Research article Available online www.ijsrr.org ISSN: 2279–0543

International Journal of Scientific Research and Reviews

Development of Secure Framework for Cache Using Phishem Guard Algorithm

[*1]G. Gopu and [2]K. Thangadurai

[*1]Ph.D. (Category B) Research scholar, Bharathiar University, Coimbatore – 641 046, Tamilnadu, India. Email: vggopu@gmail.com,
[2]Research Supervisor Bharathiar University, Bharathiar University, Coimbatore – 641046, Tamilnadu, India. Email: ktramprasad04@yahoo.com

ABSTRACT:

Today, The security is a technical term which is more important part of human day to day life and being involves in securing their day to day life from theft and malicious activities. Due to an increasing of technological advancements, the domestic and international suspicious activities has been increased. The electrical and electronic devices can be a tool to commit a targeted activity such as phishing. Phishing involves in gathering sensitive information through e-mail and websites. The article presents the role of abnormal traffic in electronic mail and intention with their way of thinking, planning and performing attacks for their personal gain and as well as others grow up. We mainly focused on security measures to prevent phishing on E-mail and to address "secure framework for cache" using Pish EM algorithm.

KEYWORDS: Introduction, Related work, Statement of the problem, Classification of Phishing, Research objectives, Proposed work, An Algorithmic Approach, Performance Evaluation.

[*]Corresponding author:

G. Gopu

Ph.D. (Category B) Research scholar,
Bharathiar University,
Coimbatore – 641 046,
Tamilnadu, India.
Email: vggopu@gmail.com,

1. INTRODUCTION

People all over the world practices e-mail to stay in touch with relatives and share movements of jay and sorrow, transfer important documents, forward meaningless junk to friends, play tricks and even maintain conditional business deals, all within a matter of seconds. More and more individuals and organizations depending upon e-mail for their daily dose of critical communication, some of which may contain personal information, company secrets and sensitive information. Because of it's an importance, the suspicious attacks involving in different method by a professional and others for their purpose or for others benefits. All these are happens only the intension of gaining something from others work[1].

2. RELATED WORK

The study describes an algorithmic approach for a secure framework for a cache and relates the review of literature to predict phishing attack on electronic mail service[1,3]. There are many similarities between E-mail and physical mail while exchanging information as mentioned bellow.

To **send** and **receive** email message, requires an Internet connection and access to a mail server with standard protocol called SMTP.

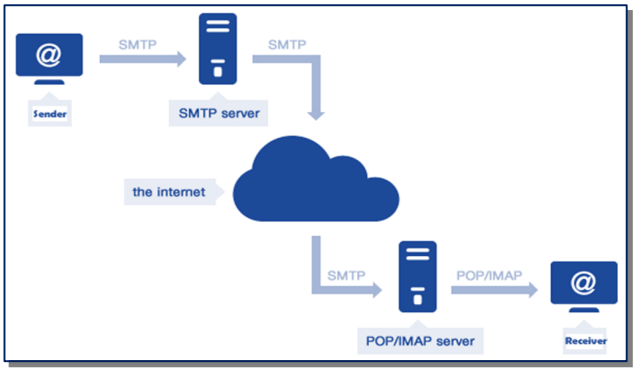
When a message is sent, the email client sends the message to the SMTP server. If the recipient of the email is local the message is kept on the server for accessing by the POP (Post Office Protocol), IMAP (Internet Message Access Protocol) or other mail services for later retrieval.

If the recipient is remote, the SMTP server communicates with a Domain Name Server (DNS) to find the corresponding IP address for the domain being sent to. Once the IP address has been resolved, the SMTP server connects with the remote SMTP server and the mail is delivered to this server for handling.

If the SMTP server sending the mail is unable to connect with the remote SMTP server, then the message goes into a queue. Messages in this queue will be retried periodically. If the message is still undelivered after a certain amount of time (30 hours by default), the message will be returned to the sender as undelivered [3].

3. REVIEW OF LITERATURE

The intention of study is to present a comprehensive detail about different approaches related to review of phishing email to achieve the maximum accuracy result and improve the whole system.



3.1 Anomaly detection approach

The anomaly detection approach implies its choices on a profile of traditional network or system behavior by applying statistical or machine learning techniques. Any event that doesn't match to the current profile is taken into account as anomalous. Several experimental anomaly detection systems exist and the non-technical developers also adopt the anomaly detection approach.

