Distributed Software agents for antiphishing

Sarika S[1] ,Dr.Varghese Paul[2]

[1] Computer Science and Engineering
, Sree Narayana Gurukulam College of Engineering
, Kolenchery, Kerala,India

[2] Information Technology , Cochin University
Thrikkakara,Cochin,Kerala,India

Abstract

Phishing attacks are one of the most popular problems of online security. The impact of phishing is quite dramatic when it involves the threat of identity theft and financial losses. There are a number of anti-phishing solutions proposed so far but those methods are inefficient for the new type of phishing attacks as the phishers use images and moving objects to impersonate a webpage. This paper proposes a new algorithm for detecting new type of phishing attacks mainly focusing on tabnabbing attack using distributed software agents. Tabnabbing is a phishing attack which prompts the user to enter login details to well known websites by impersonating those websites when the webpage is idle for some time. The proposed method continuously monitors the reload of a webpage or the change in page layout using multiagent systems and prevents those attacks.

Keywords: *Phishing, distributed software agents, perceiving, phishing detection.*

1. Introduction

In recent years, online security has received critical attention from both academia and industry. As the data network becomes more pervasive and its scale becomes larger, network intrusion and attack have become severe threats to network users.Phishing attacks are one of the most popular problems of online security. It deceive the users to capture sensitive information. It is difficult to provide secure web services due to emergence of new threats. Expensive security mechanisms can lead to reduced effectiveness and may consume excessive network resources leading to opportunities for new type of attcks. There has been a greater focus on the subject of securing web services in the context of increasing interest in online security. Many research works and discussions has been conducted in this area. From these discussions and research works ,different security issues in the field of web services has been arisen and many papers have been written describing many antiphishing approaches that defend against phishing attacks that websites face. However, the majority of these approaches did not provide a complete solution for all the attacks .

Phishing occurs in varied forms like basic URL obfuscation, hyperlink based attacks and email based attacks. Now phishing has become a proliferate threat to internet privacy and data security. Many a research has been conducted in this area and a number of algorithms have been developed for protection against each type of attack. But unfortunately the phishers search new ways to deceive the users to obtain sensitive information. They implement new techniques to mimic well known websites and capture sensitive data from a legitimate user.

One of the new type of phishing attack is tabnabbing where the phisher tries to capture the login information of a user if the website is unattended for some time. When the website is idle for some time, the phisher places the malicious page over the original page. When the user returns to the page, he thinks he has logged out and types in the login information which is stolen straight away. Now the phisher redirects the page to the original website.

There are basically two types of Phishing detection .List-Based approach and Heuristic-Based approach[12]. List-based approaches are classified into blacklist based and whitelist based. In blacklist based method,the browser hold URLs that refer to websites that are considered phishing.The browser queries the blacklist to determine whether the currently visited URL is on this list and appropriate actions are taken if the URL is in blacklist. Heuristic-based approaches detect by checking one or more characteristics of a website. These characteristics can be the uniform resource locater (URL), the hypertext markup language (HTML) code, or the page content itself.

In existing methods, the layout changes are detected but it fails if the window is resized because all the webpages are not designed to re-layout themselves. But in this method,

layout changes and other reloads are captured using an image retrieval method. This method also performs URL verification using blacklists of phishing domains. This helps in detecting phishing attacks other than tabnabbing. The phishing domains which are not blacklisted are detected using layout changes or content changes. The hyperlinks in the webpage are also verified for content change. The algorithm starts to work when a webpage is unattended for some time and monitor for some layout changes or content changes. It then checks the URL of the webpage to determine whether it is blacklisted. Then the visual changes are monitored and store a particular value in an array according to the changes. If the stored value is greater than a certain predefined threshold value, an attack is detected and a proper message is displayed.

Agent technology has generated lots of excitement in recent years because of its effectiveness as a new paradigm for solving critical problems.This is particularly attractive for creating software that operates in environments that are distributed and open, such as the internet. Now, most of the agent-based systems consist of a single agent. As the technology advances and new problems arises, it is needed to implement systems that consist of multiple agents that communicate in a peer-to peer fashion. Distributed agent-based architectures offer a promising basis for practical solutions.This paper provides an idea about distributed agent-based architecture that resists different types of phishing attacks.

