether for competitive advantage or basic operations. Online security is the main challenging problems to consider for any business or organization that utilizes online functionality for a successful ecommerce service.

Research shows that consumers are more apprehensive about ecommerce security due to the increase in SQL injection attacks. SQL injection has become an increasingly common technique used to expose the flaws in the databases typically used by ecommerce web applications. A report by Clearinghouse suggests that 312 million data records have been lost since 2005, and 83% of hacking-related data breaches were executed via SQL injection attacks (Cruz, 2012).

Other research being done has focused mainly on the prevention of these attacks. Patel, Mohammed and Soni (2011) provide an array of diverse detection mechanisms against SQL injection attacks (SQLIA's) such as taking user input from predefined choices, bind variable mechanisms and parameterised statements.
This research paper will focus on accurately detecting SQL injection attacks in order to determine the most efficient methods.

2 Proposed Approaches for SQL Injection Detection

The detection of SQL injection attacks is the first line in the defence of web applications. There are two paramount concepts to consider for the protection of web applications. Primarily the need for a mechanism incorporated within these web applications to detect and categorize SQL injection attacks. Secondly, an understanding of the vulnerabilities linked to SQL injection attacks.

Several techniques are already in operation to prevent the vulnerabilities that can lead to the occurrence of a SQL injection attack. "Many existing techniques, such as filtering, information-flow analysis, penetration testing, and defensive coding, can detect and prevent a subset of the vulnerabilities that lead to SQLIAs" (Halfond & Orso, 2005).

Numerous research papers have proposed many different methods for the detection of SQL injection vulnerabilities such as CANDID, this approach attempts to "dynamically mine the programmer-intended query structure on any input" (Bisht, et al., 2010). These frameworks include static analysis tools, dynamic analysis tools, combined static and dynamic analysis tools and black-box testing tools.

2.1 Static Analysis

Fu, et al., (2007) have conducted research that has outlined a static analysis framework entitled SAFELI that would be used to identify the SQL

injection vulnerabilities within web applications. This framework intends to detect the vulnerabilities at the compile time of ASP.net application, by examining the byte code using symbolic execution. The approach has two distinct offerings. Firstly the approach proposes a white box analysis of the byte code of applications and secondly employs a hybrid constraint to find out the user input that could lead to the breach in the data security. SAFELI analyses the byte code using the white-box techniques that relies primarily on a string analysis procedure that handles the most popular operations such as Boolean, integer and string variables. Once insertion points within the code are identified SAFELI replaces all uninitialized variables with symbolic constraints. SAFELI also includes a Test Case Generator that activates when symbolic execution reaches any hotspots within the code. This randomly posts some values into HTML text fields through a web page back to a web server backend database and will then scrutinize the results. Any vulnerability that is detected is highlighted with a step by step error trace.

Fu et al., (2007) provide insufficient evidence of experimental analysis or results therefore they can't justify any clams they make. However they do illustrate the proposed model through the use of extensive examples to show the different stages of symbolic execution as well as formal proofs. The quality of this approaches generated test cases solely depend on the completeness of the attack pattern library. This part needs input from human testers and security experts therefore leaving it open for human error and leading to an incomplete attack library. To further justify that the SAFELI framework is an efficient SQL detection tool, Fu, et al., (2007) could have arranged a controlled experiment, in a laboratory setup which will resemble a real world web application.

2.2 Combined Static and Runtime Monitoring

The combined static and dynamic technique employs the benefits of both to detect SQLIA's. AMNESIA is a proposed mechanism by Halfond and Orso (2005) that uses a model based approach, conjoining the static analysis technique with runtime monitoring. The proposed approach consists of four steps, the identification of hotspots where that SQL queries interact within the underlying database. For every one of these hotspots an algorithm is built that represents potential SQL queries that may be generated. Then calls are added to a monitoring function to check the queries at runtime. This monitor is enclosed by two parameters, the string that includes the submitted query and an ID which classifies the hotspot. Finally as the application is running the dynamically generated SQL queries are checked against the SQL query model which will prevent and detail all the queries that infringe the database. The experiments that Halfond and Orso (2005) implemented consisted of an empirical study. These studies were done using five open source web application projects on GotoCode.com and two other solutions developed by different teams of students where they implemented a prototype of the AMNESIA tool. The results they obtained ranged from 140 to 280 detected elements for each web application that are fairly clear-cut with no false negatives, as shown below.

Subject	*Unsuccessful*	*Successful*	*Detected*
Checkers	1195	248	248 (100%)
Office Talk	598	160	160 (100%)
Employee Directory	413	280	280 (100%)
Bookstore	1028	182	182 (100%)
Events	875	260	260 (100%)
Classifieds	823	200	200 (100%)
Portal	880	140	140 (100%)

Fig 1: Experimental outcomes. (Halfond & Orso, 2005)

Results from these tests also showed that the AMNESIA mechanism managed to detect 100% of a total of 1,470 attacks that they performed on the seven different applications. These subjects were all customary websites that accept input by users via a form, which then sends queries to an underlying database.

The method that Halfond and Orso (2005) proposed has some limitations in regard to the

fact that assumptions are made. This techniques success is dependent on the precision of the string analysis component for building the query models. Therefore certain types of coding could make the technique less accurate and result in both false positives and negatives. The method is also restricted to string and number values with a possibility of being compromised when collating SQL fragments from external sources. Experiments were very detailed and well thought out with a number of web applications tested. The attacks that were created to test the applications were done without bias as they employed a master's student who was an experienced programmer who was not familiar with the techniques execution. Halfond and Orso (2005) claimed the experiments could be easily re-run. This notion is well supported as Sun and Beznosovy (2009) used the testbed suite from project AMNESIA as a base of analyses and test a tool called SQLPrevent as well as partially incorporated by Shin, et al., (2006) in their SQLUnitGen approach discussed in section 2.4 and other methods . This shows that the experiments developed for the AMNESIA tool, offer a stimulating base for testing most SQL injection detection methods.

2.3 Web Application and Error Scanner

A web application and error scanner mechanism was proposed by Huang, et al., (2005) that uses a black box testing technique to detect attacks against web servers and automated web applications. It includes a fault injection crawler, where specially crafted malicious data input is used to identify vulnerable points that can be used to inject the SQL injection attack, with emphasis on observing the behaviour of the software under attack. "We identify vulnerabilities in Web applications by observing the output resulting from the specially prepared SQL injection patterns" (Huang, et al., 2005). The researchers employed behaviour monitoring for the detection of malicious content while also performing behaviour stimulation to entice any malicious actions within the examined components. The experiments that researchers formed included real world situations, 14 different popular sites were selected that reflected a variety in nature, design and overall size. Firstly they tested for thoroughness of the WAVES tool by contrasting the amount of pages recovered with three other crawlers. The results show that WAVES out performs Teleport which is one of the most thorough of these crawlers according to the researchers. "On average, WAVES retrieved 28% more pages than Teleport when tested with the 14 sites" (Huang et al., 2005). Then the injection algorithm was tested by configuring WAVES to identify all forms of interest e.g. those containing text fields to perform a negative response extraction (NRE), fill in and send the form, then make an incisive decision on the submission success based on the retrieved NRE and the reply page.

The researchers grouped a list of classified impacts of vulnerabilities into four wide-ranging groups; these groups were restricted access to resources, arbitrary execution of commands disclosure of private information and the denial of services. A total of 26 exploits were found that displayed impacts associated with the first three of these categories. They used these categories to implement a site to examine their behaviour monitoring mechanism. To test this mechanism WAVES was installed onto an unpatched version of a Windows 2000 operating system. WAVES had 100% accuracy rate in the first two categories, the table below shows the detection ratios for the four impact categories.

Class of Impact	Exploits Detection	Ratio
1) Restricted resource access	9	9/9
2) Arbitrary command execution	9	9/9
3) Private information disclosure	6	0/6
4) Denial of service	2	0/2

Fig 2: Class of impact detection ratios (Huang, et al., 2005)

The results collated conclude that the method is not entirely efficient as it could not detect all the vulnerabilities present. Within the last two categories at toal of 0 out of 8 vulnerabilities were detected. The use of the black box tool provides a faster assesment for indentification of vulnerable sites but this process does not provide critical analysis of any code. Therefore only giving an overview without thorough analysis for effective generated test cases. The experiments that Huang, et al., (2005) carried out were within a real world setting adding a

level of sophistication to their testing. However this was just to test the web crawler, self learning and NRE mechanisms, the behaviour monitoring mechanism was tested on a site that they had created , which meant different parts of the experiment would incur ambiguous results. If all mechanisms were tested on the real world applications a more accurate qauntitative analysis of the whole method could have been obtained. A combined experiment of all mechanisms together would have been more appropriate. The approach was also susceptable to false negative due to some conditions in regard to the forms being vulnerable or validates or major problems with the implementation of WAVES, either not correctly completing the forms or the NRE process failing. These considerations need to be looked at in further research in order to completely remove these threats to the WAVES model.

2.4 Automated Analysis

Shin, et al., (2006) proposed an automatic testing tool for the detection of SQL injection input manipulation vulnerabilities. In many ways this approach is similar to Halfond and Orso (2005) AMNESIA approach it uses static analysis but also with runtime detection via the use of the automatic testing feature. The approach tracks the values that are input by users in the AMNESIA SQL query model. The approach is split into three phases, within the primary phase test cases of which hotspot-reaching test cases are generated using a Jcrasher tool created by Csallner and Smaragdakis (2004). These test cases are then refined to include input values which contain SQL injection attack input. The final phase executes these test cases to detect vulnerabilities and a summary of test results are created.

The researchers used SQLUnitGen on two diminutive web applications as a means of evaluation. For both applications login modules were used to initially test and subjects were modified. One of these was developed in 2004, called cabinet. “Cabinet has approximately 2000 lines of code in 14 classes. Cabinet allows users to register, login, and order cabinets” (Shin, et al., 2006). The other web application, Bookstore was used as part of the evaluation for ANMESIA model taken from an open source website. To evaluate the generated test cases the researchers modified these applications in order to have a multitude of levels of input filtering, so that a set of controlled fault injections could be performed. Each subject was divided into three versions, the first has no input filtering function, the second incorporates input filtering for part of the input from users and the third has comprehensive input filtering for every user input.

Applications	Tools	Vulnerable hotspots	Vulnerabilities found	False positives	False negatives
Bookstore 1	SQLUnitGen	1	1	0 (0%)	0
	FindBugs	1	1	0 (0%)	0
Bookstore 2	SQLUnitGen	1	1	0 (0%)	0
	FindBugs	1	1	0 (0%)	0
Bookstore 3	SQLUnitGen	0	0	0 (0%)	0
	FindBugs	0	1	1(100%)	0
Cabinet 1	SQLUnitGen	5	3	0 (0%)	2
	FindBugs	5	5	0 (0%)	0
Cabinet 2	SQLUnitGen	1	1	0 (0%)	0
	FindBugs	1	5	4(80%)	0
Cabinet 3	SQLUnitGen	0	0	0 (0%)	0
	FindBugs	0	5	5(100%)	0

Fig 3: Comparison of results with a static analysis tool (Shin, et al., 2006)

The SQLUnitGen is shown to be quite proficient due to the fact that false positives were not encountered within the experiments conducted. Although the experiments show promising results in terms of the number of false positives, the tool does fall short in respect to the amount of false negatives encountered. The experiments were based around to smaller web applications which meant that the protocols were not detailed enough to determine an accurate depiction of the SQLUnitGen model. The model proposed could have been applied to two or three other web solutions, either created by the researchers or taken from the open source website as Halfond & Orso, (2005) used. The researchers also state that the technique will post less false positives than all static analysis methods. “The result of preliminary study shows that our technique is promising in detecting vulnerabilities with less false positives than static analysis tools” (Shin, et al., 2006). However they do not demonstrate that this statement is factually correct as they only compared the results with one static analysis tool called FindBugs. Therefore they should not categorically state that it outperforms most or all static analysis tools. Shin, et al., (2006) need to consider comparing their results with a selection of static analysis techniques before making this declaration.

3 Comparison

It is very challenging to give a definitive answer to which of the discussed approaches would render the most efficient overall. Each one, having their proven benefits for the specific type of environments (JAVA, PHP or ASP.NET) they were developed for. Therefore researchers should propose innovative techniques that can be easily adopted in the development process of applications to meet the deployment constraints. This is because the specifics of each web applications coding structure and syntax can affect the performance of much source code analysis in the SQL injection attack detection tools.

SAFELI was designed to only discover vulnerabilities on .Net Microsoft based products therefore will not detect any vulnerability on converse frameworks such as Java or PHP, whereas AMNESIA and SQLUnitGen are tools that analyse the source code of Java Web Applications. Huang et al., (2005) implement a type state-based algorithm (WAVES) for identifying vulnerabilities within PHP application code.

It is noted that many of these techniques are not in operation within real world environments due to their prerequisites. "The requirements for analysis and/or instrumentations of the application source code or acquisition of training data, limit the adoption of these techniques in some real world settings" (Sun and Beznosovy, 2010).

Results obtained by each of the researchers were varied with some having more success than others. This was partly down to the different techniques employed by each.

- SAFELI- String analysis technique
- AMNESIA – String analysis and runtime
- WAVES – Black box testing technique
- SQLUnitGen – Automated

4 Conclusions

The research papers discussed with this literature review have explored the methods proposed to discover SQL injection attacks on various web applications. This research review has revealed that the approaches and techniques are similar to one another, and often use the same strategies to ascertain a perfect model for SQL injection detection. Conclusions drawn from reviewing these research papers have encapsulated that the most effective method could be the AMNESIA approach based on results provided and large experimental analysis conducted. The results showed that the method could detect 100% of the vulnerabilities with no false negatives captured using a large quantity of web applications to test the recommended tool. The other research by Huang et al., (2005) and Shin, et al., (2006) found that their methods incurred more problems with false negatives or couldn't detect all of the vulnerabilities within the web applications they used. (Fu, et al., 2007) provided no evidence of any experimental analysis of their SAFELI methodology and therefore there is no definitive answer to whether or not the technique was efficient in detecting the SQL vulnerabilities.

Each method on their own have their benefits and limitations, however by combining techniques to detect vulnerabilities in applications code and using runtime monitoring would vastly increase the integrity of data and reduce the risk of database being compromised.

5 References

Bisht, P., Madhusudan, P. and Venkatakrishnan, V.N. (2010) 'CANDID: Dynamic Candidate Evaluations for Automatic Prevention of SQL Injection Attacks', *ACM Transactions in Information and System Security* , vol. 13, no. 12, pp. 1-39.

Cruz, V. (2012) *Black Hat is Over, But SQL Injection Attacks Persist*, [Online], Available: http://www.wired.com/insights/2012/08/black-hat-sql-injection/.

Csallner, C. and Smaragdakis, Y. (2004) 'JCrasher: an automatic robustness tester for Java', *Software Practice and Experience*, vol. 34, no. 11, pp. 1025-1050.

Fu, X., Lu, X., Peltsverger, B., Chen, S., Qian, K. and Tao, L. (2007) 'A Static Analysis Framework for Detecting SQL Injection Vulnerabilities', *Computer Software and*

Applications Conference, 2007. COMPSAC 2007. 31st Annual International, vol. 1, pp. 87-96.

Halfond, W.G.J. and Orso, A. (2005) 'AMNESIA: Analysis and Monitoring for Neutralizing SQLInjection', *Proceedings of the 20th IEEE/ACM international Conference on Automated software engineering* , pp. 174-183.

Huang, Y.-W., Tsai, C.-H., Lin, T.-P., Huang, S.-K., Lee, D.T. and Kuo, S.-Y. (2005) 'A testing framework for Web application security assesment', *Computer Networks* , vol. 48, no. 5, pp. 739-761.

Patel, N., Mohammed, F. and Soni, S. (2011) 'SQL Injection Attacks: Techniques and Protection', *International Journal on Computer Science and Engineering*, vol. 3, no. 1.

Shin, Y., Williams, L. and Xie, T. (2006) 'SQLUnitGen: Test Case Generation Detection', *Proceedings of the 17th IEEE International Conference on Software Reliability Engineering*, vol. 17.

