IJIREEICE

ISSN (Online) 2321–2004
ISSN (Print) 2321–5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 4, April 2018

Voice Based Surveillance System For Phishing Detection

Aneesh Gowtham.P[1], Balaji.H[2], Benin Moses.E[3], Harish Babu.B[4], Dr.V.Sujatha[5], Dr.M.Subba[6]

UG Scholar, Department of Instrumentation and Control Engineering,
Sri Manakula Vinayagar Engineering College, Puducherry[1,2,3,4]

Associate Professor, Department of Instrumentation and Control Engineering,
Sri Manakula Vinayagar Engineering College, Puducherry[5]

Professor & Head, Department of Instrumentation and Control Engineering,
Sri Manakula Vinayagar Engineering College, Puducherry[6]

Abstract: Phishing is the fraudulent electronic communications (emails, text messages, or instant messages) that appear to come from legitimate businesses to entice you to provide personal information. The fraudster then uses your information to commit identity theft or lure you to further engage in scams. Voice based surveillance system keeps track of certain voice based keywords like PIN Number, ATM card number, Password, Credit card number, lucky draw winner etc., in the conversation occurring between the potential attacker and the victim. When these keywords are detected, the system sends an alert message to the user's mobile phone with a vibration for 2 to 3 seconds and switches to the surveillance mode and starts tracking various details about the potential attacker like, the origin/location of the attacker's call, phone number, service provider details and stores them in a temporary memory. If any transaction happens with the user's bank account within a predefined threshold time after the phishing conversation has been detected, then the system sends an alert message to both the customer and bank about the transactions. The alert message comprises of various details regarding the bank transaction IP address of the machine from where the transaction was carried out, transaction ID, time stamp of transaction, phone numbers details (service provider names, phone user names and geographic location) of both the user and the potential attacker.

Keywords: Potential attacker, victim, surveillance mode, alert message, transaction.

I. INTRODUCTION

Phishing is a type of social engineering attack often used to steal user data including login credentials, credit card numbers and passwords without any actual hacking tool. Phishing is popular with cyber criminals, as it is far easier to track someone into clicking a malicious link in a seemingly legitimate phishing email than trying to break through a computer's defenses. Phishing attacks typically rely on social networking techniques applied to email or other electronic communication methods, including direct messages sent over social networks, SMS text messages and other instant messaging modes. Voice phishing, also known as vishing, is a form of phishing that occurs over voice communications media, including Voice Over IP (VoIP) or Plain Old Telephone Service (POTS). A typical vishing scam uses speech synthesis software to leave voicemails purporting to notify the victim of suspicious activity in a bank or credit account, and solicits the victim to respond to a malicious phone number to verify his identity thus compromising the victim's account credentials. Messages that claimed to be from bank users to dial a phone number regarding problems with their bank accounts. Once the phone number (owned by the phisher, and provided by a Voice Over IP service) was dialed, prompts told users to enter their account numbers and PIN. Vishing (voice phishing) sometimes uses fake caller-ID data to give the appearance that calls come from a trusted organization. SMS phishing uses cell phone text messages to induce people to divulge their personal information.

Ankit Kumar Jain et al., [1], 2016, proposed a comprehensive analysis of phishing attacks, their exploitation, some of the recent visual similarity based approaches for phishing detection, and its comparative study. Visual similarity based phishing detection techniques utilize the feature set like text content, text format, HTML tags, Cascading Style Sheet (CSS), image, and so forth, to make the decision. These approaches compare the suspicious website with the corresponding legitimate website by using various features and if the similarity is greater than the predefined threshold value then it is declared phishing.M.Moghimi et al., [2], 2016, proposed a new rule-based method to detect phishing attacks in internet banking. Our rule-based method used two novel feature sets, which have been proposed to determine the webpage identity. Our proposed feature sets include four features to evaluate the page resources identity, and four features to identify the access protocol of page resource elements. We used approximate string matching algorithms to determine the relationship between the content and the URL of a page in our first proposed feature set. Our proposed

DOI 10.17148/IJIREEICE.2018.6417

features are independent from third-party services such as search engines result and/or web browser history. We employed support vector machine (SVM) algorithm to classify web pages. Our experiments indicate that the proposed model can detect phishing pages in internet banking with accuracy of 99.14% true positive and only 0.86% false negative alarm. Output of sensitivity analysis demonstrates the significant impact of our proposed features over traditional features. We extracted the hidden knowledge from the proposed SVM model by adopting a related method. We embedded the extracted rules into a browser extension named Phish Detector to make our proposed method more functional and easy to use. Evaluating of the implemented browser extension indicates that it can detect phishing attacks in internet banking with high accuracy and reliability. Phishdetector can detect zero-day phishing attacks too.