Hui Keng Lau et al. (2009) proposed an anomaly detection mechanism with adaptive normal model that is inspired by immune system for email classification. This is motivated by the analogy between immune system and the mail system architecture and also the learning and tolerance ability of the immune network.

The strength of the anomaly detection approach is that prior knowledge of the security defects of the target systems is not required. Thus, it is able to identify not only known but also unknown suspicious activities. In addition, this approach can identify the suspicious actives that are caused by the abuse of legitimate users or masqueraders without breaking security policy.

3.2 Hybrid approach

The hybrid approach consists of an identity-based detection component and a keywords retrieval detection component both manipulating the DOM after the mail has been rendered in SMTP to get around intentional obfuscations. It relies on individuality identification to find the domain of the page's stated identity, and inspects the authenticity of the header of page by comparing this extracted domain with its own domain via executing the query of the form [4].

Ma et al. Proposed a research on developing tools and techniques to detect phishing email using hybrid features. Phishing has become much suspicious, complex and sophisticated that it is able to avoid filters and anti-phishing systems. Email servers now can be installed with malicious detection devices since phishing emails have initiated a lot of researchers to work on creating these techniques[4].

3.3 Cache Management Approach

A cache feature is incorporated in DB2 universal database by modifying the engine code and leveraging existing data functionality. This allows us to take advantage of DB2's sophisticated query processing power for a secure data caching. As a result, the user queries can be executed at either the local database cache or the remote backend server, or more importantly, the query can be partitioned and then distributed to optimum execution.

Umamaheswari S (2014) proposed a research on Hybrid IDS that detects the anomalies and DCS locates and fetches the required data by cache discovery mechanism. The integration of Hybrid IDS

and cache management is done in the data cache framework. This framework consists of the Cache Management architecture and the Cross Layer. The data packets are embedded with the control packets CLCp and AISp to handle the node failure and misbehaviour.

3.4 Ant Colony Optimization (ACO) Approach

The importance of the Ant System (AS) resides mainly in being the prototype of a number of ant algorithms which collectively implement the ACO paradigm. ACO is a class of optimization algorithms modeled on the actions of an ant colony. ACO methods are applied to the problems that need to find paths to goals.

Nada M.A.Al Salami (2009) proposed a hybrid algorithm to solve combinatorial optimization problem by using Ant Colony and Genetic programming algorithms.

Subodh M. Iyengar et al. (2010) analysed the ant algorithms ARA and AntHocNet. The performance of these algorithms is compared with AODV based on the effect of its routing mechanisms on its routing efficacy. A modified version of ARA has also been suggested.

AGO strategies can be applied for knowledge acquisition other than routing. A novel algorithm attempts to formulate a solution for the well-known knowledge acquisition problem of losing interest in content based documents due to low familiarity. Documents are found via pheromones deposited by such ant colony.

4. STATEMENT OF THE PROBLEM

Even if the electronic mail service is widely used mode of communication it is not the safest and most reliable. Virus infected e-mails and eavesdropping on network lines are the main reasons that affect the reliability of internet e-mail communication[23].

At this stage, it is an important to point out the insecurity in e-mail delivery pathway:

A user's email messages can be compromised in four locations: the sender's device, the network, the server, and the recipient's device. The first and last should be comprehensible to anyone, regardless of tech savviness; email accounts are usually always logged in, so anyone sitting at a computer or holding a phone should be able to read any email message they choose. Email services rarely encrypt saved messages, so reading emails and attachments is as easy as opening the program or navigating to the webpage.

An email message might travel through dozens of routers and switches on its way to a recipient, and each transfer is an opportunity for suspicious activities. There is no guarantee that each connection is equally secure. Email servers are rarely encrypted, because of the overhead costs of encryption as well as the value of saving messages in plaintext, so hackers with admin passwords or

access through security flaws can search vast swaths of emails for personal data. From sending to receiving, saving to deleting, email is unbelievably insecure.

Even if the E-mail technology is equipped with strong anti-virus application and spam filters to stop spam, viruses, and other malicious content reaches the network infrastructure with incredible programming knowledge and their motivations such as phishing.

5. CLASSIFICATION OF PHISHING

Phishing emails are sometimes focused on conjunction with spam emails. The two categories do share one unifying attribute in that they are both unsolicited contacts. But there are substantive differences between the two.