Sun, S.-T. and Beznosovy, K. (2010) 'Retrofitting Existing Web Applications with Effective Dynamic Protection Against SQL Injection Attacks', *International Journal of Secure Software Engineering (IJSSE)*, vol. 1, no. 1.

An Analysis of Current Research into Music Piracy Prevention

Steven Hodgson

Abstract

This paper briefly looks at current problems when it comes to preventing piracy and identifying the main culprits. Current research is then reviewed and compared to give a better insight into the various causes of piracy, which can then be used in the real world to aid piracy prevention. Real world application of this theory is mentioned and the paper concludes with various methods that could be implemented to reduce piracy rates in younger people.

1 Introduction

There have been several attempts in recent years to solve the problem of music piracy. This paper will look research conducted on various groups of people and what may drive them to pirating music. Once it is discovered the type of people and why those people pirate digital music, research into more effective ways to prevent piracy can be done.

Current piracy prevention methods, such as strict Digital Rights Management, can actually have the opposite effect. Heavy-handed methods can inconvenience customers so much that they resort to pirating music in order to easily gain access to the content they want. This means it is very difficult to create a method that allows for the consumer to have an easy experience but also restricts the content to only the people who legally purchase it.

Therefore, this paper will look at various studies that have been conducted to find out which people illegally download music. Variables such as gender, self-control, willingness to pay, risk factors and others. The paper will discuss the methodology and findings of this research and critically evaluate the strength of the claims made by this research. These methods will also be analysed and compared to find any areas of overlap which could lead to new knowledge, with consideration of the application of this theory. Conclusions will be made throughout the paper and will all be drawn together at the end.

1.1 Current Research into Piracy Prevention

As this paper will be looking at various factors that can cause people to pirate music, literature on the research that has been carried out needed to be found. A brief description of the research conducted from this literature can be found below, before the in-depth look at each of these studies.

Chiang and Assane (2008) conducted research that looked at the role gender played in how a group of students perceived music piracy, and whether or not they were likely to pirate music. Malin and Fowers (2009) examined the self-control of high school students to learn their attitude towards music and movie piracy over the internet. Sinha et al. (2010) studied a group of college students and their willingness to pirate music based on the services offered by a fictional online music store. Sinha and Mandel (2008) examined the impact of three different

factors, risk of being caught, embarrassment if they were caught and public versus private consumption, on people with regards to downloading music illegally. These factors were given different scales and the participants were asked to rate them against how it would affect their willingness to pay for music. These four studies will be what this paper will look at.

During the search for relevant literature, some similar studies were discovered however they were not perfect for this paper. Teston (2008) and also Jamwal and Gupta (2012) investigated software piracy amongst IT and Technology students respectively. Piolatto and Schuett (2012) present arguments that point towards music piracy being beneficial for some artists, however harmful to others. Gunter (2008) looked at how students would react to certain piracy scenarios, with questions regarding morality, peer pressure and others.

2 The Current Problem

Music piracy is a problem many companies have been facing since the late 90s and has led to sales dropping by 53% in the past twelve years, with approximately 30 billion songs being illegally downloaded between 2004 and 2009 (RIAA, 2012). It is difficult for music companies to discover patterns in music piracy, and find the cause of problem. However, people have researched the area extensively to try and find certain correlations between piracy and various social, physical and biological factors.

3 Piracy in Different Genders

Chiang and Assane (2008) conducted a study that looked at how gender can affect a person's view towards piracy; how likely they were to pirate music, and how much of their music collection was made up of pirated music. Data was collected from students at three large universities, Florida Atlantic University, University of Nevada Las Vegas and New Mexico State University. These Universities were chosen for the diversity of students.

Of the 456 selected students, 50.7% were female and 49.3% were male. Surveys were given to the participating students and included questions about general views towards copyright law as well as more focussed questions regarding music file-sharing. There were also further questions relating to the perceived risk of being caught and willingness-to-pay for alternatives. The survey was anonymous and designed carefully to avoid use of the word 'piracy' which allowed for more truthful answers.

The results from this study show that, generally, male students use file-sharing to illegally download music more often than female students. Male students also have a larger percentage of their music collection acquired via file-sharing. 62% of the males used in the study admitted to using file-sharing as a means of getting digital music whereas only 54% of females used file-sharing. Furthermore, 44% of the males' music collection was made up of file-shared music and only 36% of the females' collection was file-shared music. The results also showed that increasing the threat of being caught from 5% to 10% decreased the extent of the file-sharing in males, by 6%, and females, by 7%. Increasing the value of a song, from $0.25 to $1.00 also reduced the extent of file-sharing by 4% in males and 8% in females.

The researchers claim that, on average, male students are more likely to use file-sharing as well as having a larger percentage of their music collection acquired via file-sharing. It also shows that male students are less likely to care about greater value. Female students, although they still use file-sharing as a way to get music, do so on a smaller scale, with less file-sharing and a lower percentage of their music collection coming from file-sharing. Female students respond more positively to greater value. Both male and female students did, however, react similarly to a higher risk percentage, with a similar reduction in file-sharing. Because of this, the researchers claim that the music industry should continue enforcing threats or actions towards people who use file sharing.

The methods the authors used in this study were good. The sample of male and female students was well balanced and the wording in the survey, to avoid negative language, was a good decision. The word 'piracy' may have deterred some students from answering honestly. The size of the sample of students is sufficient to provide enough evidence to understand the scale of the problem. The authors claim that an increased risk of being caught reduces the amount

of piracy in students; however the percentage of reduction in both male and female students is less than 10%. This shows that further investigation is required to find piracy prevention methods that further reduce the rate of piracy.

It can be concluded from the results of this study that increasing the chance of getting caught and prosecuted for illegally downloading music reduces the likelihood that someone will choose to pirate a song rather than pay for it.

4 Piracy and Adolescent Self-Control

Malin and Fowers (2009) conducted a study of 200 high school students from New York to primarily understand self-control levels when it comes to music and movie piracy, among other things. The study included 18 freshmen, 76 sophomores, 58 juniors and 48 seniors from various ethnical backgrounds. The participants were given a survey to complete.

The survey included questions regarding attitudes towards piracy, self-control, computer experience and peer-pressure. Questions regarding attitudes towards piracy were answered using a 1 to 4 scale, with 1 being 'strongly disagree' and 4 being 'strongly agree' with statements such as ''I think it is okay to use copied music and/or movies because the community at large is eventually benefited.''. The self-control section also used the 1 to 4 scale used in the attitudes towards piracy section with statements such as ''I often act on the spur of the moment without stopping to think.''. Computer experience included questions asking how often the students used computer software, such as word processing, spread sheets, web browsers and email. The students answered the questions using a scale of 1 to 3; with 1 being never, 2 being sometimes, and 3 being often. The final section was related to peer-pressure, which contained a single statement: "My friends are often in trouble" to which the response would be yes or no. This section also included demographic information to complete including race, gender and school grade.

The results of this study found that the school grade that the students were in correlated with their attitudes towards piracy, with 12^{th} grade students accepting piracy more than 9^{th} grade students. It also found that male students were more accepting of piracy than female students. Those students that indicated that their friends are often in trouble were also more accepting of piracy. The race of the students did not make a difference to their acceptance of piracy; however students that had higher internet usage were more accepting of piracy. It was found that the older students with lower self-control were the most likely to pirate.

The authors of this paper claim that self-control is the biggest factor when it comes to piracy in adolescents. They state that educating high school students on the effects piracy has on the artists and reiterating that piracy is not a victimless crime. Stating that piracy is theft may help the students realise that it is unethical, which might help to lower piracy rates. Finally it is claimed that piracy rates increase with age.

The sample of students used for this study is mostly varied with students from each year of high school but they all were selected from only one school, however including more schools for the sample to be selected from, as in Chaing and Assane's (2008) study, would increase the variety of students. The claim that freshman students are less likely to pirate may be down to the majority of older students in the study, with only 18 of the 200 participants being freshmen.

Although not its primary focus, this study also shows that males are more likely to pirate digital music than females. This correlates with Chaing and Assane's (2008) study. The two studies had a similar method of giving a survey to the participants and then analysing the results to back up their claims. This information could be used to target anti-piracy campaigns more towards the male population as it will correspond with the majority of internet pirates. Online stores could also aim their marketing more toward the male population in an attempt to reduce the piracy rates of males.

5 College Students' Willingness to Pay

Sinha et al. (2010) studied 849 college students on their willingness-to-pay for a music track based on the attributers of a fictional online music store. The method included four conditions of Digital Rights Management (DRM), such as no DRM restrictions or very strict DRM restric-

tions. The participants were then given a scenario relating to a new online music store to get their opinion on its attributes. The participants were asked to state how much they were willing to pay for their favourite songs on the new store. The fictional store was presented in a general manner, without any association with brand names. Before the participants were given the surveys a presentation was given and any questions were answered to allow the participants to be fully clear on what the store was offering.

The results of this study showed that individual DRM restrictions, such as reproducing the song downloaded on multiple players, did not affect the participants' willingness-to-pay. However it did show that shared DRM restrictions, such as being able to share the song with a friend, were received positively and increased the participants' willingness-to-pay for a legal music download, with 70% of participants responding positively to that attribute.

The authors of this paper claim that reducing or even fully removing Digital Rights Management from online music stores would be received positively. They state that DRM is not in fact a control mechanism but is, however, another factor consumers will consider when deciding their willingness-to-pay for music. Complete removal or very relaxed DRM restrictions are recommended by these authors to increase sales of digital music.

The methods used by the authors of this paper in their study to obtain the willingness-to-pay for digital music were good. The fictional music store will have allowed the participants to feel they were part of a focus group rather than a piece of research, which will have produced more honest answers. The results clearly showed that being able to share your purchased music with your friends, via an online webpage or by other means, would be a feature that dramatically increases a consumers' willingness-to-pay for music. The claims that Digital Rights Management restrictions should be severely relaxed or removed altogether are not justified by the research carried out. The research and the results they produced do not state how much, if at all, the students reacted positively to the removal of DRM on the fictional music store.

6 Risk and Embarrassment

Sinha and Mandel (2008) studied the effect of three factors: risk, embarrassment and public versus private consumption when it comes to college students' willingness-to-pay for music downloads. 386 students were selected to take part in the study and were each given a questionnaire to complete. The first page included a news story about a recent lawsuit that had been filed against over 500 students for illegally downloading music. The participants were then asked to imagine they were going to make a compilation CD and gave estimates on how willing to pay they were for six scenarios. The different scenarios had varying levels of risk and embarrassment of being caught. They also answered questions regarding musical tastes, internet habits and demographics.

The results of this study showed that the willingness-to-pay increased with the increase in risk of being caught. The average the students were willing to pay for a music track with low risk was $0.34, medium risk was $0.48 and certainty to be caught was $0.62. The demographic questions resulted in a median age of 21, with age ranging from 18 to 48 years old. 86% of the sample admitted to have used file sharing applications to obtain music in the past. 37% of the sample indicated that they did not like buying CDs as you have to pay for songs you do not want. 42% stated that downloading was easier than buying a CD. 10% said they used downloading as a way of sampling songs before buying them. The results also indicated that females are willing to pay more for music.

The authors of this paper claim that an increase in legal action towards music pirates would be an effective method of reducing piracy.

The method the authors used to conduct this study allowed for the participants to further express their views towards paying for music and the questionnaire that was handed out was sufficient enough to gain an understanding of their willingness-to-pay for music downloads. This study did only take university students into consideration, which could result in a lower overall willingness-to-pay than non-students and those with full time jobs because of the money they have available. The results showed that increasing the risk of being caught also increased the average each participant was willing to pay for

the music download, meaning their claim of increasing legal action against pirates does mean that people are more inclined to pay for music downloads.

This study, like Chiang and Assane's (2008) and Malin and Fower's (2009), found that males were less likely to pay for music. Both this study and the study into gender affecting piracy claim that more threats action should be taken against pirates as it is seen to be a very good deterrent and increases the likelihood that someone will pay for a music download.

Based on the results gained from this study, and the 10% of the sample that indicated they illegally downloaded music to sample new music, it would be wise for online music stores to offer a better method of allowing users to sample music before making the purchase.

7 Real World Implications

The results of these four pieces of research have real world implications and they can be used to improve digital music stores. Sinha et al.'s (2010) claim that completely removing or offering music downloads with relaxed DRM restrictions could be implemented by any digital music store however it would require approval of the music company. This approval could be difficult to obtain as most record companies do not like the idea of DRM free music, despite being favoured by consumers.

Chiang and Assane's (2008) and Sinha and Mandel's (2008) research suggested that more legal action needs to be taken towards music pirates as this will deter people from pirating. This would be difficult and expensive to achieve in the real world. Piracy detection systems would need to be implemented which would require further research to get working.

Malin and Fower's (2009) research suggested educating high school students about the harm music piracy causes and that it is, in fact, a crime. This would require people to visit lots of different high schools and give talks on piracy, which would be expensive and may not even be affective at lowering piracy.

8 Conclusions

The research that has been looked at in this paper has allowed me to conclude that increasing the chance of being caught and prosecuted when you illegally download music would act as good deterrent to pirates. Piracy could be lowered if people are educated about the effects it can have on individual artists and the music industry as a whole when they are younger. This is because younger people may not be aware that piracy is a crime. As males are the more likely to pirate music, anti-piracy campaigns could be aimed more toward males as it will have the biggest effect on piracy reduction. This information could also be used to target online music marketing more towards the male population as it might increase their willingness-to-pay. Finally, digital music stores should offer a better way of sampling music before the consumer has to pay as this is commonly used as an excuse for piracy.

References

Websites

RIAA, 2012, *Scope of the Problem*, [online] Available at: <http://www.riaa.com/physicalpiracy.php?content_selector=piracy-online-scope-of-the-problem> [Accessed 22 November 2012].

Academic Journals

Chiang E, Assane D, 2008, Music Piracy Among Students on the University Campus: Do Males and Females React Differently?, *The Journal of Socio-Economics*, vol. 37, pages 1371–1380.

Gunter W, 2008, Piracy on the High Speeds: A Test of Social Learning Theory on Digital Piracy among College Students, *International Journal of Criminal Justice Sciences*, vol. 3, pages 54-68.

Jamwal S, Gupta N, 2012, Software Piracy among IT students of J&K: Ethical or Unethical, *International Journal of Computer Applications*, vol. 1, pages 33-36.

Malin J, Fowers B, 2009, Adolescent Self-Control and Music and Movie Piracy, *Computers in Human Behavior*, vol. 25, pages 718–722.

Piolatto A, Schuett F, 2012, Music piracy: A case of ‘‘The Rich Get Richer and the Poor Get Poorer’’, *Information Economics and Policy*, vol. 24, pages 30–39.

Sinha R, Machado F, Sellman C, 2010, Don’t Think Twice, It’s All Right: Music Piracy and Pricing in a DRM Free Environment, *Journal of Marketing*, vol. 74, pages 40–54.

Sinha R, Mandel N, 2008, Preventing Digital Music Piracy: The Carrot or the Stick?, *Journal of Marketing*, vol. 72, pages 1–15.

Teston G, 2008, Software Piracy among Technology Education Students: Investigating Property Rights in a Culture of Innovation, *Journal of Technology Education*, vol. 20, pages 66-78.

Real Time On-line Analytical Processing: Applicability Of Parallel Processing Techniques

Kushatha Kelebeng

Abstract

With the recent development of data warehouses, on line analytical processing (OLAP) has become the core of the decision support system. OLAP techniques are used for analysis of data in decision support systems since complex queries that require different view of data are used. OLAP uses multidimensional data model to provide effective decision making. The expansion of the data warehouses has led to the use of parallel processing techniques in order to reduce the time taken to analyse data thus allowing for systems that analyse data in real time and are proficient on larger data warehouses. This paper provides an evaluation of the use of parallel computing to provide performance enhancement in OLAP systems and high performance provided by parallel processing techniques that will allow for real time analysis of data in larger databases.

1 Introduction

Over the years business intelligence has become one of the keystones of the information technology sector. Business intelligence incorporates a diversity of systems which comprise of data warehousing, data mining and analytical processing.