C.Clavel et al., [3], 2005, proposed The present research deals with audio events detection in noisy environments for a multimedia surveillance application. In surveillance or homeland security most of the systems aiming to automatically detect abnormal situations are only based on visual clues while, in some situations, it may be easier to detect a given event using the audio information. This is in particular the case for the class of sounds considered in this paper, sounds produced by gun shots. The automatic shot detection system presented is based on a novelty detection approach which offers a solution to detect abnormality (abnormal audio events) in continuous audio recordings of public places. J.-H.Chang et al., [4], 2009, proposed an effective voice phishing detection algorithm based on a Gaussian mixture model (GMM) employing the minimum classification error (MCE) technique. The detection of voice phishing is performed based on the GMM using decoding parameters of the 3GPP2 selectable mode vocoder (SMV) 1codec directly extracted from the decoding process of the transmitted speech information in the mobile phone. The authors’ approach is further improved by the MCE scheme in that different weights are assigned to each likelihood ratio and is considered to be new. The experimental results show that the proposed method is effective in discriminating between true statements andlies.

G.Ramesh et al., [5], 2014, proposed an efficacious method for detecting phishing web pages through target domain identification. They present a novel approach that not only overcomes many of the difficulties in detecting phishing websites but also identifies the phishing target that is being mimicked. We have proposed an anti-phishing technique that groups the domains from hyperlinks having direct or indirect association with the given suspicious webpage. The domains gathered from the directly associated web pages are compared with the domains gathered from the indirectly associated web pages to arrive at a target domain set. On applying Target Identification (TID) algorithm on this set, target domain is zero. We then perform third-party DNS lookup of the suspicious domain and the target domain and on comparison we identify the legitimacy of the suspicious page. G.A.Montazer et al., [6], 2015, their aim is to provide a phishing detection system to be used in e-banking system in Iran. Identifying the outstanding features of phishing is one of the important prerequisites in design of an accurate system, therefore, in first step, to identify the influential features of phishing that best fit the Iranian bank sites, a list of 28 phishing indicators was prepared. Using feature selection algorithm based on rough sets theory, six main indicators were identified as the most effective factors. The fuzzy expert system was designed using these indicators, afterwards. The results show that the proposed system is able to determine the Iranian phishing sites with a reasonable speed and precision, having an accuracy of88%.

This paper is organized as follows: Section 1 discusses introduction, literature survey and organization of the paper. Section 2 discusses the methodology of the proposed system. Section 3 discusses about the result and discussions and followed by conclusion and future scope in section4.

II. METHODOLOGY

The voice from the caller during the phone call conversation is converted to word text using an android app named Bluetalk which is readily available for any user. With the help of this android app, the phone call conversation is continuously checked with certain keywords such as ATM card number, Debit card number, PIN number, etc. which is already fused with the Arduino nano. If such keywords are detected, the arduino nano switches the mobile vibrator ON for three seconds, which will intend the mobile users to look on to their phones during which an alert message is displayed from the GSM module. This alert message can prevent the mobile users from further carried away by the conversation, exposing all their credentials. If any further transaction happens between the potential attacker and the victim, then the attackers location from where the actual conversation took place and the transaction ID information are sent to the victim for further finding the phisher easily. The flowchart of the proposed system is shown inFig.1.