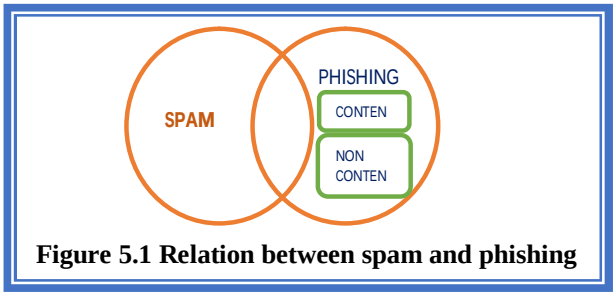


Figure 5.1 Relation between spam and phishing

The key different between spam email and phishing email attack is the intended end action on the part of the targeted user. With spam emails a webpage hires an intermediary spam to deluge large numbers of email addresses with unsolicited offers for services. Some of these are of suspicious activities. But the spammer only cares about the raw number of spam emails set and their conversion rate into hits on the client's web site.

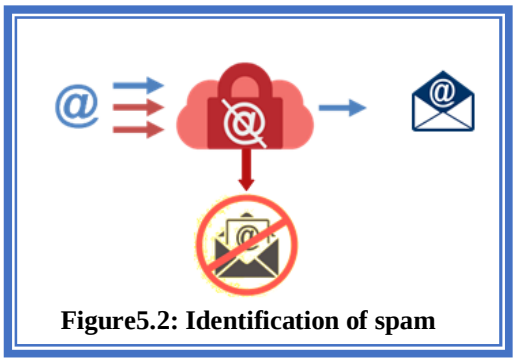


Figure5.2: Identification of spam

Phishers, as opposed to spammers, are serving their own goals. If a phisher sends out a phishing email then he/she wants to collect the targeted user's information or resources for personal use. Sometimes the attack itself belies the intended use of the user's information or resource. A phishing email that requests details of financial accounts likely wants to use that information to steal money out of the account in question. Whereas a phishing email that tries to exploit a software vulnerability

on a user's computer wants the resources that come with being able to run software on someone else's machine; bandwidth and disk space to name two.

Spam is the unsolicited emailing of offers for products and services. The users receiving the spam emails may not be of the target audience for the products and services in question.

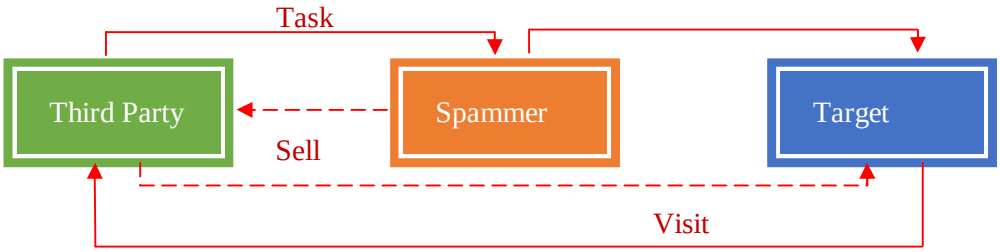


Figure 5.3: The sequence of events in a spam campaign

Figure 5.3 shows the events and messages involved in an end-to-end spam email campaign. In this example, a third party has hired a spammer to send unsolicited product offers to email addresses across the Internet. The goal is to entice some number of spam recipients into paying customers of the third party. The target user doesn't need to ever interact again with the spammer again.

The theft of user login credentials is where phishing began. Once an attacker had a user's login name and password the attacker could log in as that user and steal whatever resources that account might have access to. In the case of the AOL account theft described previously, the attacker wants to steal Internet access.

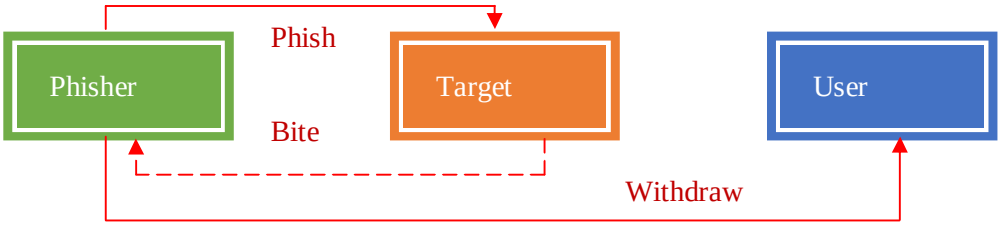


Figure 5.4: Phishing for account information

Figure 5.4 makes it immediately obvious what the difference between spamming and phishing. The phisher wants to steal a target user's information for the purposes of making fraudulent withdrawals from the target's user account.