The rest of the paper is structured as follows. Section II describes about related work.Section III gives an overview of multiagent systems. Section IV discusses about a novel scheme for detecting phishing attacks using distributed software agents. Section V describes about implementation & ananlysis and section VI concludes the paper.

2. Related Work

AntiPhish[9] is a phishing protection system that prevents sensitive user information from being entered on phishing sites and generates warnings whenever the user enter this information to an unauthentic web site .This approach is not efficient because the system requires cooperation from the user and it may seldom generate false warnings.

Phishpin [8] is an Identity Based Anti Phishing Technique. In this, both user and online entity validates each other's identity during handshake. It is an anti-phishing technique that integrates partial credentials sharing and client filtering technique to prevent phishers from easily masquerading as legitimate online entities. There is no need for users to re-enter their credentials as mutual authentication is followed. This method provide mutual authentication for server as well as client side. This approach fails when an intruder is succeeded to gain access to the client computer and disable the browser plug-in.

In Genetic Algorithm Based Anti Phishing Techniques[10] phishing hyperlinks are detected using the rule based system formed by genetic algorithm. These rules are used to differentiate legitimate website from suspicious website. There is a ruleset generated by Genetic Algorithm that matches the suspicious links.This approach is effective to detect phishing links with minimal false negatives but is not much efficient because it need multiple rule set for only one type of URL based phishing detection and sometimes it may lead to complex algorithm.

In visual similarity based approach[2], the page features of a legitimate site and fake site is compared visually.It compare the page style,images and the text that is embedded in the page etc.Whenever a suspected phishing page is found, the suspected link is extracted And the corresponding legitimate page is retrieved .now a comparison of both page is performed and check for similarity.If not similar,an alert is raised.This method is having less false positives but it fails if the webpage contains moving objects.

LinkGuard[11] is a character based antiphishing approach.In this,the generic characteristics of the hyperlinks in phishing attacks are utilized. LinkGuard, extracts the DNS names from the actual and the visual links and then compares the actual and visual DNS names, if these names are not the same.If attackers use dotted decimal IP address in actual DNS, it is a possible phishing attack.This technique is not much efficient as it may result in false positives.

GoldPhish[4] and CANTINA[6] are a content based approaches. The advantage of GoldPhish is it does not result in false positive and provides zero day phishing.But it delays the rendering of a webpage. It is also vulnerable to attacks on Google's PageRank algorithm and Google's search service. CANTINA [4] is a content similarity based approach to detect phishing websites.It calculates the lexical signature of suspicious page using TF-IDF and feed it to a search engine .It can be determined as a phishing site by checking the sorting order of suspicious pages in the search results.

There is an approach to predict phishing websites using neural networks as explained in [5]. The advantage of this method is when an element of the neural networks fails, it can continue without any problem because of its parallel nature. But to fully implement a standard neural network

architecture would require lots of computational resources and requires lots of time.

All these approaches focus onto old type of phishing attacks.This paper proposes a solution to new type of attacks mainly focusing on tabnabbing attack.Very less researches are conducted to defend against this attack and one of the technique is to develop a Firefox add-on [1] to watch the open tabs and indicates whether the page has changed its layout, favicon and title. But this method fails if the user resizes the browser

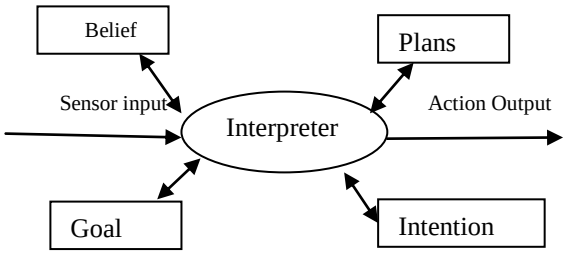

Fig 1.BDI Architecture

3. Multiagent Systems

An intelligent agent can be used as a powerful tool for solving and handling complex issues like phishing attacks. This paper suggests a new algorithm for resisting new types of phishing attacks using multiagent systems or distributed artificial agents. Multiagent systems help in achieving modularity of the problem. The decomposition allows agent to use most appropriate technique for solving the problems. The multiple agents in the system must co-ordinate with each other to solve some interdependent problems. These multiple agents are autonomous and heterogeneous in nature.