DOI 10.17148/IJIREEICE.2018.6417

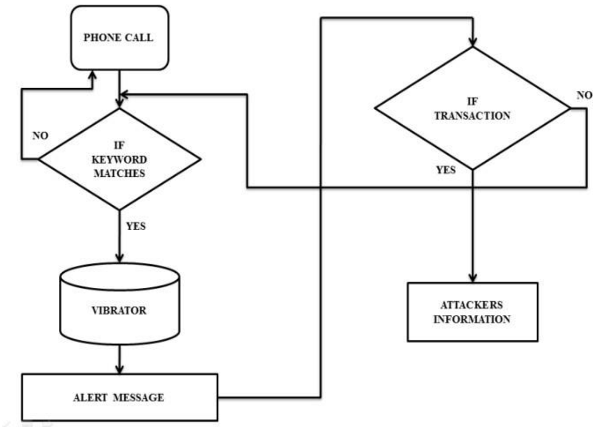


Fig 1.Flowchart of the proposed system

The interconnection between different components is explained using the architecture of system. The architecture diagram is shown in the Fig 2. The phone call conversation is monitored continuously by connecting an android application called Blue Talk, which recognizes the voice from the caller, to a text format which is fed to the arduino nano to compare each and every word with the prescribed keyword which will be already fused to the arduino nano by using arduino IDE (Integrated Development Environment). A vibrator is connected with the digital pin 2 of arduino nano and the rest is grounded. The SIM800L GSM module is connected with the transmitter and receiver pin of arduino nano with the transmitter and receiver pin of the GSM module. If the voice of the attacker matches the keyword, the vibrator vibrates for 3 to 4 seconds and the GSM module sends an alerting message to the user. If further transaction happens between the potential attacker and the victim then the GSM module will send the attackers information about the exact location during where the phone call occurred, to the victim.

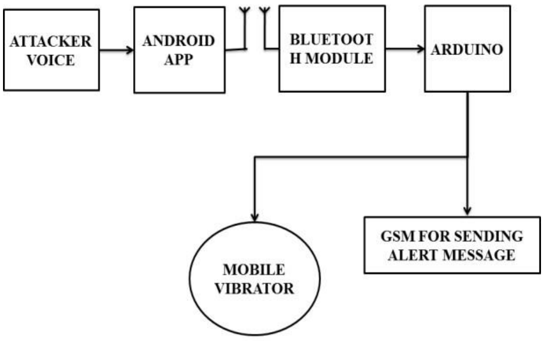


Fig 2.Architecture diagram

The design of the system is divided into two parts: Hardware components and software components.

A. Hardware components

1) *Bluetooth module (HC-05):*HC-05 module shown in Fig.3 is an easy to use Bluetooth SPP (Serial Port Protocol) module, designed for transparent wireless serial connection setup. Serial port Bluetooth module is fully qualified Bluetooth V2.0+EDR (Enhanced Data Rate) 3Mbps Modulation with complete 2.4GHz radio transceiver and baseband. HC 05/06 works on serial communication. The Arduino Bluetooth module at the other end receives the data and sends it to the Arduino through the TX pin of the Bluetooth module (connected to RX pin ofArduino).

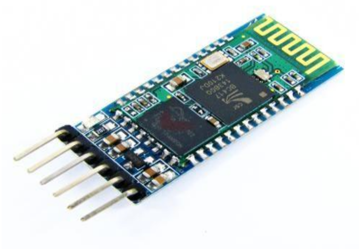


Fig 3.Bluetooth module (HC-05)

DOI 10.17148/IJIREEICE.2018.6417

2) *SIM800L GSM/GPRS module:* GSM (Global System for Mobile communication)shown in Fig.4 is a digital mobile telephony system that is widely used in Europe and other parts of the world. GSM uses a variation of time division multiple access (TDMA) and is the most widely used of the three digital wireless telephony technologies (TDMA, GSM, and CDMA). A GSM modem requires a SIM card to be operated and operates over a network range subscribed by the network operator. It can be connected to a computer through serial, USB or Bluetooth connection. GSM modem is usually preferable to a GSM mobile phone.

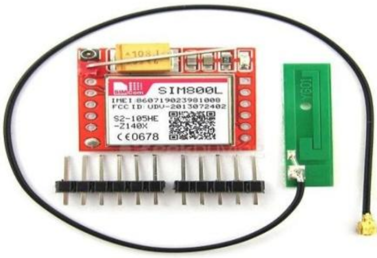


Fig 4.SIM800L GSM module

3) *Arduino nano:* Arduino nano shown in Fig.5 consists of both a physical programmable circuit board (often referred to as a microcontroller) and a piece of software, or IDE (Integrated Development Environment) that runs on your computer, used to write and upload computer code to the physical board.Itis a small, complete, and breadboard-friendly board based on the ATmega328P (Arduino Nano). It has more or less the same functionality of the Arduino UNO, but in a different package. It lacks only a DC power jack, and works with a Mini-B USB cable instead of a standard one.