6. RESEARCH OBJECTIVES

The main objective of the study is to design and develop a framework for secure cache using Phish EM guard algorithm for mail security awareness system[1, 3, 11].

- To give the nature of Email security on SMTP, and the fact that data security threats only escalate.
- To perform the analysis of existing security algorithms, clustering schemes and routing protocols in SMTP.

- To design and develop a Hybrid Detection System model for communication network using Ant Colony Optimization (ACO) to detect the anomalies in the network in order to make the devices in email network safe and secure.
- To design and develop a Data Caching Scheme (DCS) that improves the data accessibility ratio and reduces the query delay by proposing a cache discovery algorithm using ACO.
- To design an enhanced data cache framework by that improves security during the data retrieval. This is achieved by embedding immunity into data packets thereby increasing data accessibility and reduces query delay with the support of cross layer design approach and clustering.
- To address the development of framework to ensure and secure the data access on SMTP.

7. PROPOSED WORK

We propose to adopt the research work is to produce new idea, knowledge and awareness in the form of research methods. The study will design and develop framework for securing a data from attack. These security methods will include authenticating routing updates, erecting a firewall, detecting intrusion attempts and responding with group rekeying measures designed to isolate the attacker[1, 3].

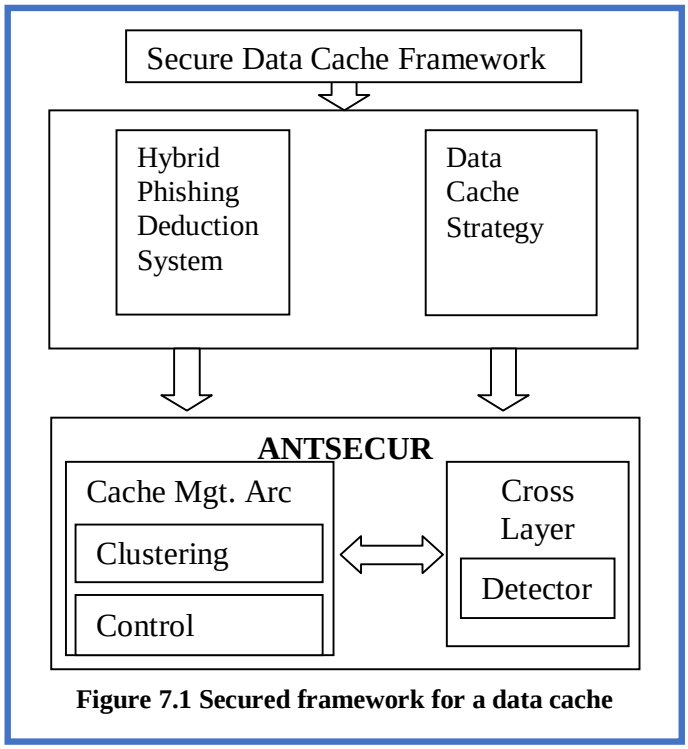


Figure 7.1 Secured framework for a data cache

It improves solutions to a problem. Here the study divides into two models theoretical and simulation model. In the theoretical model study implies different security issues and their solutions. In the simulation model run simulation with configuration and try to learn mechanisms which will

help us to enforce security in SMTP[3].

7.1 Secured architectural framework of a data cache

The Hybrid Phishing Deduction System detects the anomalies and Data Cache Strategy locates and fetches the required data by cache discovery mechanism. The integration of Hybrid PDS and cache management is done in the ANTSECUR Framework [3][4]. This framework consists of the Cache Management architecture (Cache Mgt. Arc) and the Cross Layer. The data packets are embedded with the control packets Cross Layer protocol and detector set to handle the node failure and misbehavior.