A data warehouse is a collection of data used for the decision support systems in an organisation. Research shows that a data warehouse is a collection of decision support technologies used to help an organisation in making better and faster decisions (Dayal et.al 2009).

The expansion of data warehouses has led to the emergence of some analytical tools such as Online Analytical Processing (OLAP) for efficient analysis of data in warehouses. OLAP technology is used to analyse data on large data warehouse environments. Xu et.al (2010) argues that OLAP enhances the power of data analysis in data warehouses since it allows users to efficiently access data and allows for meaningful data to be derived from large amounts of data.

The major limitation of the current static OLAP system is performance due to the expansion of data warehouses, as new data gets into the system it is not updated instantly and this lead to the data warehouse not always being up-to-date as updates are only done periodically.

A call has been made to provide a real time OLAP system that will provide a system that is updated instantly thus providing an up-to-date data warehouse. Buchmann (2005) argues that a real time system is a system designed to handle data that is constantly changing. Real time processing means that a transaction is processed faster and the results are available instantly thus improving the decision support systems.

Parallel processing techniques can be used to provide an OLAP system that is real time since they offer concurrent data processing tasks for the point of escalating the computational speed of the system and also enabling the use of large data set. The focus is on distributing data across different computing nodes to be processed in parallel. Real time data access using OLAP is fundamental since it yields reliable results.

This paper focuses on evaluating how the application of efficient parallel processing

techniques in OLAP query processing can provide the chance for real time OLAP on larger data warehouses, thus allowing for effective decisions that empower users to be made.

This paper is organised as follows; in section 2 the recent work in the area is evaluated mainly focusing on the techniques that are available, OLAP and parallel processing techniques are also critically evaluated in relation to the recent research. The synthesis of the lessons learnt is covered on section 3. Recommendation on the future work is done in section 4 and the conclusion then follows in section 5.

2 Algorithms Evaluation: Real Time OLAP Using Parallel processing Techniques

The traditional online analytical processing (OLAP) research pursues the static data cube which permits users to analyse data from diverse insights and at a diversity of summarization levels. However there are problems that are related with this kind of approach and one of them is the performance issue. These problems have led to what has come to be known as real time OLAP system that will be updated as the latest data comes into the system hence providing an up-to-date database.

Recent research has come up with a way of implementing real time OLAP using different approaches. According to Dehne and Zaboli (2012) the construction of a parallel real time OLAP data warehouse is intricate though it's deemed proficient for larger databases.

Fragmentation of data and its distribution across different computing nodes to be processed in parallel allows for real time loading of OLAP queries hence improving the performance. Usman et.al (2009) argues that parallel computing can be used as an approach of augmenting the performance of OLAP as it reduces the computational time by allocating tasks between a range of nodes.

The sections below will be looking at the different parallel processing techniques that can be applied for parallel OLAP query processing that can allow for real time OLAP.

2.1 OLAP and a cluster of workstations

According to Dehuri and Mall (2007) parallel processing has appeared as a resolution to the problems encountered by OLAP queries due to the expansion of data warehouses. The authors presented a parallel algorithm for the OLAP queries using a group of machines. The algorithm will allow for execution of OLAP operators in parallel. The reason why the authors came up with such a technique was that they wanted to shift from big computing amenities and focus on a cluster of workstations since they provide networks that are of high speed and processor performance that is advanced.

The authors presented an algorithm for parallel processing of OLAP queries using a cluster of computers. The cluster had two digital alpha workstations and the parallel virtual machine was used for message passing. The authors evaluated performance from two different perspectives; the size of the database and the number of attributes supplied to the operator. Two database sizes were used; the 0.2 and the 4.7mb and the number of attributes ranged from two to six. The results obtained that the speedup is higher for larger databases. It showed that when the database size increases to either gigabytes or terabytes a higher speedup can be expected.

The results showed that with a larger database and more number of nodes, higher speed can be achieved. The authors concluded that for larger databases and large cluster sizes the speedup could be achieved by parallelization of the operators.

Though the algorithm can be a fine answer for the decrease of time taken to respond to queries, the authors failed to look at the aspect of failure. Their algorithm is not fault tolerant as data is not replicated at different nodes to achieve a fault tolerant algorithm. In the event of failure of individual workstations, the state of the computations will not be saved.

2.2 OLAP and multi-core processors

Just recently a new contribution in real time OLAP by the use of parallel computing techniques was proposed. Dehne and Zaboli

(2012) suggested a real time OLAP that is parallel using processors that are multicore. The major problems identified by the authors were the major performance issues associated with the large scale databases. This research was aimed at addressing these issues through the use competent parallel processing techniques. Authors designed a DC tree that is parallel for architectures that are strictly multicore, with the intention of showing that the performance of the tree improves as the processor cores are increased.

The method involved two parts; an extension of the DC tree data structure and the new algorithms being the parallel_OLAP_insert and the parallel_OLAP_query. The authors' solution consisted of three parts being the minimal locking scheme that locks only the node that is currently being updated, a time stamp that is a mechanism that allows concurrent transactions to discover when they are working on invalid data. A sandy bridge multicore processor was used and the system was implemented in C++ and executed in Linux kernel. 400 000 parallel-OLAP_insert operations were executed on 1, 2, 4,8,16 cores. The speedup was 40% of the maximum possible speedup.

The authors then observed that the speedup increases with the increasing number of cores and then concluded that it is better for large data warehouses.

The major problem of the DC tree is possible interference between the parallel and the insert operations. The strength of this work is that it permits for multiple insert and query operations to be executed both in parallel and in real time. The algorithm was evaluated using different set-ups and it demonstrated that the performance can be improved by increasing the number of processor cores hence achieving a significant speed up in transaction response time.

Authors proposed different techniques that can be used for OLAP query enhancement. Dehuri and Mall (2007) reported on the use of a cluster of workstations for parallel processing of OLAP queries. The authors identified that a lot of processing overheads occur during the analysis of data using OLAP due to the large sizes of the data warehouses and this then leads to unacceptable response times. The authors' choice of the processing technique was good since clusters provide flexibility because their node capabilities can be enhanced and also extended. For the analysis the authors used the partitioning schemes, they partitioned the available processors into a number of clusters and the data was also partitioned and allocated to the partitioned group of processors.

However the authors failed to take into consideration the possibility of having more processor groups and having to deal with extra processors. The data set that was used by the authors for experimental purposes was also insufficient and that made it difficult to come to conclusions using the data set.

In addition to the OLAP enhancement work using parallel computing, Dehne and Zaboli (2012) also proposed the use of multi core processors. Their work proved to be good as the authors tested their method using different settings such as the use of diverse proportions of the query and insert transactions, contrasting system loads in order to demonstrate that their method improves the response time. The analysis of the results presented by the authors show sufficient evidence to justify the claims made.

The authors clearly stated the hardware platform that they were using for the experiments and how they are going to implement the system. Though Dehne and Zaboli (2012) presented a number of experiments, they failed to provide comparison data as they claimed that it is the first method to be parallelized using multicore processors.

Looking at research discussed above the work by Dehuri and Mall (2007) seemed substandard as compared to the one by Dehne and Zaboli (2012) as the authors failed to mention some of the aspects like how the system will be implemented and the hardware platform to be used for the experiments was not clearly stated and also the authors only evaluated the performance from only two perspectives.

2.3 OLAP and the cgmOLAP server

In addition to the enhancement work using parallel processing techniques Chen et.al (2005)

presented a technique known as the cgmOLAP server which is an competent parallel generation and querying of large ROLAP data cubes that can be measured up to terabytes. This technique is a parallel system that is fully functional and it is also able to make data cubes at a speed of more than a terabyte in an hour.

The server consist of an application interface, a parallel query engine that will support parallel rollup, drill down, slice, dice and pivot queries, a parallel cube materialization engine that will support the generation of partial iceberg and full cubes, metadata and cost model repositories. The system will run on two 32 processor with Linux based clusters. The server integrates a number of parallel computing techniques. The technique will increase the computational power through the use of multiple processors and multiple parallel disks.

The strength of this piece of work will be that it is fully parallel and has been designed to capably expand the computational power of memory clusters that are cheap and that it is highly scalable looking at its dimensions and processors. The authors also took into consideration the issue of fault tolerance and in case the two clusters are not available the system will run locally on a single processor machine.

2.4 OLAP and the SIDERA server

Due to the performance limitations presented within large databases, in 2007 In 2007 Eavis et.al (2007) presented a technique almost similar to the one by Chen et.al (2005) known as the Sidera, which is a server that is cluster based for online analytic processing. The model consists of a network accessible frontend server and a series of backend servers that handle user requests. This technique proposes platform with many nodes and also consisting of a chain of largely autonomous close servers combined as one, presenting what is known as the as an OLAP server that is fully parallelized.

The strength of this work is that it provides an expressive database management systems support back end cluster nodes and requiring less implementation effort.

The authors concluded that though this technique can be viewed as one of the best techniques its major limitation is that conventional relational database management systems have only restricted maintenance for highly developed OLAP functionalities such as hierarchical querying.

2.5 OLAP and data fragmentation

Pereira et.al (2012) also added a brick to the castle of knowledge by proposing a real time data loading and OLAP queries using data fragmentation across many computing nodes. The technique applies horizontal fragmentation to the fact table to resolve the issue of data loading. It updates the data warehouse from operational data sources hence allowing queries to be executed in real time. The architecture decreased the response time of loading operations and it also executes queries faster.

The architecture relies heavily on the use of distributed and parallel techniques over a cluster of computers. The potency of this work is that it is optimised to achieve real time insertions of data.

Eavis et.al (2007) and Chen et.al (2005) both presented a similar approach of using fully parallelized servers to deal with the issue of data parallelisation in order to come up with a real time OLAP system. What motivated both authors was the growth of data warehouses and the processing power and thus leading to single CPU servers to be strained. The choice of hardware that was used to implement the system is of high-quality and proved to be excellent as they are easy to administer though they can be quite expensive to purchase and to maintain. Looking at the factor of scalability the CgmServer by Chen et.al (2005) proved to be better than the Sidera server as the authors stated that it is fully scalable in terms of both the processors and dimensions.

It is eminent that a significant amount of work has been done with regard to applying parallel processing methods in OLAP query processing in order to have real time OLAP which can lead to effective data analysis in data warehouses

3 Synthesis Of Lessons Learnt

The area of business intelligence encompasses a variety of important aspects and one of them being the analytical mining. Many tools have emerged for analysis of data to enable options for decision making purposes. OLAP is one of the techniques used in decision support systems by finding interesting information from large data warehouses and analysing it to be used by an organization. Many techniques have been created in order to improve OLAP system for it to yield reliable results.

4 Recommendations

Parallelism of data plays a momentous part in analysis of data as it speeds up the rate of data analysis in warehouses; this is achieved by parallelism of OLAP operators. Since increasing the number of processors improve performance and the running of a data warehouse in a fully parallelized server can be an excellent resolution in the reduction of query processing time, in future the two should be hybridized that is, using both the parallelised server and multicore processors in order to permit OLAP systems that operate in real time and proficient for larger databases. Another possible extension that can be considered is replicating data at different computing nodes to deal with the issue of fault tolerance.

5 Conclusions

The fast growth of data warehouses and OLAP technology has paved the way towards effectual analytical mining. Various algorithms were discussed in the body of this paper and it showed that the improvement of performance hardware is no longer based on how faster the processors are but on increasing the number of the processor cores. It is expected that the application of parallel processing techniques will augment the power of data analysis in data warehouses and also lead to real time OLAP which will reduce the computation time of analysis. This will help in achieving better performance improvement to the current OLAP system.

References

Buchman A. (2005) 'Real Time database systems.' *Encyclopaedia of Database Technologies and Applications*.

Chen Y., Dehne F., Eavis T. and Rau-Chaplin A. (2005) 'cmgOLAP: Efficient Parallel Generation and Querying of Terabyte size ROLAP Data Cubes.' *Proceedings of the 31st VLDB conference*, Trondheim, Norway.

Dayal U., Castellanos M., Simitsis A. and Wilkinson K. (2009) 'Data Integration Flows For Business Intelligence.' *International Conference on extending Database Technology: Advances in Database Technology*, Saint Petersburg, Russia Federation, pages 1--11.

Dehne F. and Zaboli H. (2012) 'Parallel real time OLAP on multicore processors.' *12th IEEE/ACM International Symposium on Cluster, Cloud and Grid computing.*

Dehuri S. and Mall R. (2007) 'Parallel Processing Of OLAP Queries using a Cluster of Workstations.' *International Journal of Information technology and Decision Making*, 06(02): 279--299.

Eavis T., Dimitrov G., Dimitrov I., Cueva D., Lopez A. and Taleb A. (2007) 'Sidera: A cluster-Based Server for Online Analytical Processing.' *part 2, LNCS 4804*, pages 1453--1427.

Eavis T. and Taleb A. (2012) 'Query optimization and execution in parallel analytics DBMS.' *IEEE 26th International Parallel and Distributed Processing Symposium*.

Muhammad U., Asghar S. and Fong S. (2009) 'A conceptual model for combining enhanced OLAP and data mining systems.' *Fifth International Joint Conference On INCIMS and IDC.*

Xu B., Bing W., Haode L. and Jian L. (2010) 'The Design And Implementation Of Web Based OLAP Drilling Analysis System.' *Fuzzy Systems and Knowledge Discovery (FSKD)*

Seventh International Conference, pages 2570--2573.

Evaluating Authentication And Authorisation Method Implementations To Create A More Secure System Within Cloud Computing Technologies

Josh Mallery

Abstract

Cloud computing is a relatively new way of storing data and resources. It has been the subject of massive growth worldwide. Although security systems on cloud networks can be very secure and effective, there still are some security issues. Security issues like authentication and authorisation can still be a problem. Within this survey paper contains description, analysis and evaluation of five authorisation and authentication security methods which have been suggested by current researchers. These security methods are then analysed and evaluated. Also included are suggestions of further reading. This paper reaches an overall conclusion on which method would be suggested to implement to secure their cloud network.

Key words: cloud computing, authorisation, authentication, security methods

1 Introduction

Cloud computing is a way of communication and resource sharing which has had a meteoric rise in popularity during the past few years. Despite the massive growth, there is one major disadvantage to this data sharing method that could, potentially, be very harmful to the users transferring data; data security within cloud computing. Data security is integral to cloud computing; without it there would be sensitive data manipulated from the cloud network and could potentially be leaked which would be damaging for all parties involved, especially the company and their clients.

Many people could be trying to obtain the private resources held on the cloud network (whether it is a private cloud or a hybrid cloud), (Sood, 2012), without the owner of the resource's consent. From engaging with clients from companies who actively cloud compute, this is an issue on a worldwide scale, so this could consequently affect millions of people. This survey paper focuses on the identity management factors within cloud computing; the authentication and authorisation of the cloud interactions.

This paper will evaluate authentication and authorisation security methods that have been designed by other researchers to help achieve the aim of developing more secure cloud computing networks.

1.1 Research Evaluated

There have been ten journals gathered altogether for this research. The conclusion reached is that only five will be used within this research as the other four are not focused to the topic authentication and authorisation within cloud systems, these papers just discuss cloud computing security within a generalized sense. The authors of these papers are Kadam et. al. (2012), Rong et. al. (2012), Shaikh and Sasikumar (2012) and Singh and Jangwal (2012).

Research that Thiyagarajan and Dinesh (2012) conducted was research into product authentication via cloud computing. Although this can prove to be a good theory for retail companies selling their products, this will not be relevant within this paper as the focus here is specific to authentication and authorisation methods in the cloud specifically to make it more secure, and not about product authentication.

2 Authorisation and Authentication implementations

Below will be an investigation into the chosen research journals, which will contain discussion of the research methodology along with analysis and evaluation of the experiments, experimental results and research conclusions.