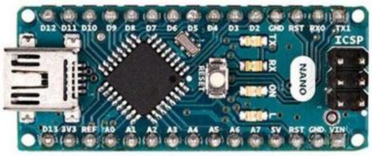


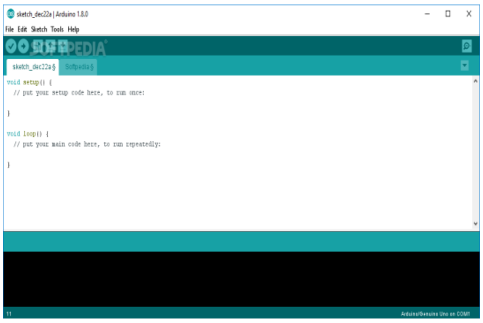
Fig 5.Arduino nano

4) *Mobile vibrator:* Vibration motor shown in Fig.6 is a compact size coreless DC motor used to informs the users of receiving the signal by vibrating, no sound. Vibration motors are widely used in a variety of applications including cell phones, handsets, pagers, and so on.

Fig 6.Mobile vibrator

B. Software components

1) *Arduino IDE (Integrated Development Environment):* Arduino consists of both a physical programmable circuit board (often referred to as a microcontroller) and a piece of software, or IDE (Integrated Development Environment)shown in Fig.7 that runs on your computer, used to write and upload computer code to the physical board. In fact, you already are; the Arduino language is merely a set of C/C++ functions that can be called from your code. Your sketch undergoes minor changes (e.g. automatic generation of function prototypes) and then is passed directly to a C/C++ compiler (avr-g++).



sketch_dec22a | Arduino 1.8.0
File Edit Sketch Tools Help
sketch_dec22a §
Softpedia §
11
Arduino/Genuino Uno on COM1

```
void setup() {
  // put your setup code here, to run once:

}

void loop() {
  // put your main code here, to run repeatedly:

}
```

Fig 7.Arduino IDE window

2) *Bluetalk android app:* It is used to control your Arduino, via Bluetooth. The App works by pressing the mic button, then the it will wait for you to say a command. The app will then display the word's that you've stated and will send data strings for the Arduino to process.

III. RESULT&DISCUSSION

Fig.8 shows enhanced miniaturized model alerts the user through the vibration of the vibrator for the user to look at their phone when the alert message pops up. Fig.9 shows the alert message that is sent to the user. Problem definition of our underlying system which is basically useful for monitoring phishing attacks over phone call conversation and to alert the user from further being carried away by the phisher by exposing their important credentials.

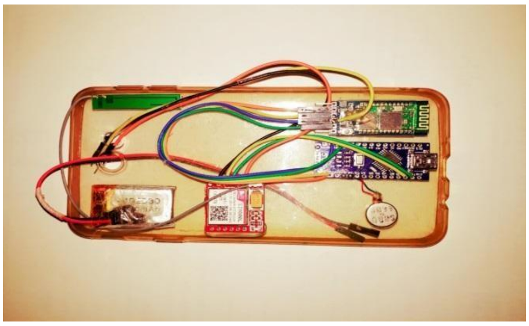
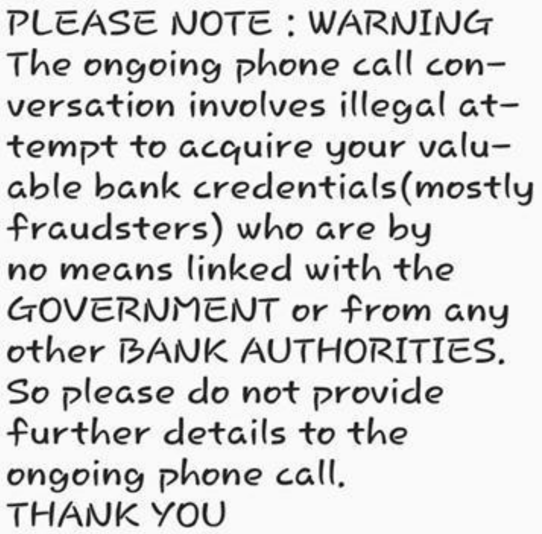