8. AN ALGORITHMIC APPROACH

There are many approaches under Swarm Intelligence, out of which Ant Colony Optimization (ACO) Algorithm is proved to produce an optimized solution in automatic knowledge gathering problems for awareness, robust routing, screening, tracking etc[19,21,22]. ACO is a class of optimization algorithms modeled on the actions of an ant colony. ACO methods are applied to the problems that need to find paths to goals.

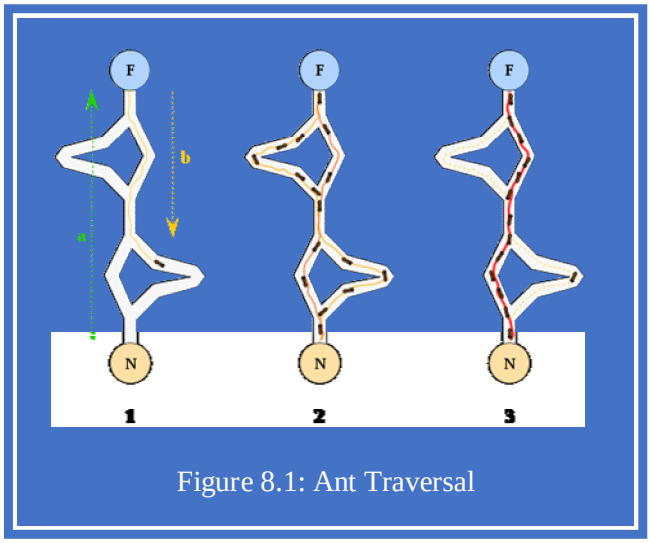


Figure 8.1: Ant Traversal

Real ants lay down pheromones directing each other to resources while exploring their environment. The simulated 'ants' similarly record their positions and the quality of their solutions, so that more ants trace better solutions in later simulation iterations. More the density of the pheromone trails indicates that there is a food source present and each and every time, the ants travel in the same path, they increase the density of the pheromones and thus making it as a strong source of food. The paths less travelled by the ants have low density of pheromone trails which do not indicate the strong source of food. This concept of ACO can be brought in analogous to the email

networks where the devices in the network exchange data packets between the other devices. These type of exchanging of information is brought in by the ACO algorithm to route the information in the shortest path.

The following is the pseudo code of ACO.

```
start procedure ACO_MetaHeuristic
do while(if not termination)
antGeneration_and_Activity()
phishActions()
phishUpdate()
end do
end procedure
```

The ant Generation_and_Activity function deals with the generation of ants, its movement and activities related to it. Phish Actions() are used to implement centralized actions, which could not be implemented by a single ant, such an invocation of localized optimized procedure or the update of global information to be used to decide whether to bias the search process from a non-local perspective. Phish Update is used to disappear trail values over time to avoid the unlimited accumulation of trail values over a connection[19,21, 22].

9. PERFORMANCE EVALUATION

The proposed ANTSECUR framework was evaluated in an NS-2 simulation environment [30]. Different simulation scenarios have been performed. The performance of the proposed framework is measured with packet delivery ratio and end to end delay. In the simulation, each host moves in the simulation area following the random waypoint mobility model. The random waypoint model is used for simulating the movement pattern of email data in SMTP. For 150 and 200 nodes simulation was carried out in 1500*1000m. In Table 6.2, the simulation parameters are listed.

Table 9.1. ns-2 Simulation Parameters

Parameter	Value
Transmission Range(M)	300
Bandwidth(Mbps)	3
Node Speed(M/s)	0-10
Routing Protocol	Ant Phish Net
Pause Time(S)	100
Cache size (KB)	300
Average TTL (S)	100-3000
Zipf-like Parameter ( )	0.5-1.0
Number of Data items	1000
No. of nodes	150,200
Request Interval(S)	10
Simulation Time	2000 sec.

To evaluate the performance of the proposed ANTSECUR framework, two scenarios based on

number of client nodes have been considered. There are five Tcl simulations run have been conducted in 150 and 200 nodes, based on which average end to end delay and packet delivery ratio has been obtained. Generally, misbehaving node may cause raise in delays, packet drop and reduce throughput of the network. To reduce this factor AISp has been embedded within the data packet. This prevents the network from nodes that cause abnormal behaviour.