2.1 Mutual Authentication Framework

Nakak et. al. (2012) suggested a mutual authentication framework for users logging into the cloud network. Compared to the existing frameworks used in the industry currently, this suggestion has a two way authentication process which will include an update password procedure. The first section of this process is the registration phase is where the new user enters a valid ID into the system. The next stage is the user login and authentication stage; this model shows another method called the password change phase which the user can activate at any time to change their current password. The user will have to use their email to access the option of a password change. It says within this research that the evidence found proves that it is secure and faster; (Nakak et. al. 2012).

The registration phase within this system assumes that the user is providing the correct email ID information. The change password stage highlights more security issues; it could be better to implement that option within the user's system offline, so will have no threat of any unauthorized access via the internet at the emails. Although it will be more secure via the two way authentication, as both the server and the user will have to prove their identity, it will be slower compared to today's framework implementations because there are more processes taking place, which will prove to be more time consuming. The system will only be faster if they take advantage of the use of more servers.

The registration phase can prove to be bad practice because if this was an implementation live in the industry, this would be the first level of security that has been penetrated by a potential hacker, with ease.

This model does seem to be very secure in the way it authenticates the system and user; however, we won't be able to say for definite that this is the case and we can't say the conclusions Nakak et. al. (2012) draw from this are fully correct until we experiment and test this framework. Because they haven't conducted any experiments yet, they could firstly develop a prototype framework and test this via a lab experiment. They would be able to control most aspects of the experiment and remove extraneous variables, this would be a start as they could iron out any initial faults as they would know the reasons behind the faults, and rectify those. This would be more controllable than a natural experiment and is more realistic.

They could improve this further by placing the prototype within a fully operating company and conduct action research so they could experiment further; by doing this, the results and conclusions they have drawn from the research paper will be validated further.

2.2 New User Authentication and File Encryption

Nafi (2012) et. al. also proposed a new framework in cloud computing to ensure the user's password into the system is secure. They propose that the system will implement a one-time password for the user to log into the system with. The system will generate a new password randomly which will only be used once by the user, then will be deleted straight from the system. The password will be given via an authorized email. The password will be generated by a MD5 hashing method (Jayarana et. al., 2012). "The main purpose of MD5 hashing is that this method is a one way system and unbreakable" (Nafi et. al., 2012). One of their advantages for this proposed system they said would be that the probability of information leakage would be low compared to all the current models in use which have a probability of medium (Nafi et. al. 2012).

There is no experimental evidence to support these claims. The case studies and the lab experiment that they conducted were for the downloading times of the system.

Within this research Nafi et. al. (2012) could have started off by conducting a lab experiment. This is a beneficial approach because it would be very controlled and the results from this would be reliable. The next experiment they conducted were case studies in which they worked with different users at different times. Although this method ensures that there is a wider sample base of users, this lacks consistency. It would have enhanced this research project by conducting action research as it would have been more relevant to what they are actually testing; they would be able to gain results from the actual security framework working on the company's machines, to evaluate if it is more secure than the current models.

The authorized email can prove beneficial because this is the first wall of security; because if the user does not have an authorized email account, then the system will remove attempted unauthorized access straight away. The MD5 method does have certain vulnerabilities which they haven't stated and this could be seen as biased and could have an impact on the reader's opinion when studying their research.

2.3 Authorisation System Structures regarding Privacy within the Cloud

Chadwick and Fatema, (2011) conducted some research into authorisation structures and designed their own structure. They tested the correct authorisation decisions are being made and then they implemented a performance check to see how quickly they are being made. To test the correct authorisation had been chosen by the system the following methods had been developed; test one requested to store personal data. The system evaluated the three policies (user's, legal and controller's) and was stored by the authorisation system with a sticky policy. In test two all the policies were evaluated again and the request was to read the personal information. It was granted. The final test requested to transfer the data to different cloud network via evaluating the policies. A sticky policy returned with the granted request (Chadwick and Fatema, 2011).

"Each test was run sequentially 500 times and then the mean time and standard deviation were calculated" (Chadwick and Fatema, 2011). The authors went on to say that any anomalies / outliers within the data were removed from the results. "Any results that varied over 3 times the standard deviation from the mean time were removed as outliers." (Chadwick and Fatema, 2011) The experiment recorded the speed of the authorisation via the sticky policies.

Although this is very much related to authorisation within cloud computing, this experiment they proposed does not fully test the security of the cloud network, only the speed of the authorisation is measured. To test for security an experiment could be set up where researchers time how long it would take to penetrate their security implementation; this experiment will analyse security as well because the quicker it takes to penetrate the system from the controlled hacker, then the less secure the system will be. Each phase of this experiment was tested 500 times which provides a large sample data which makes the results gained from this valid. However, the outliers were removed, which added up to 3% of the overall results (Chadwick and Fatema, 2011). It would make the experiment more reliable if the authors kept these in the final results as it would prove there was no external influence on the recording of results. It would further improve the reliability of the experiment if the median, mode and range of results were recorded; because the conclusion would have more results to prove their theories.

2.4 Integrity and Authentication Framework

Sood's (2012) proposed model deals with two phases of cloud computing; storing the data within the cloud and retrieving the data from the cloud which includes "...double authentication, verification of digital signature and integrity..." (Sood, 2012).

The authentication within this proposed model is very simple, yet effective; a password and a digital signature are sent to the cloud and a search request is sent and the result is returned.

The user will then attempt to download the request and the reaction of this will be the encrypted file is returned. The user will then use the decrypting key (provided by the owners of the data) to decrypt and consequently view the file. The author then analyses the security of this model; the data and keywords are encrypted which are stored on the cloud (Sood, 2012). This model has changed from a 128-bit encryption to a 256-bit encryption to add more security from a brute force attack. The MAC encrypts data sent from the owner which will reduce the chance of and will easily show evidence of data tampering. Overall this proposed method is the resultant of several other proposed methods from different researchers. Sood (2012) has included the successful parts of other implications of the others, and developed those on his methodology. Previous methods use either index and encryption, classification of data methods or MAC, whereas Sood's (2012) proposed design has implemented all of these aspects.

The results show of the experiment that was conducted on a cloud computing simulator called Hadoop, that this method of incorporation holds the best security value. The results prove that this is a more secure implementation compared to the current research already out there, which was what Sood (2012) was trying to develop. This was mainly down to the fact that the proposed model makes very good use of current research by implementing and improving on that in its own respect. Within the research the experiment wasn't fully explained and as a result some may question the validity of the results, as they would not know entirely how the researcher gained these. Overall however, this is the strongest proposed model so far with regards to authorisation and authentication within cloud computing. This is because the author has the results to prove the theory is correct and has a strong implementation of already tried and tested methods within this research.

This paper also makes suggestions to further the security of the cloud with regards to other possible implementations for example "Off-site back-up for disaster recovery" (Sood, 2012).

2.5 Cloud System Trust Model

Zissus and Lekkas (2010) proposed a cloud user system that meets specific requirements in terms of security issues; strong and secure authentication and authorisation methods for the users to login to the cloud network. This system is a trust based system and operates in a top-down fashion, the dependent layer has to trust the determinant layer, if not, the system would break down (Zissus and Lekkas 2010). Zizzus and Lekkas (2010) identified other requirements to meet other security issues; however, this paper will be focused on authentication and authorisation. This research proposes that the whole system is focused around certificate-based authorisation.

When the user logs into the cloud network they have to get their ID authenticated to gain authorized access. The application that they access within the cloud also has an authorisation certificate. The data will then be encrypted (Bhati et. al., 2012) and will travel through the cloud via the virtual datacenters, which will have a virtual server certificate. The final stage will have a hardware authorisation when it is located in either datacenters; Europe or Asia (Zissus and Lekkas 2010).

The system seems very secure, which is a consequence of the amount of authenticated layers the system holds. This methodology of the system could be seen as a domino effect. As the system is very much based on layer dependents, if one layer's certificate of authentication does is not validated, then the result would be the user won't be able to achieve access. Zizzus and Lekkas (2010) have identified a couple of limitations within the system they have proposed. This method implements a lot of encryption and decryption techniques so consequently the outcome would result in the system running slower and "…inducing additional processing consumption" (Zissus and Lekkas, 2010).

This could be a very effective method if it is put into practice and implemented onto a functional, working system. However, there haven't been any experiments stated within their research article to back up their claims.
They could conduct a lab experiment to test each individual authentication layer to make sure it is as secure as possible. Dividing each layer would make it easier to pinpoint any problems compared to testing the whole system.

The next stage would be to design and develop a prototype system with all the layers included which had just been tested, and test this within a functional company (action research).

3 Application of Method

Many companies use clouds as part of their business i.e. storing information and interacting with colleagues. It is imperative for the company that they implement a security feature that will keep all their data within the cloud network secure. However, depending on their wealth and resources, each company may choose a different approach to security. A large company may choose to implement the best possible feature as money has no relevance, whereas a smaller company may choose for a potentially less secure network, which is more financially viable.

As these models haven't been fully implemented, it would be difficult to say how much they would cost and what type of company would benefit from that particular framework. It would be the company's responsibility to investigate what security framework would match their specific needs.

4 Conclusions

In comparison to all the authentication and authorisation structures examined, Sood's (2012) Integrity and Authentication Framework would be the most efficient to enhance security in a cloud computing network. This is primarily because it has been tested and experimental results prove it's a secure implementation, and as a consequence will enhance current systems to improve security.

Contrastingly, Nakal et. al.'s (2012) Mutual Authentication Framework contains no experimental results which won't be able to fully justify their conclusions.

The implementation could prove to be slower than other systems evaluated and have performance issues (Nakal et. al. 2012) The New User Authentication and File Encryption proposed by Nafi (2012) on the other hand did conduct an experiment and gained results, however, these were more focused on the speed of the cloud network, rather than security.

The author said that the Information leakage probability was low, however, didn't conduct a suitable and relevant experiment for this aspect within cloud networks. Additionally, Chadwick and Fatema's (2012) also had the same problem; this was not specific with regards to security features. There was a recurring theme from the experimental results which was primarily to do with the speed of the cloud network and not based around security.

Due to rapid increase of Cloud Computing, there will be a corresponding amount of technologies trying to gain unauthorized access. To protect our systems, now is the time to act upon cloud computing securities.

References

Bernabe, J.B; Perez, J.M.M; Calero, J.M.A; Clemente, F.J.G; Perez, G.M; Skarmeta, A.F.G; 2012, 'Semantic-aware multi-tenancy authorisation system for cloud architechtures', *Future Generation Computer Systems*, ScienceDirect [Online], 'In Press, Corrected Proof', http://dx.doi.org/10.1016/j.future.2012.05.011

Bhati, S; Bhati, A; Sharma, S.K; 2012, 'A New Approach towards Encryption Schemes: Byte – Rotation Encryption Algorithm', *Lecture Notes in Engineering and Computer Science*, Vol.2201, No.1, P979-983

Chadwick, D.W; Fatema, K; 2011, 'A privacy preserving authorisation system for the cloud', *Journal of Computer and System Sciences*, Vol.78, No.5, P1359-1373

Jayarana, I.G; Cahyawan, A.A; Sasmita, G.M; 2012, 'Dynamic Mobile Token for Web Security using MD5 and One Time Password Method', *International Journal of Computer Applications*, Vol.55, No.6, P1-6

Kadam, K.D; Gajre, S.K; Paikrao, R.L; 2012, 'Security issues in Cloud Computing', *National Conference on Innovative Paradigms in Engineering and Technology*, Vol.ncipet, No.11, P22-26

Nafi, K.W; Kar, T.S; Hoque, S.A; Hasham, M.M.A; 2012, 'A Newer User Authentication, File encryption and Distributed Server Based Cloud Computing security architecture', *International Journal of Advanced Computer Science and Applications*, Vol.3, No.10, P.181-186

Nayak, S.K; Mohapatra, S; Majhi, B; 2012, 'An Improved Mutual Authentication Framework for Cloud Computing', *International Journal of Computer Applications*, Vol.52, No.5, P.36-41

Rong, C; Nguyen, S.T; Jaatun, M.G; 2012, 'Beyond lightning: A survey on security challenges in cloud computing', *Computers and Electrical Engineering*, ScienceDirect [Online], 'In Press, Corrected Proof', DOI Link 'http://dx.doi.org/10.1016/j.compeleceng.2012.04.015'

Shaikh, R; Sasikumar, M; 2012, 'Security Issues in Cloud Computing: A survey', *International Journal of Computer Applications*, Vol. 44, No.19, P4-10

Singh, S; Jangwal, T; 2012, 'Cost breakdown of Public Cloud Computing and Private Cloud Computing and Security Issues', *International Journal of Computer Science and Information Technology*, Vol.4, No.2, p17-31

Sood, S.K; 2012, 'A combined approach to ensure data security in cloud computing', *Journal of Network and Computer Applications*, Vol.35, No.6, P1831-1838

Thiyagarajan, M; Dinesh, K.K; 2012, 'QR CODE AUTHENTICATION FOR PRODUCT USING CLOUD COMPUTING', *Journal of Global Research in Computer Science*, Vol.3, No.2, P5-8

Zissus, D; Lekkas, D; 2010, 'Addressing cloud computing security issues', *Future Generation Computer Systems*, Vol.28, No.3, P583-592

A Detailed Analysis Of Current Computing Research Aimed At Improving Facial Recognition Systems

Gary Adam Morrissey

Abstract

Facial recognition is an active area of research for many reasons, such as security and surveillance and is the only biometric identification and surveillance method that does not require restraint of the 'target'. In this paper the use of Neural Networks within facial recognition systems are analysed, compared and evaluated. This paper also makes recommendations on possible application of the methods in a 'real-world' environment.

1 Introduction

In the last 10 years, significant improvements have been made to facial recognition systems.

Biometrics has become an area of active research due to a wide variety of possible applications, such as security. Three types of biometric recognition systems exist; identification, verification and surveillance. While there are many strands to biometrics, facial recognition is the only method that does not require 'restraint' of the individual and can be used passively.

This is particularly beneficial for surveillance techniques and is more accurate than fingerprint recognition (El-Bakry, Abo-Elsoud and Kamel, 2000).

Many different techniques and algorithms have been created such as adaptive classifications (Connolly, Granger and Sabourin, 2012), neural networks & fuzzy logic (Fadzil and Choon, 1997) and dynamic particle swarm optimization (Connolly, Granger and Sabourin, 2012). The main downfall of facial recognition systems currently is that they are unable to achieve a successful recognition rate of 100%. Whilst a 100% match rate may not necessarily be required for some purposes, for example surveillance applications where further investigation can be carried out by humans once a potential match is made, security systems in terms of verification and identification would require a perfect system in place to ensure only authorized persons are able to access the information these systems protect, or to employ these biometric constraints as an extra barrier of security to current methods.

The contribution of this paper is to investigate and discuss various aspects of facial recognition systems and techniques, rather than the difficulties in face detection, i.e. determining if a face is present or not, in the media being used. Papers by El-Bakey, Abo-Elsoud & Kamel (2000) and Rowley, Baluja and Kanade (1998) discuss these issues in further detail.

2 Facial Recognition Systems

Face recognition algorithms can be categorized into two main groups, holistic (global) feature and local. In a holistic system, the entire face image is saved in the system as input data, such as Eigenfaces (Turk and Pentland, 1991) and Fisherface (Belhumeur, Hespanha and Kriegman, 1997). Local approaches use extraction of features such as the eyes and nose which are segmented and used as input data. Figure 1 shows a typical facial recognition system processes.

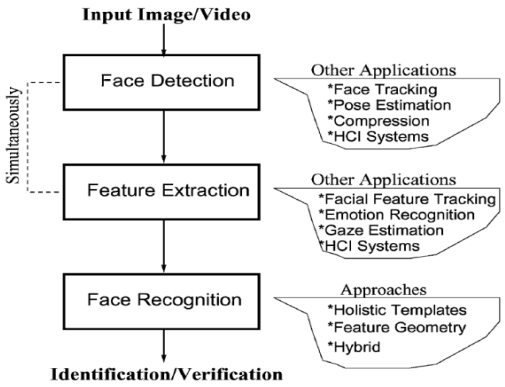


Figure 1 – Configuration of a generic face recognition system
(Zhao et al., 2003)

2.1 Neural Network Based Systems

Neural Network systems involve simple processing elements, passing information through parameters to determine an output from the information fed into the system. See figure 2 for an example of a Neural Network.