An agent in this system has a belief-desire-intention (BDI) type architecture. Each BDI agent or deliberative agent has planning, scheduling, execution, information gathering and co-ordination with other agents. It also has its own reasoning architecture and modules that operate asynchronously.

In planning module, a set of goals are taken as if and finds out a plan that satisfies the goals.In scheduling the plans that are produced during planning module are scheduled in this phase, ie, a step by step execution of each of the plan is created in this module. The execution of plans may sometime require co-operation from other agents. In such situations, agents may communicate with each other by sending messages. This task is done by communication module. The messages sent to other agents may contain request for services. These requests are taken as goals of the agent who has received the message. The agent has a layered architecture and has three main software layers. The layers are 1)Lowest layer or reactive layer 2)Middle Layer and 3)Uppermost layer

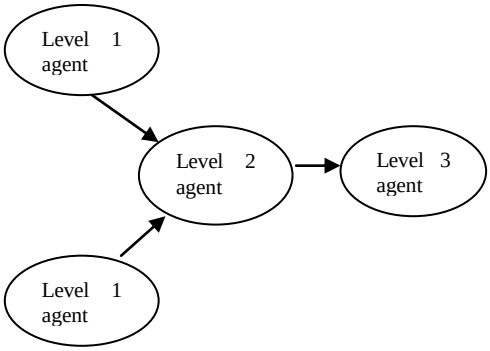

Fig 2.Levels of agents

- **The lowest layer or reactive layer**

This layer makes decisions about what is to be done when a problem arises.

- **The middle layer**

It abstracts about the inner view of agent's environment

- **The uppermost layer**

It represents communication and co-ordination with other agents.

Resolving the conflicts between agents is also one of the steps to be followed. The inconsistencies and disparities in agents' goals, belief, desires and results can happen in some situations. Negotiation is a method of resolving the disparities among multiple agents.

4. Agent Based Phishing Detection

This method is based on the inco-operation of autonomous distributed agents with strong level of intelligence. A distributed multi-agent system presents a great capacity for high level of learning , distribution of tasks and responsibilities, fault recovery, and adaptation to new changes. Mainly consists of agents in 3 levels and they communicate each other as needed.The level1 agents checks the URL of the webpage and confirms whether it is not blacklisted and checks the layout of webpage and confirms whether it is not changing when the webpage is idle for sometime. The level 2 agent detects the webpage as a phishing page or a legitimate page.In this mechanism level1 agent acts as an agent for detecting URL obfuscation and detects tabnabbing attack.level2 agent is acting as a webwatcher when the threshold values of actual webpages differ from the values received from its phished

webpage.A proper message will be displayed when the webpage is detected as a phished one.

The agents in different levels communicates with each other for achieving their goals.They learn and perceive from the environment and reacts accordingly.

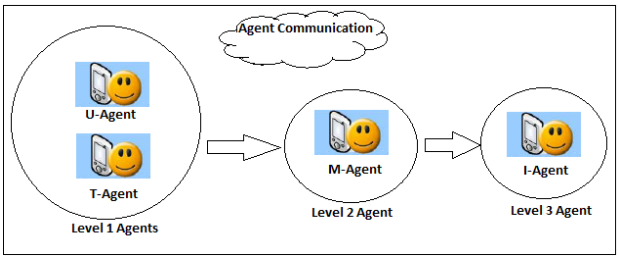

Fig 3.Distributed Agent Architecture

Level 1 Agents: Level 1 agents are acting as sensor agents.They are responsible for handling the incoming requests from the browsers for webpages.Two agents are there in level 1.A URL check agent(U-agent) and a Tabnab agent(T-agent).The U- agent is detecting URL based phishing attack and the T-agent is detectingtabnabbing attack of a webpage.The results of level 1 agents are sent to the agent at the next level in the hierarchy of the classification process.