Table 9.2. Network Performance on 150 nodes

(150 Nodes) Scenario1	Packet Delivery Ratio		Delay	
	ANTSECUR	AODV+COCA	ANTSECUR	AODV+COCA
1	96.51	91.38	195.33	223.39
2	95.03	90.25	234.53	258.04
3	92.92	89.57	142.70	160.41
4	96.06	89.42	254.68	321.93
5	97.38	93.96	191.44	238.47
Mean	95.58	90.92	203.74	240.44
SD	1.71	1.87	43.28	58.41

The mean value of packet delivery ratio and end to end delay of ANTSECUR framework presented in Table 6.3 shows good performance than AODV+COCA as it contains cross layer control CLC. This CLC manages path distraction due to node failure which certainly reduces end to end delay. As there is decrease in delay, packet delivery ratio gets increased. But there is no control for managing Path distraction in case of AODV+COCA. The delivery ratio was 95.58% in the ANTSEC framework where it was 90.92% in case of AODV+COCA. ANTSEC framework has 22.91% efficient in reducing the delay and 7.70% efficient in improving the packet delivery ratio than AODV+COCA while transmitting the packets.

Table 9.3. Statistical Analysis for 150 Nodes

150-nodes	**Mean**	**SD**	***t-value***	**df**	***p-value***
PDR-ANTSEC vs. PDR-AODV+COCA	6.83	1.99	11.45	6	0.003
DELAY-ANTSEC vs. DELAY-AODV+COCA	-53.73	29.71	-6.06	6	0.024

Table 9.4. Network Performance on 200 nodes

(200 Nodes) Scenario2	**Packet Delivery Ratio**		**Delay**	
	ANTSEC	**AODV+COCA**	**ANTSEC**	**AODV+COCA**
1	125.32	120.24	204.41	228.39
2	121.51	116.81	359.91	405.39
3	113.01	108.60	406.24	450.43
4	128.27	123.47	274.17	319.96
5	126.96	122.83	227.59	276.99
Mean	**123.01**	**118.39**	**294.47**	**336.23**
SD	**6.1**	**6.07**	**86.3**	**91.15**

ANTSECUR framework shows good performance in the Tcl simulation runs as specified in Table 6.5. The delivery ratio was 92.26% in the ANTSEC framework where it was 88.79% in case of AODV+COCA. In case of delay, ANTSEC transmits the packet within 220.85 ms where as

AODV+COCA takes around 252.17ms. The proposed framework has 12.42% efficient in reducing the delay and 3.9% efficient in improving the packet delivery ratio than AODV+COCA while transmitting the packets.

Table 9.5. Statistical Analysis for 200 Nodes

200-nodes	Mean	SD	*t-value*	df	*p-value*
PDR-ANTSEC vs. PDR-AODV+COCA	4.63	0.36	37.92	5.33	0
DELAY-ANTSEC vs. DELAY-AODV+COCA	-41.76	10.13	-12.29	5.33	0.001