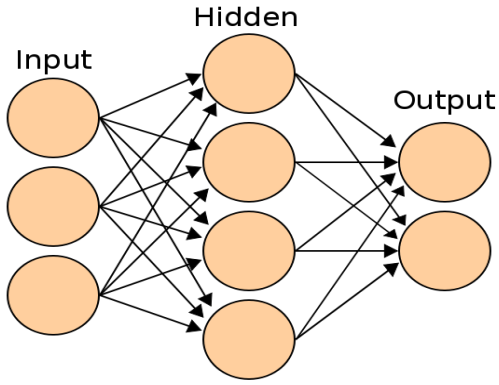


Figure 2 – Simple Neural Network example

2.1.1 Multi-Layer Perception (MLP) Neural Network Based System

A Multi-Layer Perception (MLP) Neural Network Based Face Recognition System using a Constructive Training algorithm has been tested against images on a UMIST database. The database contained 1012 images of 20 people (633 images as the training set, 40 the validation set and 339 used as the test set) for the system (Boughrara, Chtourou and Amar, 2012).

Figure 3 – Examples of images in database
(Boughrara, Chtourou and Amar, 2012)

In order to ensure consistent contrast of the images and to reduce illumination variances, all of the images in the database had undergone histogram equalization (Celik, 2012).
Boughrara, Chtourou and Amar (2012) conduct tests of their Constructive Training Algorithm using Zernike Moment and Gabor Features (for extraction of facial features for recognition application) against an MLP with back-propagation algorithm.

2.1.1.1 Zernike Moment

Zernike Moment methods have received considerable attention due to the invariance when scaling images and where rotation of images is needed.

Tables 4 and 5 show the recognition rates using the MLP and the Constructive Training Algorithm respectively.

Number hidden neuron	Recognition Rate (%)
10	89,38
15	91,15
20	91,74
25	91,44
30	89,38
35	91,15
40	91,15
45	90,26
50	89,67

Table 4 – Recognition on the UMIST database using Zernike Moment and MLP
(Boughrara, Chtourou and Amar, 2012)

Eps	epochs	N_hid	Recognition rate (%)
0.01	5000	28	83.48
0.01	10000	45	85.54
0.01	15000	28	81.41
0.01	20000	18	82.59
0.007	5000	15	90.85
0.007	10000	22	94.98
0.007	15000	63	89.50
0.007	20000	18	86.43
0.005	5000	92	87.9
0.005	10000	81	97.34
0.005	15000	28	96.16
0.005	20000	25	91.15

Table 5 – Recognition on the UMIST database using Zernike Moment and the Constructive Training Algorithm
(Boughrara, Chtourou and Amar, 2012)

The results show that the Constructive Training Algorithm can produce a higher recognition rate (97.34%) against that of the Multi-Layer Perceptron (91.74%).

2.1.1.2 Gabor Features

Gabor features have proved to be one of the most successful approaches for face recognition, due to its performance despite illumination changes and expression variances.

Tables 6 and 7 show the recognition rates using the MLP and the Constructive Training Algorithm respectively.

Number hidden neuron	Recognition Rate (%)
10	94.39
15	94.39
20	94.39
25	94.39
30	94.98
35	94.69
40	93.51
45	94.39
50	94.98

Table 6 – Recognition on the UMIST database using Gabor Filters and MLP
(Boughrara, Chtourou and Amar, 2012)

Eps	epochs	N_hid	Recognition rate (%)
0.01	500	9	86.13
0.01	1000	9	91.74
0.01	5000	16	92.92
0.01	10000	23	86.13
0.007	500	45	90.56
0.007	1000	15	92.33
0.007	5000	16	96.16
0.007	10000	47	92.03
0.005	500	178	95.87
0.005	1000	47	94.98
0.005	5000	8	97.05
0.005	10000	12	94.98

Table 7 – Recognition on the UMIST database using Gabor Filters and the Constructive Training Algorithm
(Boughrara, Chtourou and Amar, 2012)

Once again, the Constructive Training Algorithm out performs the MLP, with the highest rates of recognition at 94.98% and 97.05% respectively.

The results show an increase in accuracy using the constructive training algorithm compared with the MLP. The conclusions of the paper are in line with the results that have been achieved, the Constructive Training Algorithm has been tested using Zernike Moments and Gabor Features and out performs the Multi-Layer Perceptron.

Figure 3 showed examples of images that have been stored in the database and the experiments are conducted in a way to ensure that variables and discrepancies are minimized where possible, selecting the training set images at random, covering a range of appearances, genders and race and using PGM format images, all sized 220 x 220 pixels. This ensures that the results data is justified and, more importantly, credible.

2.1.2 Neural Network and Fuzzy Logic System

Using feature extraction, Fadzil and Choon (1997) created a system that examined three areas of the face; upper face, lower face and eye. Trained upper face and lower face neural networks are employed to recognize the upper face and lower face of an image respectively. The fuzzy system is used to check the

recognition information provided and acts as the final decision making process.

Incorporating this into a smartcard system as the second level security protection to identify the card owner, the system has been tested in both offline mode (to identify system improvements) and online mode (to reveal overall performance). Images of six people are used to train the respective networks (upper & lower face and eye).

Offline testing concluded that the eye network encounters difficulties detecting eye location when the subject is wearing spectacles. This was expected as the network was not trained with images of users wearing spectacles. Reflection from spectacles can also result in failure to detect the eye as can head orientation.

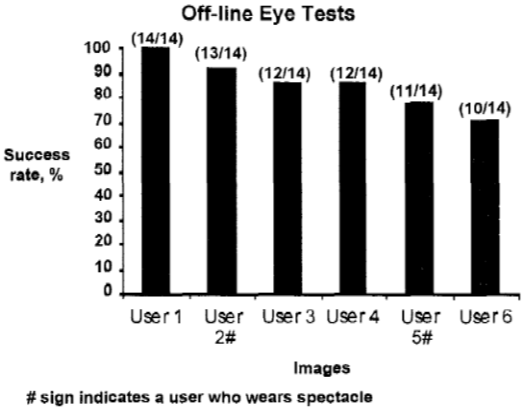


Figure 8 – Results from offline eye tests. Failures due to: i) Eye outside search area, ii) Head tilted beyond system limitations, iii) Wrong determination of eye location, iv) Spectacles problem, v) Partially closed eye(s)
(Fadzil and Choon, 1997)

Offline testing showed limitations of the system and areas to improve upon. Several failures occurred during the testing of face recognition, which were found to be because of poor recognition of lower face images. By experimenting, allowing the system to accept a greater degree of freedom in facial expression dramatically increased the recognition capability.

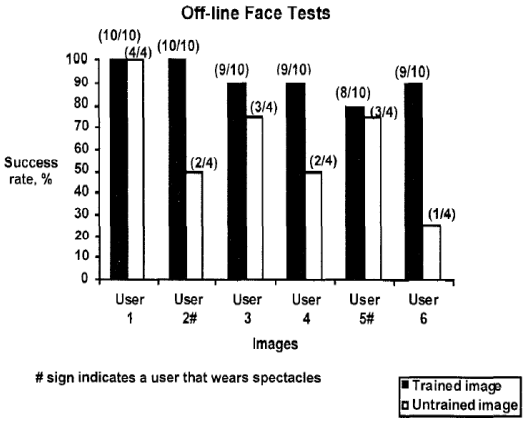


Figure 9 – Results from offline face tests. Failures due to: i) Eye detection failure, ii) Unable to recognise lower face
(Fadzil and Choon, 1997)

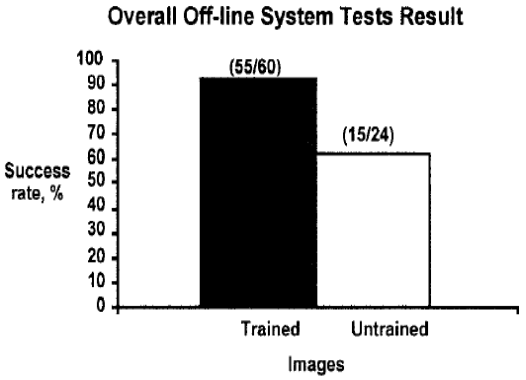


Figure 10 – Overall offline system test results for trained and untrained images
(Fadzil and Choon, 1997)

Testing their methods on an online version of the software, Fadzil and Choon (1997) conducted two major tests to evaluate the system, authorized and unauthorized card owner tests. Figure 11 shows the results of these tests, with a 94% success rate for authorized and 92.8% success rate for unauthorised cases.

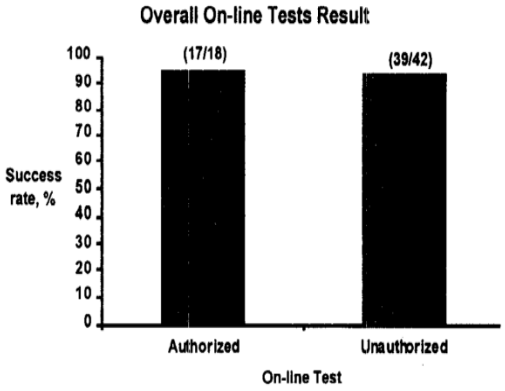


Figure 11 – Overall Online Test Result
(Fadzil and Choon, 1997)

Whilst the conclusion of the report states that the face recognition system in security could be applied to high security situations, such as automated teller machines, security room access and confidential information access facility (Fadzil and Choon, 1997), they reveal that the face recognition system would not work for images with eyes closed and images of bald people.

This system may be able to be employed in these methods stated with the limitations of eye closure, as video based methods would record the eye opening and images could be captured when requested by the user in order to ensure the eyes are completely open. However, further investigations into the issue of recognizing those faces of people who are bald will need to occur, as recent studies show that "more than 30% of adult men go completely bald by age 55" (What is the percentage of bald people in the world, 2011). This would obviously cause issues for systems using this type of biometric information.

Whilst the conclusions presented state that the system has not been trained with images of persons wearing spectacles, it has not been discounted that this will not work with spectacle wearers. Training the system to recognize faces with spectacles should be investigated. Whilst this may not cause an issue for verification and identification purposes, similar to the methods used at airports, such as Manchester, UK (BBC, 2008), where the subject interacts willingly with the system and could remove spectacles, applications such as surveillance would suffer shortcomings using this system.

3 Comparison of Methods

The tests and results made in these papers are robust and sound in terms of what the authors wished to achieve and the results they have published to the wider science community.

A rigorous criterion was drawn up by Fadzil and Choon (1997) for their training of the neural network system (11 different images for each person). During their testing, it was discovered that the system was unable to recognise faces where spectacles were present. The training data did not involve training the system with images of faces with spectacles. Further investigations into this method would be advantageous in order to discover the system's ability to handle those faces where spectacles are being worn.

The results achieved by Boughrara, Chtourou and Amar (2012) show that they have achieved a higher recognition rate to those achieved by Fadzil and Choon (1997). Testing their method against two varying feature vectors, their method of a Constructive Training Algorithm outperformed the system they were comparing it with and the Neural Network and Fuzzy Logic approach.

Whilst the work by Boughrara, Chtourou and Amar (2012) show statistically higher recognition rates compared with the work of Fadzil and Choon (1997), they have not presented any reasoning or hypothesis for the failure rates. These rates may have been caused from other grounds, but also may have been caused by the same issues Fadzil and Choon (1997) encountered. Further investigations into the failure rates are necessary to provide the next stages of research in order to address these hindrances.

Further investigations into the images and other variables of User 1 within Figure 8 of the research would also be beneficial. Does this person have predominant features or markings that make it easier for the system to distinguish this person against the other stored information within the database for example.

4 Application of Methods

Work completed by Dadgostar, Bigdeli and Smith (2011) around face enrolment and recognition within a CCTV network links in well with both pieces of research. In their system, simple constraints are added to the system, such as a person can only appear once in a single frame and "it takes at least N seconds to move from camera 1 to camera 2", in order to avoid false positive alerts. Surveillance systems are employed in order to be able to identify persons of interest, rather than as 100%, fool proof identification that can be used to access secure areas, be it compounds/locations or folders stored on PC's.

"Based on the current technologies, automatic inference of the current location of a passenger is not achievable, mainly due to limitation of facial recognition systems and limited coverage of CCTV networks" (Dadgostar, Bigdeli and Smith, 2011).

This task would be performed reliably by humans who are experts in this field, but the manpower and resources for this to be done manually would be a drawback. Implementing methods such as the ones discussed throughout this paper could enable identification of 'persons of interest' and alert authorities to investigate further, speeding up the identification process and reducing man hours spent on this task.

Unless a system that can be proved at 100% accuracy of recognition, it would be foolish to use biometric identification as a stand-alone measure for accessing sensitive information.

5 Conclusions

In this paper, two facial recognition systems have been evaluated. Both methods described in this paper show significant advancements in facial recognition but, as their conclusions and results show, obtaining 100% identification rates overall was illusive. This is not solely a technology issue however. Rates of 100% were recorded during tests completed by Fadzil and Choon (1997) – see Figure 8, User 1. Issues exist around changes in the face. When a face is detected, it may not be the same as the image recorded on the system. Ageing factors or scarring of the face can impact on the recognition rate of systems, as well as variances in head orientation.

5.1 Secure Access Systems

Both systems have proved that, whilst they could not be implemented as a stand alone access method, employing them as a second line of security in an access system can go some way to ensure that the person accessing sensitive information, entering secured environments or entering PIN details at a bank ATM, is the owner of the access "key" (be it a credit card or metallic key).

Biometrics is undoubtedly the next stage of research for security systems. "Until today, humans have proved or verified their personal identity in their interactions with machines either through owning a special possession, such as a metal key or a plastic card, or through having secret knowledge such as a password or PIN Number. (Metal keys date back 5,500 years ago to the Bronze Age, and passwords at least to Roman Centurions" (Daugman, 1997).

5.2 Surveillance systems

Integrating these systems into surveillance systems could improve the security in such arenas as airports and rail stations, which have become targets of terrorism in recent times. The ability of a system to identify possible terror suspects and alert authorities would be a massive advantage to those agencies charged with the protection of the public.

Both methods have proved their worth in this arena and whilst a false positive alert may be of an inconvenience to the target and security staff, further research should be targeted at ensuring false authorisations/detections are minimised.

References

BBC (2008) *Passengers test new face scanners*, 19 August, [Online], Available: HYPERLINK "http://news.bbc.co.uk/1/hi/uk/7568686.stm" http://news.bbc.co.uk/1/hi/uk/7568686.stm [12 December 2012].

Belhumeur, P.N., Hespanha, J.P. and Kriegman, D.J. (1997) 'Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, July, pp. 711 - 720.

Boughrara, H., Chtourou, M. and Amar, C.B. (2012) 'MLP neural network based face recognition system using constructive training algorithm', *International Conference on Multimedia Computing and Systems (ICMCS)*, 233 - 238.

Celik, T. (2012) 'Two-dimensional histogram equalization and contrast enhancement', *Pattern Recognition*, vol. 45, no. 10, October, pp. 3810 - 3824.

Connolly, J.-F., Granger, E. and Sabourin, R. (2012) 'An adaptive classification system for video-based face recognition', *Information Sciences*, vol. 192, June, pp. 50 - 70.

Connolly, J.-F., Granger, E. and Sabourin, R. (2012) 'Evolution of heterogeneous ensembles through dynamic particle swarm optimization for video-based face recognition', *Pattern Recognition*, vol. 45, no. 7, July, pp. 2460 - 2477.

Dadgostar, F., Bigdeli, A. and Smith, T. (2011) 'Demo: An automated face enrolment and recognition system across multiple cameras on CCTV networks', *Fifth ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC)*, Ghent, 1 - 2.

Daugman, J. (1997) 'Face and Gesture Recognition: Overview', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, July, pp. 675 - 676.

El-Bakry, H.M., Abo-Elsoud, M.A. and Kamel, M.S. (2000) 'Automatic Face Recognition System using Neural Networks', *The 2000 IEEE International Symposium on Circuits and Systems*, Geneva, 543 - 546.