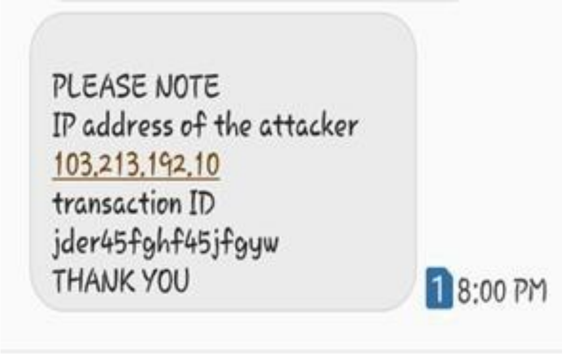
Fig.8. Proposed model



PLEASE NOTE : WARNING
The ongoing phone call conversation involves illegal attempt to acquire your valuable bank credentials(mostly fraudsters) who are by no means linked with the GOVERNMENT or from any other BANK AUTHORITIES.
So please do not provide further details to the ongoing phone call.
THANK YOU

Fig.9. Alert message to the user

Fig.10 shows the message of the attacker's information, this only occurs only when the transaction occurs.



PLEASE NOTE
IP address of the attacker
103.213.192.10
transaction ID
jder45fghf45jfgyw
THANK YOU

Fig.10. Attackers information

IV.CONCLUSION & FUTURE SCOPE

A cost effective voice based surveillance system is was proposed, designed, miniaturized and successfully implemented in this project. Along with the phishing detection, the setup also provides a more efficient way of segregating the conversation happening between mobile phones. The LED's provides the types segregation in a more vivid way. This setup when implemented in smart phones during manufacturing can save many people from further being ensnared by the phishers. This entire setup can be further reduced by fusing this program to the mobile processor during manufacturing. Since mobile phones have built-in vibrator, processor, GSM message sending feature, there will be no need of external hardware required for implementing this system to any smart phones during manufacturing.

REFERENCES

[1] Ankit Kumar Jain and B. B. Gupta, "Phishing Detection: Analysis of Visual Similarity Based Approaches", NationalInstitute of Technology, Kurukshetra, India, 10 January 2017.

[2] M. Moghimi and A. Y. Varjani, "New rule-based phishing detection method," Expert Systems with Applications, vol. 53, pp. 231–242,2016.

[3] C. Clavel, T. Ehrette, G. Richard, "Events Detection For An Audio-Based Surveillance System", Thales Research and Technology France, Domaine de Corbeville, 91404 OrsayCedex, France GET-ENST, 46 rue Barrault, 75634 Paris Cedex 13, France,2005

[4] T. Moore and R. Clayton, "Examining the impact of website take-down on phishing," in Proceedings of the Anti-Phishing Working Groups 2nd Annual Ecrime Researchers Summit, pp. 1–13, Pittsburgh, Pa, USA, October2007.

[5] A. Almomani, B. B. Gupta, S. Atawneh, A. Meulenberg, and E. Almomani, "A survey of phishing email filtering techniques," IEEE Communications Surveys and Tutorials, vol. 15, no. 4, pp. 2070–2090,2013

[6] Ramanathan and H. Wechsler, "PhishGILLNET- phishing detection methodology using probabilistic latent semantic analysis, AdaBoost, and co-training," Eurasip Journal on Information Security, vol. 2012, article 1,2012.

[7] Y. Zhang, J. I. Hong, and L. F. Cranor, "Cantina: a content-based approach to detecting phishing web sites," in Proceedings of the 16th International World Wide Web Conference (WWW '07), pp. 639–648, Banff, Canda, May 2007.

[8] K. L. Chiew, E. H. Chang, S. N. Sze, and W. K. Tiong, "Utilisation of website logo for phishing detection," Computers & Security, vol. 54, pp. 16–26,2015.

DOI 10.17148/IJIREEICE.2018.6417