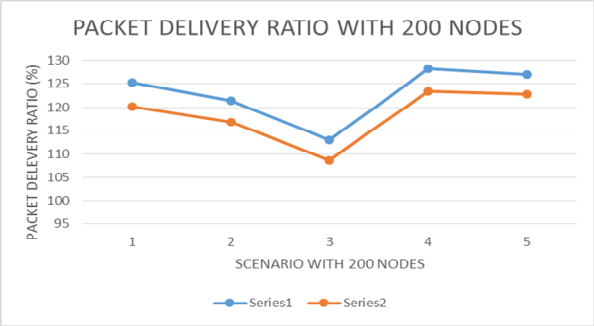


Figure 9.1 Packet Delivery Ratio with 200 Nodes

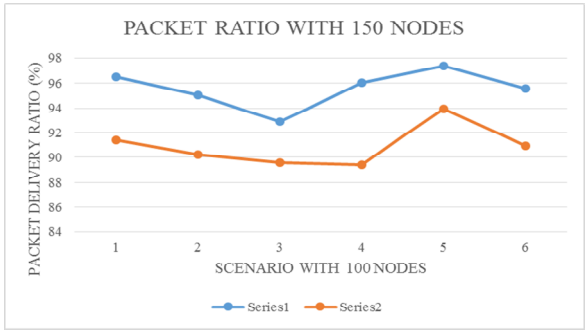


Figure 9.3 Packet Delivery Ratio with 150 Nodes

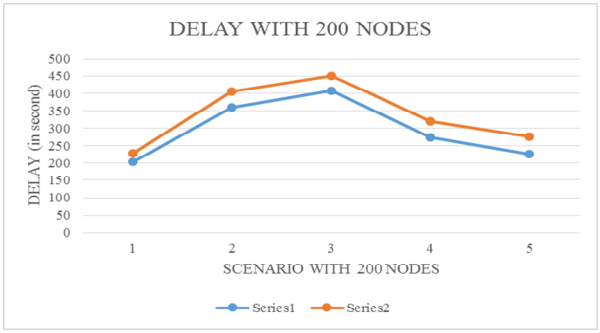


Figure 9.2 Delay with 200 Nodes

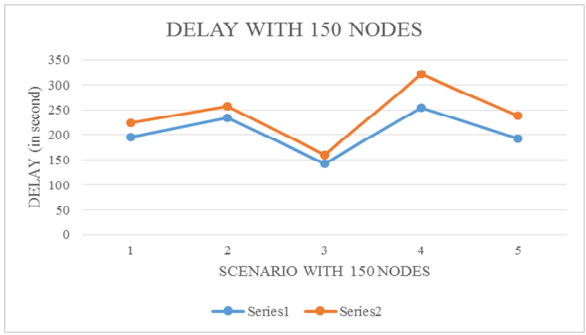


Figure 9.4 Delay with 150 Nodes

The packet delivery ratio of the network with 200 nodes and 150 nodes is depicted in the figure 9.1 and figure 9.3 respectively. The delay of the network with 200 nodes and 150 nodes is depicted in figure 9.2 and figure 9.4 respectively.

The statistical paired t-test has been proposed to evaluate the significant difference between proposed model and existing model with respect to Packet Delivery Ratio (PDR) and Packet Delay. Hence, the proposed model has found significant effect on existing model with respect to PDR and Delay in both the scenarios with 150 nodes and 200 nodes.

The ANTSECUR framework always outperforms AODV+COCA as it embeds ACO and Phish EM; both factors make the proposed framework more efficient with higher PDR (Packet Delivery Ratio). As the framework uses SMTP and the misbehaving nodes are identified by the email data which are embedded into the Data Packets (DP) the end-to-end delay is lowered. From the above results, it is proved that the proposed ANTSECUR framework works well in large network size

including malicious nodes.

CONCLUSION

E-mail, being a global phenomenon is bound to attract many crimes such as phishing. Around the world government has taken a key step in prevention of phishing E-Mail by the enactment of the Information Technology Act and by giving exclusive powers to the police and other authorities to tackle such crimes.

Similar efforts have been made to fight E-Crime by enacting national legislations but in the long run, they may not prove to be as beneficial as desired. An effort is still wanted to formulate an international law on the use of Internet to control this forthcoming danger of E-mail related crimes and to achieve a crime free Cyber Space. Prevention they say is the best cure. Cyber Laws aim to prevent cybercrimes through the use of penal provisions.

REFERENCES

1. Dr. K. Thangadurai, Gopu.G, "A secure data cache framework to deduct phishing E-Mail using ACO technique" IJRAR January 2019, E-ISSN 2348-1269, P-ISSN 2349-5138.
2. Annu K Simon, Dr, S Subasree, "Design and Development of Enhanced Optimization Techniques based on Ant Colony Systems", IJIRST - International Journal for Innovative Research in Science & Technology Volume3, Issue 04, ISSN (online): 2349-6010, September 2016.
3. Dr. K. Thangadurai, Gopu.G, "A study on Ethical Phishing on E-mail networks and its impacts in India", Indian Journal of Science and Technology, Vol 9(45), DOI: 10.17485fijst/2016/69i45/90847, December 2016.
4. Karthika Renuka D, Visalakshi P, Girish R "A Hybrid ACO Based Feature Selection Method for Email Spam Classification", WSEAS transactions on computers, E-ISSN: 2224-2871, 2015; 14.
5. Ankur Dumka, Ravi Tomar, J. C. Patni, Abhineet Anand, "Taxonomy of E-Mail Security Protocol", International Journal of Innovative Research in Computer and Communication Engineering, April 2014; 2(4).
6. P. Rohini, K. Ramya, "Phishing Email Filtering Techniques A Survey", Volume 2, Issue 1, February 2011, Nov 2014; 17(1).
7. Gori Mohamed .J, M. Mohammed Mohideen, Mrs. Shahira Banu.N "E-Mail Phishing – An open threat to everyone" International Journal of Scientific and Research Publications, ISSN 2250-3153 February 2014; 4(2).