Fadzil, A.M.H. and Choon, L.C. (1997) 'Face Recognition System based on Neural Networks and Fuzzy Logic', *International Conference on Neural Networks*, 1638 - 1643.

Jain, A.K., Ross, A. and Pankanti, S. (2006) 'Biometrics: A Tool for Information Security', *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, June, pp. 125 - 143.

Rowley, H.A., Baluja, S. and Kanade, T. (1998) 'Neural Network-Based Face Detection', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23 - 38.

Turk, M.A. and Pentland, A.P. (1991) 'Face Recognition Using Eigenfaces', *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 586 - 591.

What is the percentage of bald people in the world (2011), 9 August, [Online], Available: HYPERLINK "file:///C:\\Users\\Simon\\AppData\\Local\\Microsoft\\Windows\\Temporary%20Internet%20Files\\Content.Outlook\\35JC635S\\www.chacha.com\\question\\what-is-the-percentage-of-bald-people-in-the-world" www.chacha.com/question/what-is-the-percentage-of-bald-people-in-the-world [11 December 2012].

Zhao, W., Chellappa, R., Phillips, P.J. and Rosenfeld, A. (2003) 'Face Recognition: A Literature Survey', *ACM Computing Surveys*, vol. 35, no. 4, pp. 399-458.

A Critical Analysis Of Current Research Into Stock Market Forecasting Using Artificial Neural Networks

Chris Olsen

Abstract

Stock market forecasting is used by individuals and companies alike in order to reduce the level of risk inherent in trading. A model which can predict the value of a company, commodities or other traded stock is a useful tool and forms part of a successful trading strategy. Research in this area has focused on the improvement of artificial neural networks to detect nonlinear relationships between the various aspects of the market; this paper considers, compares and critically evaluates four of these neural network models and makes recommendations for future research as well as how they could be applied in a hybrid architecture to further improve forecast accuracy.

1 Introduction

This research paper focuses on current research into the techniques being used to forecast stock market indexes and prices. According to Atsalakis and Valavanis (2009) refining methods for stock market forecast is a key element of a successful trading strategy. As such, research focusing on the use of artificial neural networks (ANN) and their ability to detect non-linear relationships between data has become a major focus of interest in recent years (Simon and Raoot 2012). Developments in this sector have reached a point in which artificial computing models outperform conventional models in the majority of cases (Atsalakis and Valavanis 2009), (Simon and Raoot 2012). Further research involves perfecting current models to improve the level of accuracy achieved.

This paper will focus upon the approaches of modeling ANN's to assess which method (or combination) yields the best results for stock market forecasting. The main emphasis of this paper will be on the network model although consideration will be given to the learning algorithms which are used for training.

2 Artificial Neural Networks

Research into stock market forecasting has identified Artificial Neural Networks (ANN) as an effective tool for prediction due to its ability to detect non-linear patterns and adaptation to noisy datasets (Simon and Raoot 2012). There are a variety of networks which could be used; this section will look at a range of ANN and discuss the work being conducted in that field.

2.1 ANN and Standard Statistical Techniques

A standard ANN has many ways it can be trained to model specific problems, research by Vaisla and Bhatt (2010) sought to prove ANN were more effective at stock market prediction than standard statistical techniques.

The researchers used two years' worth of data (April 2005 until March 2007) from the National Stock Exchange of India, specifically the NIFTY exchange.

The statistical technique against which the NN model was being compared is Multiple Regression Analysis; Figure 1 shows the results of the experiments.

	SSE	MAE	MSE	RMSE
NN	0.001137459	0.00138	2.28E-06	0.00151
REG	0.105770461	0.010517	0.000212	0.014559

Figure 1 – Comparison of Neural Network and Regression forecasting models (Vaisla and Bhatt, 2010)

As claimed by Vaisla and Bhatt (2010) "Neural Networks performs well than compared to Statistical forecasting of daily closing Nifty values because the error in Neural Network is very less than the Statistical method." They go on to conclude that their research "depicts the fact that Neural Networks outperform Statistical technique" (Vaisla and Bhatt 2010).

While their claims are generally backed up by the evidence they have provided, the wider application of their research is somewhat limited. Despite discussing the need for training the neural network and the emphasis put on finding the appropriate architecture the researchers fail to make mention of key elements of the NN model such as the number of input nodes, what training algorithm was used, or the number of layers being used so reproducing the results or making adjustments to the model with the aim of improving its performance would not be possible using the information in their research alone.

There is also very little discussion regarding the specific experimental methodology which was carried out.

Furthermore their final conclusion in which Vaisla and Bhatt (2010) claim Neural Networks outperform statistical techniques is slightly too generalized to be drawn solely from this single piece of research. Only one type of neural network and one type of statistical technique were compared therefore it is over-reaching to say all NN's will outperform all statistical techniques, even within the somewhat limited domain of predicting the NIFTY value.

Further experiments could involve using several types of Neural Networks and comparing those to several types of statistical techniques, at which point it would be more accurate to claim that in those cases one method was more accurate than the other.

2.2 GMDH Networks

GMDH (Group method of data handling) neural networks are based on a model of self-organization, as discussed by Abdalla et al. (2012), the unbounded nature of the network allows the independent variables to randomly shift in order to find the best match for the dependent variables. This technique places high importance on the inputs and thus the neurons which are activated to achieve the most effective result. Abdalla et al. (2012) went on to assert that GMDH was the most suitable method of predicting stock market prices due to its ability to find the most fitting solution based on the external criterion.

An optimized algorithm was presented for use in which "the neurons are selected one-by-one and then added to a binary network according to an external criterion" (Abdalla et al. 2012). Simulations were carried out using this methodology with two training algorithms (Delta Rule and Genetic Algorithm) to determine which combination performed best with 141 months of data. The two models were compared to a biquadratic GMDH NN where it was concluded that the Inductive model with the Delta Rule training algorithm was the most accurate, "Inductive model A gives the best performance" (Abdalla et al. 2012).

Figure 2 shows the experimental results with the mean squared error (MSE) backing up their claim that Model A did in fact produce the lowest MSE value and was therefore the most accurate of the tested methods.

No	Method	MSE
1	Biquadratic GMDH NN	0.0004955
2	Proposed Inductive Model A	0.0002404
3	Proposed Inductive Model B	0.0009048

Figure 2 – Comparison of the Mean Square Error Results of the three tested models showing Model A has the lowest error rate. (Abdalla et al., 2012)

While it is true from the data sets and algorithms which were tested that Inductive Model A was the most accurate, the research is somewhat limited by the fact that it was compared to only one other separate method (the Biquadratic, as Model B was the same technique trained differently) whereas further experimentations and comparisons may have led to more comprehensive claims of improved accuracy.

In addition the training of the two models varied somewhat with the first model being trained for 2000 cycles for the first Adaline and 1500 for

the second, contrasted with the training given to Model B which was 3000 cycles per Adaline (Abdalla et al, 2012). As such over-fit or under-fit could be an issue with Model B performing less effectively as it should or alternatively Model A not being as effective as it potentially could be.

Furthermore, research by Abdalla et al. (2012) did not take into account the impact of noisy data. Other research in this field sought "to reduce the effects of noise through the linear and nonlinear smoothing techniques" (Dai et al. 2012). Guresen et al. (2011) also researched the influence of noise relating to stock market prediction, highlighting the problems associated with reproducing real world conditions "The noise that caused by changes in market conditions, it is hard to reflect the market variables directly into the models without any assumptions" (Guresen et al., 2011).

2.3 D-GMDH Networks

Research into GMDH models was also the topic of investigation by Zhang et al. (2012) who proposed a dynamic model with the aim of creating "noise-immune forecasting algorithms." (Zhang et al. 2012). This was attempted by calculating a diversity value and using the most disparate models as the input in order to reduce noise and improve the accuracy of the prediction, "using the most diverse models as initial input models may provide better forecasting accuracy." (Zhang et al. 2012). Figure 3 shows the generation of the new initial elementary input models.

In order to test their theory Zhang et al. (2012) compared the results to various other models such as, "the traditional GMDH model, GMDH combination model and ARMA (Auto Regression Moving Average) model." (Zhang et al. 2012).

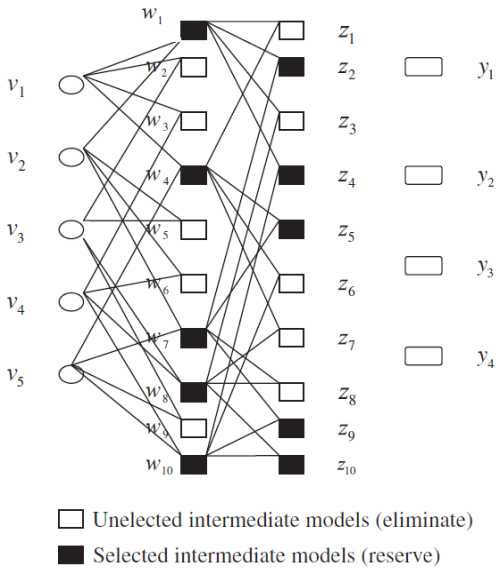


Figure 3 – Model of the First layer generation of candidate models. (Zhang et al., 2012)

	Time	June,2009	July,2009	August,2009	RMSE
	y	587.93	490.13	505.62	
(hundred million RMB)					
D-GMDH	Forecasted value	561.79	499.90	476.39	23.42
	Relative errors	−4.4%	2.0%	−5.8%	
GMDH-Combination	Forecasted value	535.00	491.27	514.86	31.01
	Relative errors	−9.0%	0.2%	1.8%	
ARMA	Forecasted value	530.25	499.05	499.90	33.87
	Relative errors	−9.8%	1.8%	−1.2%	

Figure 4 – Comparison of the results from 3 forecasting models. (Zhang et al., 2012)

The results of their experiments are provided in Figure 4 which shows the performance of each model and their corresponding root mean squared error (RMSE), the calculation for which is shown in Figure 5.

$$\sqrt{\frac{1}{n} * \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

Figure 5 – Calculation used to calculate the RMSE of each method. (Zhang et al., 2012)

Zhang et al. (2012) concluded in their research that the dynamic method outperformed the traditional GMDH as can been seen in Figure 6.

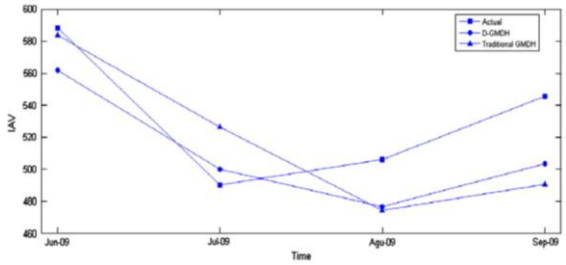


Figure 6 – Graph plotting the Actual value against the standard and dynamic models. (Zhang et al., 2012)

However despite using a GMDH combination model and ARMA in their initial experiments neither were included in the time-lapse (June 2009 until September 2009) forecast.

Therefore claims that the "D-GMDH has better forecast accuracy" (Zhang et al. 2012) are not as credible as they could have necessarily been had all of the discussed models been tested using the same parameters. Also the selected time frame is very limited, not giving the models enough time to conclusively prove they are accurate with long term predictions.

Despite this however it can be seen from research conducted by both Abdalla et al., (2012) and Zhang et al. (2012) that GMDH have great potential for stock market forecasting, claims which have been backed up in additional research by Bing et al. (2011) and Sung-Kwun and Witold (2006).

To compare the models from Abdalla et al., (2012) and Zhang et al. (2012), both present valid, if not completely definitive, results however as "economic forecasting using noisy and volatile datasets is an important topic" (Zhang et al. 2012), and giving consideration to the noisy data is essential in long term forecasting, it would be recommended that Inductive model A proposed by Abdalla et al., (2012) was refined to take into consideration noisy data, or alternatively use the output of this model as the input of a dynamic GMDH to further improve accuracy. The models should also be tested over a longer period of time before any conclusive claims regarding long term prediction accuracy can be made.

2.4 Hybrid Backpropagation Neural Networks

Backpropagation networks are one of the most popular types of ANN, "The backpropagation neural network (BPN) is feedforward network and is probably the most commonly used class of neural network in financial time series forecasting and business." (Dai et al. 2012). Figure 7 shows a typical BPN topology.

Nonlinear independent component analysis (NLICA) is a "novel feature extraction technique to find independent sources from observed nonlinear mixture data." (Dai et al. 2012).

Dai et al., (2012) research into the development of a hybrid NLICA-BPN model aimed at improving prediction methods using the non-redundant, independent statistical components which are the product of NLICA.

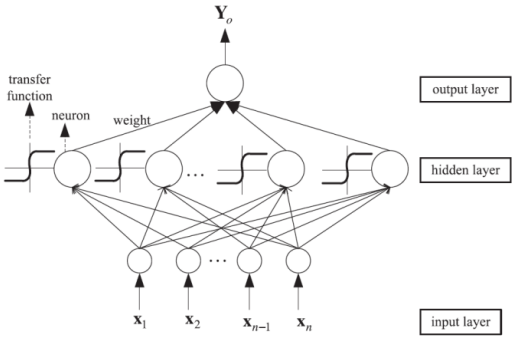


Figure 7 – An example of a typical backpropagation neural network topology. (Dai et al., 2012)

Experiments involved "creating an optimal network topology of the BPN" (Dai et al. 2012), using four variables at the input layer with the output providing the closing price of the Nikkei 225 index. The NLICA-BPN model was compared to three other models; the standard BPN (identified as Single BPN in figure 7), PCA-BPN (Principal Component Analysis) and LICA-BPN (Linear Independent Component Analysis).

Figure 8 shows the results of the experiments, from which the researchers claim that the "model can produce lower prediction error and higher prediction accuracy and outperformed the

LICA-BPN, PCA-BPN and single BPN models" (Dai et al. 2012).

From the experimental results, the error rate is substantially less than the other 3 tested methods, the experimental methodology was focused on predicting a single cash price value (Nikkei 255) using substantial datasets (Feb 2004 until March 2009), each of the models were tested using the same data and then the experiment was repeated using the Shanghai B-Stock index (Figure 9). Consideration was given to the ratio of training to testing samples during the evaluation of the model and were "based on the relative ratio of the size of the training dataset size to complete dataset size" (Dai et al., 2012).

Metrics Models	RMSE	MAD	MAPE (%)	RMSPE (%)	DS (%)
NLICA-BPN	**50.44**	**39.78**	**0.242**	**0.302**	**85.69**
LICA-BPN	202.77	143.86	0.941	1.389	73.92
PCA-BPN	198.78	145.70	0.947	1.352	74.85
Single BPN	263.60	216.50	1.353	1.671	77.77

Figure 8 – Forecast results for the Nikkei 225 of the four models tested models; NLICA-BPN, LICA-BPN, PCA-BPN and single BPN models. The proposed model NLICA-BPN is shown in bold with the lowest RMSE. (Dai et al., 2012)

Metrics Models	RMSE	MAD	MAPE (%)	RMSPE (%)	DS (%)
NLICA-BPN	**1.67**	**1.25**	**0.95**	**1.29**	**80.50**
LICA-BPN	1.73	1.37	1.01	1.29	78.26
PCA-BPN	2.39	2.00	1.48	1.78	79.50
Single BPN	2.18	1.63	1.24	1.68	72.67

Figure 9 – Forecast results for the Shanghai B-Share index of the four models; NLICA-BPN, LICA-BPN, PCA-BPN and single BPN models. The proposed model NLICA-BPN is shown in bold with the lowest RMSE. (Dai et al., 2012)

The conclusions provided by the researchers are reasonable and not over-reaching as they effectively demonstrated NLICA-BPN's accuracy over three main variants of the BPN model. Further research could focus on comparing this method to other ANN to provide more credibility to their claims.

The main criticism of the model is that the network only accepts 4 input variables, "The input layer has four nodes as four forecasting variables are used," (Dai et al. 2012). As stated by Simon and Raoot (2012), "Further research is anticipated to incorporate additional inputs that influence stock returns with this neural approach", and backed up by Yakup et al. (2011) that macroeconomic influences should be taken into consideration within an ANN model.