Level 2 agent: Level 2 agent is purely a CBR agent which is acting as a manager who is responsible for coordination, communication, decision-making and its evaluation. This agent is like a webwatcher when an attack is detected and alarms the system about the attack.It also evaluates the different decisions arrived at for the process operation and takes necessary actions immediately

Level 3 agent: This agent acts like an interface agent which deals with the interaction of the user with the system. The interface agent is having responsiveness, competence, and accessibility. Interface agents communicates to the user about the detection of phishing attack and displays proper message.They are rule based applications and can act autonomously to perform operations without explicit directions from the user, and also they can collaborate with other types of software agents.

5. Implementation

This is an ongoing research work and is currently implemented by focusing on tabnabbing attack and can be extended for resisting any type of phishing attacks.In this paper ,a conceptual distributed agent-based framework is used to protect webpages from URL based attack and tabnabbing attack. The agents in this system are based on BDI architecture. The proposed methodology is implemented using JADE software framework. The distributed multiagents are having communication and cooperation with each other and they communicate via

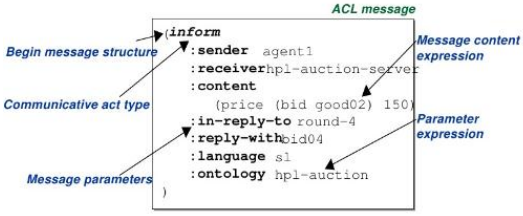

Fig 4. FIPA ACL Message

FIPA ACL.Figure shows an example of FIPA ACL message.

The FIPA works towards standardising agent systems so as to allow an easy interoperability. It is achieved by standardising the key components of the agent system, protocol transport levels and their interaction ontology. A FIPA-compliant Agent platform includes the AMS, the DF and the ACC. Startup activates all of these in the platform.

A standard message language is specified by FIPA ACL by setting out the encoding , pragmatics and semantics of the messages. It also specifies an encoded message between platforms in a textual form. This bypasses the need for a specific mechanism for transport of messages internally.

Fig 3 illustrates the architecture of the agent based antiphishing approach.The agents are deployed in different levels to completely deviate the ill effect of tabnabbing attack that can happen to a webpage.The core idea behind this framework is that a user cannot distinguish between the legitimate and the malicious webpage when the webpage layout is changed.A phisher exploits this limitation and tries to change the layout of a webpage.when it is unattended for sometime and prompts the user to enter the login details.This may cause financial losses to the user as the phisher can enter into user's personal information.

The working of the proposed method is as follows: First, the level 1 agents continuously monitors the change of the webpage URL or webpage layout as explained in [3].If it is changed,a message is communicated to the level 2 agent and alerts the system about the attack. When a malicious

webpage is found, a message is displayed to the user about the attack.

As an example, Fig 5 shows the original website of SBI. Fig 6 is detected as the fake website as it has changed the URL of legitimate site. Once a tabnabbing attack is found,the proposed method is cancelling the effect and alerting the user about the attack.This is shown in fig 7 .