8. SANS the monthly security awareness newsletter for computer users, "Email Phishing Attacks", February 2013.
9. Phishing attack against MSN/Hotmail users - a new year, but old tricks still persist by Graham Cluley, Monday, January 2013; 14.
10. N. Vijayalakshmi, E. Sivajothi, Dr. P. Vivekanandan, "Efficiency and Limitation of Secure Protocol in Email Services", International Journal of Engineering Sciences and Research Technology, Nov-2012; 539-544.
11. Umamaheswari S, Radhamani G, "Clustering Schemes for Mobile Ad Hoc Networks: A Review", IEEE sponsored Second International Conference on Computer Communication and Informatics (ICCCI 2012), 2012.
12. Sunny gill, Gaurav Rupnar, Vaibhav Ramteke, Prof. Dipit Patil, Vijay M. Wadhai. "Email Security Protocol", International Journal of Computer Trends and Technology – March to April Issue 2011.
13. Shamal Firake, Pravin Soni, Dr. B.B. Mesharam, "Phishing E-mail Analysis", International Journal of Computer Trends and Technology, February 2011; 2(1).
14. Hung-Min Sun, Bin-Tsan Hsieh, Hsin-Jia Hwang, "Secure E-mail Protocols Providing Perfect Forward Secrecy", February 2011; 2(1).
15. Mrs. K. Shanmugavadivu, Dr M. Madheswaran, "Caching Technique for Improving Data Retrieval Performance in Mobile Ad Hoc Networks", (IJCSIT) International Journal of Computer Science and Information Technologies, 2010; 1(4): 249-255.
16. Nada M.A.Al Salami, M.A., "Ant Colony Optimization Algorithm", Ubi CC 2009; 4(3): 823-826.
17. Stuart McClure, Joel Scambray, George Kurtz, "Hacking Exposed 6 Network Security Secrets & Solutions" 10th Anniversary Edition, 2009, the McGraw-Hill.
18. Chu Yan, Zhang Jian Pei, Zhao Chunhui, "Data Cache Strategy Based on Colony Algorithm in Mobile Computing Environment", IEEE International Conference on Internet Computing in Science and Engineering, 2008; 235-238.
19. Dorigo, M., Mauro Birattari, Thomas Stiizle, "Ant Colony Optimization", IEEE Computational Intelligence magazine, November 2006.
20. M. Altinet, Q. Luo, S. Krishnamurthy, C. Mohan, H. Pirahesh, B. G. Lindsay, H. Woo, L. Brown, "DB Cache: Database Caching for Web Application Servers", 612, SIGMOD 2002.
21. Cordon, O., F. Herrera and T. Stützle, special issue on "Ant Colony Optimization", Mathware and Soft Computing, November, 2002; IX: 2-3
22. James Kennedy, Russel C.Eberhart, Yuhui Shi, "Swarm Intelligence", Morgan

Kauffman Publishers, 2001.

23. Ankit Fadia, “E-mail Hacking”, Vikas Publishing House Pvt. Ltd.
24. Rajendra Maurya “HACKING MADE EASY”, SCORPIO NET SECURITY SERVICES Publication, 2nd Edition.
25. Agrawal, D. P. and Qing-An Zeng, “Introduction to Wireless and Mobile Systems”, 2005, Brooks/Cole.
26. Andrew S. Tanenbaum, David J. Wetherall, “Computer Networks, Pearson”, 2011.
27. CISCO Security White Paper “Email Attacks: This Time its Personal”, June 2011.
28. Behrouz A, Forouzan, “Data Communications and Networking”, New Delhi: Tata McGraw-Hill, 2011.
29. Artail, H., Haidar Safa, Khaleel Mershad, Zahy Abou-Atme, Nabeel Sulieman, “COACS: A Cooperative and Adaptive Caching System for MANETs”.
30. Network Simulator 2, http://www.isi.edu/nsnam/ns
31. Hacking : http://www.hackingarticles.in/
32. Mail Examiner: http://www.mailxaminer.com.
33. Email racer: http://www.cyberforensics.in/Products/emailtracer.aspx
34. Internet Engineering Task Force (IETF) www.ietf.org