As the model only accepts 4 inputs, there is a lot of external indirectly related information (such as political or environmental) which is not taken into consideration. In the long term this could affect the accuracy of predictions as the stock market is a chaotic system in which seemingly unrelated factors can affect an outcome in the future (Banik et al. 2012).

3.0 Comparison of Methods

Research performed by Vaisla and Bhatt (2010) while valuable in demonstrating the success of an ANN against a statistical method did not place a premium on improving the accuracy of the network and only achieved decreases in error in comparison to one standard method.

The work performed by Abdalla et al., (2012) and Zhang et al. (2012) however did provide novel ways of constructing a forecast model using ANN, and their work into GMDH and D-GMDH networks provided a solid foundation from which future research can be based. The noise tolerance of the D-GMDH method in particular shows promise as instead of ignoring outliers in the datasets it is able to weigh them appropriately and integrate it into the output.

Also the work conducted by Dai et al., (2012) on NLICA was significant as it provided an innovative way of identifying potential inputs for an ANN extracted from non-linear data, which was tested within a BPN. As the selected input variables are an important factor when attempting a forecast using any ANN, the work by Dai et al., (2012) has far reaching implications and application in other work.

4.0 Application of Methods

As noted, research by Dai et al., (2012) in feature extraction from non-linear datasets can be applied outside of the domain of stock market prediction and used for any ANN which is modelling a system which has a degree of chaotic behaviour and non-linear relationships.

In terms of the research by Abdalla et al., (2012) and Zhang et al. (2012), these too could be applied outside of the domain of stock market pre-

diction and would be suited not only to financial computation but systems such as weather prediction which have elements of nonlinear relationship influence.

5.0 Conclusions

This paper has surveyed four types of neural network architecture and the associated learning algorithms in relation to their ability to forecast stock market prices with the highest degree of accuracy.

Due to the chaotic nature of the stock market and the number of variables which affect individual prices, none of the methods can forecast with absolute certainty. Indeed the level of accuracy somewhat degrades over the long term, but artificial neural networks still outperform standard statistical techniques in the majority of cases and the design of the ANN architecture is integral to improving accuracy (Atsalakis and Valavanis 2009).

Of all the methods the GMDH networks have been shown to deliver accurate results (Abdalla et al. 2012), (Zhang et al. 2012) following on from this future research would not only involve more extensive testing of the models but also integration with other techniques to improve the accuracy rate.

Using the techniques discussed by Dai et al., (2012) in improving the quality of the input variables of the neural network, it would be recommended that a hybrid NLICA-DGMDH model was used. With the inputs garnered by the Nonlinear independent component analysis and then fed into the D-GMDH which has demonstrated notable noise tolerance, meaning two critical elements of the forecast (input and noise) will be given due consideration which should result in a more comprehensive model and therefore a more accurate prediction.

Additionally the number of the inputs into the neural network models could be increased, as the more facets of market conditions which are taken into consideration the more complete any prediction will be.
It is also worth noting research by Chang-le et al. (2012) who developed a predictive model which used a partially connected neural network (with potentially several hidden layers) and a genetic algorithm. This could be explored as part of a larger hybrid architecture which would provide a broader base of inputs for a NLICA-GMDH model or similar multi-layered network.

References

Abdalla T, Abdalla Abdulkareem Y, and Dexon LA, 2012, 'Stock Market Prediction using Inductive Models', *International Journal Of Computer Applications*, Volume 49, Issue 1, Pages 36 – 42

Atsalakis GS, Valavanis KP, 2009, 'Surveying stock market forecasting techniques – Part II: Soft computing methods', *Expert Systems With Applications*, Volume 36, Part 2, Pages 5932-5941

Banik S, Chanchary F, Rouf R, Khan A, 2012, 'Modeling Chaotic Behavior of Dhaka Stock Market Index Values Using the Neuro-fuzzy Model', *Recent Patents On Computer Science*, Vol. 5, Issue 1, Pages 72-77

Bing Z, Chang-Zheng H, Panos L, Xiao-Yu L, 2011, 'A GMDH-based fuzzy modeling approach for constructing TS model', *Fuzzy Sets And Systems*, Volume 189, Pages 19-29

Chang-le Z, Pei-Chann C, Di-di W, 2012, 'A novel model by evolving partially connected neural network for stock price trend forecasting', *Expert Systems With Applications*, Volume 39, Pages 611-620

Dai W, Jui-Yu W, Chi-Jie L, 2012, 'Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes', *Expert Systems With Applications*, Volume 39, Pages 4444 – 4452

Guresen E, Gulgun K, Tugrul D, 2011, 'Using artificial neural network models in stock market index prediction', *Expert Systems With Applications*, Volume 38, Pages 10389-10397

Simon S, Raoot A, 2012, 'Accuracy Driven Artificial Neural Networks In Stock Market Prediction', *International Journal On Soft Computing*, Volume 3, Issue 2, Pages 35 – 45

Sung-Kwun O, Witold P, 2006, 'Multi-layer self-organizing polynomial neural networks and their development with the use of genetic algorithms', *Journal Of The Franklin Institute*, Vol. 343, Pages 125-136.

Vaisla K S, Bhatt A K, 2010, 'An Analysis of the Performance of Artificial Neural Network Technique for Stock Market Forecasting', *International Journal On Computer Science And Engineering*, Volume 2, Issue 6, Pages 2104 – 2109

Yakup K, Melek Acar B, Ömer Kaan B, 2011, 'Predicting Direction Of Stock Price Index Movement Using Artificial Neural Networks And Support Vector Machines: The Sample Of The Istanbul Stock Exchange', *Expert Systems With Applications*, Vol. 38, Issue 5, Pages 5311-5319

Zhang M, Changzheng H, Xin G, Panos L, Bing Z, 2012, 'D-GMDH: A novel inductive modelling approach in the forecasting of the industrial economy', *Economic Modelling*, Issue 30, Pages 514 – 520

Evaluation of User Authentication Schemes

Sukhdev Singh

Abstract

This research paper shows and evaluates the abuser authentication methods which developed to improve the security of online computing. Paper enclosed the comparison of different remote user authentication schemes which proposed till now. It describes the password based, smart cards and biometric based schemes and shows the drawbacks, risks and advantages of all. The key piece of work in this paper demonstrates how to make safe communication between remote user and service provider/server. It covers the weak point of authentication methods which often results into online attacks such as masquerade attacks and server spoofing attacks. The heavy part of paper described, why and how biometric technology plays an important role in user authentication methods.

1 Introduction

Online security is the major issue in IT world which needs to be geared up. Security matters a lot when term online is comes whether its privacy of service provider or user both subject lots. User authentication is the way used by online service provider to give user an access to their online structure. Usernames and traditional passwords are not as secure to protect private information from online thieves. Simple passwords are easily broken by internet frauds by assigning dictionary attacks to gain access to people's private data.

Lamport (1981) proposed first famous remote user authentication method with smart-card which stores the password key at service provider server to verify strength of the log-in request by user, nevertheless confusion in front and requirement for password retuning reduce the appropriateness and realistic capability of lamports's scheme. Since than many similar schemes have been proposed, they all have general feature which store the password table at the authentication server. Fundamentally this property is the weakness of the security method.

Afterward a further client authentication method projected by Jan-Chan (1998) in contrast to previous schemes it avoids storing password table at system. But Li-Hwang (2000) pointed out that due to the amount of authentication tables which are relative to numeral remote consumer it is unsuccessful for server to keep

all this authentication record. Then next Hwang and Li (2000) intended remote abuser authentication technique with the use of Public-key cryptosystem which designed by ElGamal in 1985. Hwang-Le proposal keeps just one secret input and there is no need to uphold the password table at all in the system. Until now various processes for user confirmation based on passwords and smart cards has been done (Sun, H, 2000; Chien, H et al. 2002; Shen, J et al. 2003). Biometric key technology based schemes Hwang, L (2009); Zhang, K (2007); han et al. (2008); Lai-Lin (2004) are designed to fulfil the drawbacks of previews password authentication schemes. Biometric keys method is based on behavioural and physiological character of individuals for example finger, voice, iris, face and hand. Benefits of biometric keys are, it cannot be lost or forgotten and cannot be guessed easily, it's not easy to copy or share. As a result biometric based user authentication method is innately more confident than usual password based user authentication methods. But still there are many attacks which are not blocked by recent proposed biometric based schemes. Lee et al. (2002) design finger-print based client authentication method but it was unsuccessful as

well to endure masquerade attack (Lin-Lai, 2004; Hsieh et al. 2003). Moreover Lin - Lai (2004) biometric based scheme also not continue too long, the problem Zhang-Khan (2007) pointed in Lin-Lai scheme is that their scheme putting the server to risk of spoofing attack.

So far study of user authentication methods is indicate that biometric based schemes are more secure than traditional password based schemes. However problem of user authentication is still not cleared and require further study to mitigate the security risk of internet based business. Many researches were been done and several on the way and hopefully various further will begin soon to make online computing strong enough.
Rest of paper shows the comparison of recently developed user authentication methods i.e. biometric technology based schemes vs. traditional password based and smart card based schemes. Then next Conclusion of survey paper is shown at last before references list.

2 Comparisons of Recent Studies

Yoon et al. (2004) intended a user authentication method which is planned on password and smart-card. But that's not the traditional password it's the cryptographic secret key which is however stored on smart card while registration phase. The problem of cryptographic key is that it's too hard to remember therefore Yoon et al. ended a work with which user no more need to remember the cryptographic key but instead of that PIN (Personal Identification Number)/Password is needed. There it seems high risk again. However if attacker found or steal the smart card of user then it is possible for attacker to extract the information from smart card and then attacker can claim it as real user to server (Hsiang, H et al. 2009). The possible attacks which attacker can able to perform with found smart card are Masquerading attack, Password guess attack and Parallel session attack. But Hsiang et al. (2009) didn't show any experimental work which shows that Yoon et al. Scheme is really vulnerable to these attacks. There are some other remote user authentication schemes proposed as well which based on similar methods but reason for lacking behind is using password technology. Whereas Lee et al. (2002) user verification scheme also based on smart-card but instead of just password (or PIN) fingerprint biometric technology is used to verify the user identity.

However Lin and Lai (2004) found some problems which Lee et al. (2002) method faces. Lee et al. method is consist of three stages first one is Registration phase where user enrol his/her information, second is Login phase where user request for logging into system and third and final is Authentication phase where server check the information of user which he/she enter for log in request with the information which user enrolled while registration phase if the information will equivalent to each other than server allow user to logging in otherwise user request is cancelled. But the availability of changing password is not providing in Lee et al. Scheme and it is also helpless to masquerade attack (Lin-Lai, 2004). Lin and Lai (2004) improved security of Lee et al. (2002) scheme by adding change password phase, In this phase user able to change his/her password whenever he/she desire to. Before changing password user need to successfully complete the log-in phase. Again there is no experimental work shown in Lin-Lai (2004) paper which proved that Lee et al. (2002) method is helpless to stop masquerade attack. owever Lee et al. (2002) and Lin-Lai (2004) Methods are strong than Yoon et al. (2004) scheme because it consist of fingerprint biometric technology even if the criminal achieve user's smart card, it is still not possible for attacker to extract the user information because server hold the biometric template of user fingers which used to verify the identity of real user and it is measured that no two persons fingerprints will ever be same. In addition both Lee et al. (2002) and Lin-Lai (2004) schemes carried out the login request by timestamp from system. If server found alter in request timestamp it stops log in request and provide failure massage to user screen otherwise server allow user to log-in safely.

Afterward Ku et al. (2005) pointed out some setbacks of Lin-Lai (2004) scheme by exposing a fake attack, Ku et al. Study show that attacker can copy any legal user by performing this forgery attack plus they add that Lin-Lai (2004) scheme is very hard to fix. Khan and Zhang (2007) also capture some errors in Lin-Lai (2004) scheme they demonstrate that Lin-Lai (2004) scheme can be easily attacked. They

state that Lin-Lai (2004) method is susceptible to server spoofing attack because their scheme has not any feature to authenticate remote server, therefore Lin-Lai (2004) scheme also vulnerable to masquerade attack.

To overcome this problem Khan-Zhang (2007) suggested better confidence bit, which carries out combined authentication between server and user and hold out server spoofing attack. Khan-Zhang (2007) demonstrates the problem of Lin-Lai (2004) scheme by supposing adversary named as "Bob" who eavesdropping over the unsafe channel between remote server and user. When consumer sends the login request to server, bob also catch it and bob can able to spoof user with this information by Mimicking the server. So Bob here plays a task of bogus server and transmits a fake combined authentication message to remote user. Whereas Khan-Zhang (2007) proposed method verify the mutual authentication message before relaying on server. In Khan-Zhang (2007) scheme user verify the validity of timestamp acquire by server and if user found it forged operation will terminate and chances of user data spoof is zero. Till now there are many user authentication schemes are proposed but not all fits everywhere, they missed some bits to cover which attacker can used as holes to perform attacks on communication between remote user and server.

Li-Hwang (2010) projected a different remote client authentication structure using biometric technology. Li-Hwang (2010) did their best up to many extant to cover up the weakness of previous proposed schemes. By comparing their scheme with other works, it reached at more effective and improved scheme. Lin-Lai (2004) and Lee-Chius (2005) scheme required more computational the reason is that their method is based on natural logarithms and result for using their scheme is spending more money and wasting time to manage the system. Whereas schemes of Li, H (2010); Chang et al. (2006); Yoon et al. (2005); Khan et al. (2008) are very cost effective in compare to Lin-Lai (2004) and Lee-Chiu (2005) schemes because there schemes are based on just few hashing functions and no need to spend too much money on computational to manage it. Chang et al. (2006) scheme allow user to select password of own choice while registration part but after that when user want to change password they need to notify the server which increase the communication between remote user and sever and risk of possible attacks (such as masquerade attack and server spoofing attack) over insecure network are rise.

In contrast Li-Hwang (2010); Chang et al. (2006); Yoon et al (2005); Khan et al. (2008) provide the joint authentication among server and user communication. Nevertheless Yoon et al. (2005) method not provide non-repudiation as Li-Hwang (2010) and Khan et al. (2008) scheme offer because their scheme is based on biometric technology. Li-Hwang (2010) claimed their scheme provided more reliability, efficiency and security because their scheme hold great properties in compare to Khan et al. (2008) and Yoon et al. (2005) schemes the reason they showed that they don't use synchronised clock as them. In addition they state disadvantages of synchronised clock first it is very intricate to get time period concurrency and second it consist a few problems which occurs like supply latency and dissimilar time-zone of remote user with remote server.

3 Conclusions

At the end of this paper it's been found that biometric based user authentication schemes are more secure than traditional password based schemes. Every User is looking for three main things to rely on service provider i.e. safety measures which protect their network, good organization system which fulfil their needs securely and dependability which keep their own information/data confidentially on secure network. Study discuss in this paper shows that Li-Hwang (2010) remote abuser authentication method which is based on biometric arrangement enclose at least all features which are required to sheltered remote user and server from online frauds. Their scheme is solid enough to withstand attacks such as masquerade attack and server spoofing attack as they discuss in paper. An extra thing, their scheme is based on biometric technology which is very strong to limit the access of real user to their private data. As it is mentioned above there is no possibility of equality in biometrics of two individuals therefore it's not possible for other than user to access his/her private data. But experimental work should to

done before launching any method so it proved that this is the right thing to protect user from online criminals and In this case Li-Hwang (2010) also not discussed any experimental work to back up their method theory. So there is need to experiment methods like user authentication to prove people it is the right for them to safe and securely using online computing.

Further study which been found from this paper is to improve synchronised clock which been used by some of user authentication schemes discussed above. The problems inspecting i.e. different time zone and delivery latency need to fix in synchronised clock to protect user and server communication.

References

Chun-TaLi a, Min-ShiangHwang, (2010), "An efficient biometrics-based remote user authentication scheme using smart cards", *Journal of Network and Computer Applications*, Volume 33, Pages 1–5.

H. M. Sun, (2000) "An efficient remote user authentication scheme using smartcards," *IEEE Trans. Consumer Electronic*, Volume 46, Issue 4, Pages 958-961.