Fig 5. Legitimate site of SBI

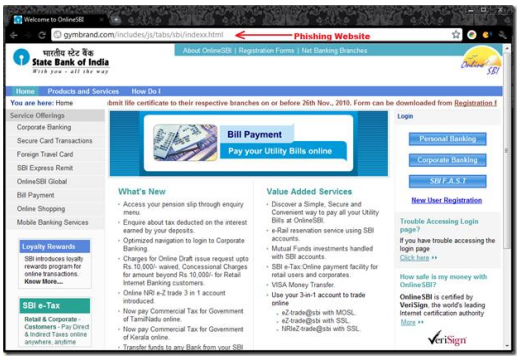

Fig 6. Fake site of SBI

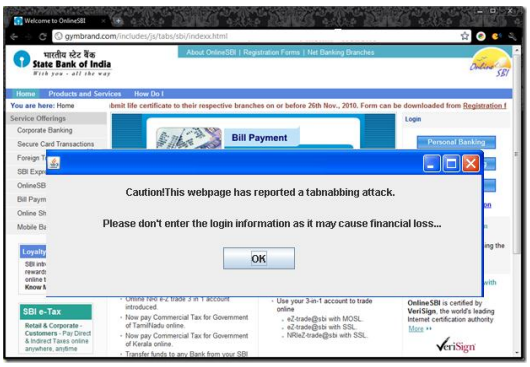

Fig 7. Phishing alert generated by agents

4. Conclusion

In this paper, an attempt is made to discuss about a distributed agent-based architecture that resists different types of phishing attacks. In recent years, online security has gained a lot of attention from industry.This paper suggests a solution to new type of attacks mainly focusing on tabnabbing attack.Agent technology has generated lots of excitement in recent years because of its effectiveness as a new paradigm for solving critical problems. This is particularly attractive for creating software that operates in environments that are distributed and open, such as the internet. Now, most of the agent-based systems consist of a single agent. As the technology advances and new problems arises, it is needed to implement systems that consist of multiple agents that communicate in a peer-to peer fashion. This research presents a software framework for detecting and rectifying web phishing attacks. The goal of security in a network is to provide the layer with sabotage resistance. Sabotage resistance means robustness against different types of attacks, such that an attacker cannot deceive a user.

It is a feasible approach as it ensures online security and resistance from phishing attack. The long term goal of this research is to develop a framework that can be used for resisting all types of phishing attacks .As it is an agent platform, it will be capable of perceiving and learning according to environmental changes.

References

[1] Seckin Anil Unlu, Kemal Bicakci, "NoTabNab: Protection Against The Tabnabbing Attack",IEEE 2010.

[2] Eric Medvet, Engin Kirda, Christopher Kruegel, "Visual similarity-based phishing detection", Computational intellligence in cyber security 2009. CICS '09. IEEE Symposium pages: 30-36

[3] A. Rosiello, E. Kirda, C. Kruegel, F. Ferrandi. "A Layout-Similarity-Based Approach for Detecting Phishing Pages".In IEEE International Conference on Security and Privacy in Communication Networks (SecureComm), 2007.

[4] M. Dunlop, S. Groat, and D. Shelly," GoldPhish: Using Images for Content-Based Phishing Analysis", in the Fifth International Conference on Internet Monitoring and Protection, 2010.

[5] A.Martin, Na.Ba.Anutthamaa, M.Sathyavathy, Marie Manjari Saint Francois,Dr.Prasanna Venkatesan, A Framework for Predicting Phishing Websites UsingNeural Networks, International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011,pp 330-336

[6] Y. Zhang, J. I. Hong, and L. F. Cranor, "CANTINA: a Content-based Approach to Detecting Phishing Web Sites", Proceedings of the 16th International Conference on World Wide Web (WWW'07), Banff, Alberta, CA, 2007

[7] Behrouz H. Far, "Distributed Software Agents for Network Fault Management," University of Calgary. Internal Paper.
[8] Hicham Tout, William Hafner "Phishpin: An identity-based anti-phishing approach" in proceedings of international conference on computational science and engineering, Vancouver, BC, pages 347-352, 2009
[9] Engin Kirda ,Christopher Kruegel, "Protecting Users Against Phishing Attacks with AntiPhish", Computer Software and Applications Conference, 2005.
[10] V.Shreeram, M.Suban, P.Shanthi, K.Manjula "Anti-phishing detection of phishing attacks using genetic algorithm" in proceedings of Communication control and computing technology(ICCCCT),IEEE international conference, Ramanathapuram , pages 447-450, 2010.
[11] Juan Chen, Chuanxiong Guo-"Online Detection and Prevention of Phishing Attacks (Invited Paper)"in proceedings of Communicational and networking in china, first international conference, Beizing, pages 1-7, 2007.
[12] Nourian, A.; Ishtiaq, S.; Maheswaran, M.;" CASTLE: A social framework for collaborative antiphishing databases", in Proceedings of IEEE- 5th International Conference on Collaborative Computing:Networking, Applications and Worksharing, 2009.

Sarika S is a B-Tech. degree holder in Information Technology,M Tech in Computer Science and Engineering and doing her research works. Currently she is the Assistant Professor in Computer Science and Engineering at Sree Narayana Gurukulam College of Engineering, Kolenchery, Kerala, India. Her research interest includes Mobile Ad hoc networks, Network security, Internet security.

Dr.Varghese Paul is a B-Tech. degree holder in Electrical Engineering, M Tech in Electronics Engineering and PhD in Computer Science. Currently he is the Professor in information Technology at Cochin University of Science and technology, Thrikkakara, Cochin, Kerala, India. His research interest includes Fault Tolerent Computing and Data security.