H. Y. Chien, J.K. Jan and Y. M. Tseng, (2002), "An efficient and practical solution to remote authentication: smart card," *Computer & Security*, Volume 21, Issue 4, Pages 372-375.

J. J. Shen, C. W. Lin and M. S. Hwang, (2003), "A modified remote user authentication scheme using smart cards", *IEEE Trans. Consumer Electronic*, Volume 49, Issue 2, Pages 414-416.

Jan J-K, Chen Y-Y, (1998), "Paramita wisdom password authentication scheme without verification tables", *The Journal of Systems and Software*, Volume 42, Issue 1, Pages 45–57.

Hsieh B-T, Yeh H-Y, Sun H-M, Lin C-T, (2003), "Cryptanalysis of a fingerprint-based remote user authentication scheme using smart cards", *Proceedings of 37th IEEE conference on security technology*, Pages 349-350.

Lee J-K, Ryu S-R, Yoo K-Y, (2002), "Fingerprint-based remote user authentication scheme using smart cards", *Electronic letters*, Volume 38, Issue 12, Pages 554-555.

Lin C-H, Lai Y-Y, (2004), " A flexible biometrics remote user authentication scheme", *Computer standards and interfaces*, Volume 27, Issue 1, Pages 19-23.

Chang Y-F, Chang C-C, Su Y-W, (2006). "A secure improvement on the user-friendly remote authentication scheme with no time concurrency mechanism", *Proceedings of 20th international conference on advanced information networking and applications, IEEE CS.*

Khan MK, Zhang J, (2007), "Improving the security of a flexible biometrics remote user authentication scheme", *Computer Standards and Interfaces*, Volume 29, Issue 1, Pages 82–85.

Khan MK, Zhang J, Wang X, (2008), "Chaotic hash-based fingerprint biometric remote user authentication scheme on mobile devices", *Chaos, Solutions and Fractals*, Volume 35, Issue 3, Pages 519–524.

Yoon E-J, Ryu E-K, Yoo K-Y, (2005), "An improvement of Hwang–Lee–Tang's simple remote user authentication scheme", *Computers and Security*, Volume 24, Issue 1, Pages 50-56.

An Evaluation of Biometric Security Methods for Use on Mobile Devices

Joe van de Bilt

Abstract

Mobile devices currently employ number or pattern based security methods to 'lock' a device, preventing it from being used by unwanted parties, however these systems are not flawless and can be susceptible to a number of attacks, this paper looks into methods of biometric security currently available and if these methods can be implemented onto a mobile device and if so, how well the mobile device will function and how well it would employ the biometric security method. Looking in depth into new areas of research and as well as new takes upon old methods, we establish if a mobile device can employ biometric security methods

1. Introduction

Mobile devices contain a lot of sensitive user information and cryptography methods are already used to keep that information secure. However if the mobile device were to physically fall into the wrong hands there is little in the way of securing data on a local scale, mobile devices can employ locking technology, such as security patterns, passwords and pin codes, but these methods are old and easy to break through using brute force methods and also carry the potential to be forgotten (Al-Assam et al. 2012)

In this paper we look into research currently being undertaken into using biometric security methods on mobile devices, in particular we will be looking into methods such as using facial recognition (with an onboard camera) or fingerprint detection.

We will draw upon a number of research papers already written and critically evaluate the conclusions reached in those papers, thinking about the evidence and scientific methods used, then correlating those conclusions in order to reach a conclusion of our own based on the knowledge gained.

2. Biometric Security Methods

Biometric security has been around for long enough to judge what is possible and what is not given the current technology. We have analyzed a variety of research papers and highlighted key topics that are relevant and show promise in the field of biometric security.

2.1 Keystroke Dynamic Authentication

Keystroke Dynamic Authentication (KDA) is a method of combining physical interaction with a device and a physical password. A device currently uses a passcode or pattern to determine legitimacy of a user, KDA assesses a user's physical behavior when entering that password to enhance security meaning "even if the password is revealed by dictionary attacks or shoulder surfing attacks, the probability of breaking authentication is reduced." (Ting-Yi et al. 2011)

The system employs a training period, where the device 'learns' the behavior of the user, taking into account the pauses they take between keystrokes and the pressure they apply to the device, as to build a profile from which they can evaluate (once the training period is over) if a user is genuine or an imposter.

However, Xuan et al. (2012) raises the issue that a stable profile cannot be maintained over long periods of time, claiming that determining a user as an imposter is not such a simple procedure, and that research in the past has only yielded positive results due to flaws in their experimentation.

The system has been in development and refinement for decades and has only yielded a small amount of positive gains, as stated by Ting-Yi et al. (2011) the chances of breaking authentication is reduced, but the system still faces challenges in terms of the degradation in the biometric profile over time.

Evaluation

The methods used to confirm the effectiveness of KDA in experiments carried out by Ting-Yi et al. (2010) do indeed show that a biometric profile can be made and implemented on a mobile device to prevent imposters from gaining access, the profile is built from a user's own personal habits and as such its effectiveness at security is not diminished by the knowledge of the user, for instance somebody not technically literate may not use a very complex password but a biometric profile is secure no matter what the knowledge of the user. The experiment does show that an imposter can be detected using this method, but as suggested by Xuan et al (2012) the imposters were already on file, and a binary decision would have been easier to reach by the program given that information.

2.2 Fingerprint Recognition

Fingerprint recognition is a key area of research in the field of biometric security, as it is one of the most fundamental uniqueness' that being a human can carry (no two people will have the same fingerprints) Methods already exist to detect finger prints, typically using a CCD, the details of the finger are shown as they reflect the light emitted from within the device and the image is then analyzed and matched against a database containing authorized users.

Hiew et al. (2009) determined in experiments carried out in their research paper that their proposed 'touch-less' method of fingerprint scanning is possible, with the use of a digital camera. The research provided positive test results, however their experiments were carried out in locations that had a constant and consistent source of light, their method was never conducted in different environments, and so there is not enough evidence to conclude this method as completely effective, however with more research it may yield positive results. Their experiment used a Canon The PowerShot Pro1 digital camera with an 8megapixel output and used standard RGB, features that can now be found in some mobile devices, however pictures were always taken of the thumb and were taken between 4-6cm away from the lens of the camera (see figure 1 for diagram).

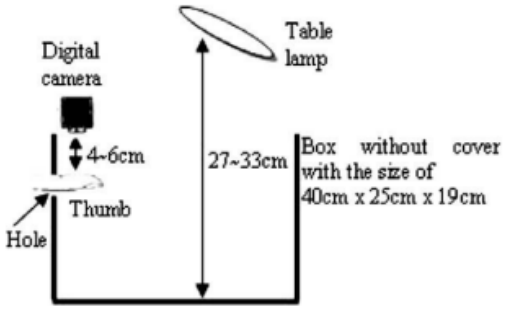


Figure 1 diagram of experiment apparatus (Hiew et al. 2009)

Evaluation

In these experiments it was shown that the scanning of a fingerprint is a valid way of determining a person's true identity and is possible on a mobile device. The equipment used in the experiments done by Hiew et al. (2009) shows us that it is possible with today's modern technology, better in fact considering technology has advanced for 3 years further now. However due to the method of the experiment we cannot say for certain if the method could be used on a mobile device. In the diagram we see that everything is kept stationary with a consistent amount of light and the conditions remain the same in every instance, this is not possible in real world applications.

2.3 Iris Recognition

As with fingerprint scanning technology, the human iris is unique in each and every human being, and as such is also a prime method of determining the true identity of somebody.

Kang (2010) tells us that we need a high resolution image of an iris, as well as a focused one in order to properly assess the details of the iris, however even on low quality images algorithms can be used to detect parts of the eye, including 'noise' around the eye (eyelashes, eyelids and facial features) and can be trained to disregard this noise in order to know where to focus. He claims his research shows

encouraging results in terms of mobile iris detection.

However as we know from normal iris recognition devices "There should also be a certain point that each person should look at so that the device can obtain a consistent reading from the same angle each time" (Brown 2012) This will prove to be difficult if using a digital camera as the angle at which a device is held cannot be consistent as in most cases, the device taking the picture is stationary.

Evaluation

Iris detection showed more promise in mobile devices, experiments carried out by Kang (2010) yielded a high volume of positive results, despite issues raised that images need to be high quality, the system designed worked well taking a picture of an eye, clearing up the image and analyzing it properly, and his test subjects were varied and tests were carried out in different environments in and out of the lab. However the system was only used to detect people already in the database, it was not used to detect and intruder and as such I cannot conclude the system works, as not only does this system need to identify authenticated users but also deny imposters.

2.4 Hand Based Biometrics

Hand based biometrics focus on hand gestures, they are currently used in practice as a means of verifying a person's identity. They gather a range of information on a person's hand, generally fingerprints, knuckles, palm, contours and even veins, all to generate a profile of a person's hand (usually only one hand is used)

The hand is a viable way of identifying a person's true identity because "Once a person has reached adulthood, the hand structure and configuration remain relatively stable throughout the person's life" (Goh Kah Ong et al. 2012)

Most hand recognition devices use touch based reading devices, but in the research carried out by Goh Kah Ong et al. (2012) they proposed using simple off the shelf webcams to identify hands without the need for touch screens, which as they point out carry drawbacks, including hygienic reasons. The webcams take an image of the hand and analyze it looking at all the features available, as it only looks at the inner hand it focuses' on reading the palm, knuckle indents, fingertips and contours/gaps between fingers, which is referred to as hand geometry. To ensure that the hand was always visible a series of normal lights was used to illuminate the hand, the capture device is shown in figure 2.

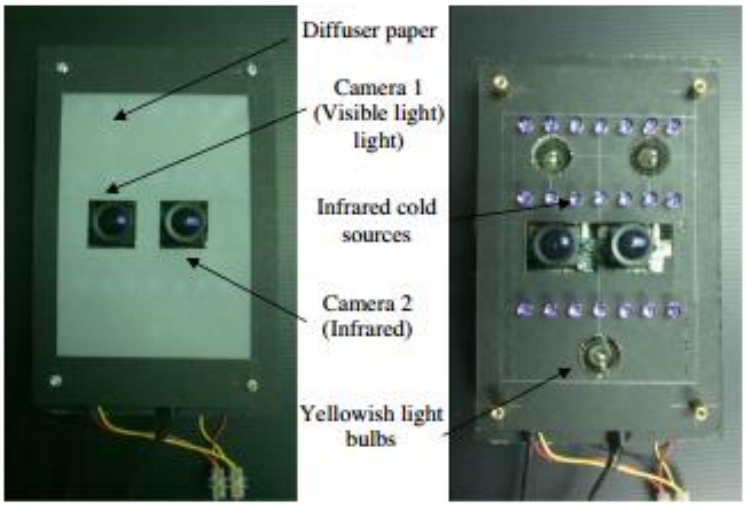


Figure 2 Capture device used in experiment (Goh Kah Ong et al. 2012)

Evaluation

Goh Kah Ong et al. (2012) brought a new look onto an old technique. In the research by Adan et al. (2008) it is indicated that hand verification technology is cumbersome, however it would appear over the last 4 years the technology has evolved and is now able to be performed by such simple off the shelf items. The software itself is more complex and required a lot more processing power, a quad core processor was used and still verification took an average of 2 seconds.

The experiment itself was conducted in the laboratory and consisted of 10 test candidates from different ethnicities and genders, this means the range of test candidates varies and the results collected cover success rates and speed of the system. The only potential flaw with the system is that it has not been used to identify imposters, and has only been used to test if such a system is possible, though the high success rate is achieved; it is not mentioned if the system will reject imposters as well as identify authorized users.

3 Conclusions

KDA was developed *for* mobile devices and as such would be extremely easy to implement into practice on existing devices, maintain it however would prove to be difficult as its reliability would degrade over time. For it to become a viable method of securing a mobile device it may be an option to repeat the training period every few months to ensure that the biometric profile generated is up to date and does not suffer degradation.

Fingerprint detection proves to be possible on mobile devices by using onboard capture devices to scan a fingerprint, however the camera must always be focused and light must be sufficient, which may prove to be a hindrance, the picture of the finger must also be taken with a suitable distance from the lens and at a suitable angle, which may be difficult when the device is being held by a human arm, it would also prove difficult keeping the capture device steady while trying to capture an image.

Similar problems would also occur in iris recognition, though the quality of the image does not have to be as precise and the image can be 'cleaned' to determine the parts of the eye easier, it still requires the device to be extremely still, and well lit, things that cannot always remain constant when using mobile devices, also from the experiments carried out by Kang (2010) The equipment used required an extra lens be but onto the device, which should be considered, as if this were to be implemented onto a smart phone, say for instance an iPhone that focuses on sleek design, it may be a poor selling point to have bulky extra equipment built into it.

Hand based biometrics are suitable for mobile devices because "of their low processing time and real time response" (Adan et al. 2008) however the equipment needed to recognize a hand gesture cannot be easily combined with a mobile device, it also in some instances requires small receivers to be fitted to the hand. The experiments carried out by Goh Kah Ong et al. (2012) showed us however that the system can function on less complicated technology, however this resulted in a dramatic increase in processing power and response time, which when using a mobile device is relied upon quit often. As such I believe this system is not a practical solution.

Most of these methods rely heavily on proper lighting stability and high focus images (does not apply to the KDA method) mobile devices are often used while 'on the go' as it is their main purpose as such it is necessary for them to allow quick access. Imagine if you receive an email and you wish to read and respond to it, but before you can begin you must take a picture of your fingerprint, or iris or indeed your whole hand. First the image must be steady, which may be difficult to achieve if you are walking or standing, the image may not always autofocus correctly and even so, it is not uncommon for mobile devices to be used at night or in the dark, if this is the case taking a picture may become difficult, a mobile camera can have a flash function on it, but it would be difficult for the image to focus properly when there is not ample lighting beforehand.

Even if all those problems are resolved the research does not show conclusively that imposters would be properly denied access to

the mobile device, only in the instance of KDA was biometrics fused with existing mobile security and was successful in identifying and stopping intruders, but as was shown, the biometric profile degrades over time, and there is a fear that the security may lock out authentic users, and so the system requires refinement.

Ultimately I find that a mobile device cannot rely solely on biometric security methods in order to safeguard the device, it can only be used to enhance existing security methods and should not be relied on wholeheartedly. Biometrics do hold great potential to being the key of allowing computers to identify individuals and with more refinement over the next few years we may begin to see biometric security starting to weave itself into our existing technology.

References

Adan, M., Adan, A., Vazquez, A., Torres, R., (2008). 'Biometric verification/identification based on hands natural layout'. *Image and Vision Computing, Vol. 26, Issue 4, Pages 451-465*

Al-Assam. H., Jassim, S., (2012). 'Security evaluation of biometric keys'. *Computers & Security, Vol. 31, Issue 2, Pages 151-163.*

Brown, D., (2012). 'A Detailed Analysis of Current Biometric Research Aimed at Improving Online Authentication Systems'. *University of Sunderland, Selected Computing Research Papers, Vol. 1, Pages 7 – 11*

Goh Kah Ong, M., Tee, C., Andrew, T., (2012). 'A contactless biometric system using multiple hand features'. *Journal of Visual Communication and Image Representation, Volume 23, Issue 7, Pages 1068-1084*

Hiew, B Y., Teoh, A., Yin, O., (2009). 'A secure digital camera based fingerprint verification system'. *Journal of Visual Communication and Image Representation, Vol. 21, Issue 3, Pages 219-231*

Kang, J., (2010). 'Mobile iris recognition systems: An emerging biometric technology'. *Procedia Computer Science, Volume 1, Issue 1, Pages 475-484*

Ting-Yi, C., Cheng-Juang, T., Jyun-Hao, L., (2011). 'A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices'. *Journal of Systems and Software, Vol. 85, Issue 5, Pages 1157-1165.*

Xuan, W., Fangxia, G., Jian-feng, M., (2012). 'User authentication via keystroke dynamics based on difference subspace and slope correlation degree'. *Digital Signal Processing, Vol. 22, Issue 5, Pages 707-712*