ISSN No:-2456-2165

Enhancing Mail Server Security Using Machine Learning

Swati S Maddur, Richul N Prasad, P Jyothi Priya, Pruthvi Sainath Reddy, Dr. Sumithra Devi K.A
Department of Information Science and Engineering
Dayanand Sagar Academy of Technology and Management
Bengaluru, India
swatism312@gmail.com, richulnprasad@gmail.com, jyothi.priya2710@gmail.com, sainathreddy98@gmail.com, deanacademics@dsatm.edu.in

***Abstract*— Even with the advancement in technologies for providing at most security to users, there is always a glitch in the implementation or the algorithm used, which means compromise on user's security or rather privacy. Most of the standalone servers do not provide much protection beyond password security and spam control. The paper aims to build a machine learning model that moulds according the user's e-mail usage pattern ,and hence, no third-party data sets are needed to train the model. As we progress, further, in the paper we focus on building a unsupervised machine learning model to improvise the security of mail server protocols such as SMTP, IMAP, SSH, etc.**

Keywords - Mail Server Security, Machine Learning , Security SMTP, IMAP and SSH protocols.

I. INTRODUCTION

Whenever we speak of mail related security issues, we think of it as message applied security measure, and more often to antivirus and antispam protection. However, this is only one stage in the complicated process of securing the mail sever. Most of us use password based system to access our mails which provides only a layer of security. If the password is known to someone it can be easily manipulated to send spams or used to the extract the user sensitive information. One way to enhance security is limiting the number of connections or authentication errors, the maximum number of commands or setting a time-out for the sessions, number and size of email messages.

Most common Mail Server security measures these days is only concerned with authentication of user and building security policies based on a fixed set of rules. In this project, as an alternative to these simple and fixed rules we take a different approach by building a Machine Learning System that will be trained with the email habits of a user through its mail server and then provide security against account compromise vulnerabilities. Unlike, most Machine Learning models which need to be trained with large data sets in order to further use it for decision making, here, we won't rely on third party data sets, instead we will collect data from each mail server. Thus, for each instance our ML Algorithm will start from Zero-Data and then grow upwards. Also, Most Machine Learning algorithms require a data collection system which can in turn act as a Surveillance system.

Thus, our algorithm will entirely run on a local network in conjugation with the mail server. The mail server data will neither be shared to third parties nor will be stored there. in this way we can provide high level security by leveraging the benefits of Machine Learning to those who care about data privacy and don't want to systems. With most of the proprietary software used for email transaction, almost all of them illegally track the email behavior of the software user's and then collect these data-sets and sell it out in the market for higher shares. This is being done without the user consent , thereby violating their civil liberty rights.

So for those of the user's who are concerned about safe-guarding and protecting their data from the third party services and run their own mail-servers to allow mail transmission we build a machine learning model that will incorporate unsupervised methodologies to build individually according to the user's email behavior.

II. IMPLEMENTATION

In This Survey Paper the existing methodologies of Spam Detection and Authentication led us to arriving at a much more reliable way of authenticating users to their mail accounts.

A. Recognizing Abused Mail Accounts

The methodology used in [1] detects compromised mail accounts using the Metadata from MTA (Mail Transfer Agent e.g: SMTP)with information like IP, country and Delivery Status Notification (DSN) , Sent time , Message Identifier (Message-ID):

- IP-Geolocating – Country of the authentication point From Source IP
- Delivery Status Notification - Status of Delivery from one MTA to Another.
- Message-ID -Correlates information between incoming and outgoing messages through MTA.
- Timestamp- To correlate information in a chronological order.
- Further Information- Like, No. Of Recipients, Message Size etc.

The Metadata from the MTA (Mail Transfer Agents) and MDA (Mail Delivery Agents) can be used to detect abused

mail accounts by drawing an analysis between connection times , and IP addresses used to calculate difference between authentication point from MTA and access time on the MDA within the same account and IP address in a given time interval. Hence , being able to differentiate between a normal and abused mail account upon describing the line between their individual behaviors.

This paper highlights the fact that delivery of emails has a specific behavior that distinguishes them from spam which makes it possible to detect abused email accounts after delivery of one or more emails, but spams are blocked at their generation site and not at destination site. Hence, this paper suggests using only a small amount of metadata to get geographical destination and authentication. Therefore, this method has faced successful implementation in one the university where it's real-time emails were tested and detected approximately 27,000 abused accounts in April 2015.

B. Keyword Extraction

This survey paper focuses on formulating implicit queries to a just-in-time-retrieval system[2]. It is mainly used to extract the required keywords for the model from a large dataset with much larger number of words in the logs. Just-in-time-retrieval system is used here as it is helpful to extract only the necessary data which is then clustered to form different types.

The methodology used here is the novel keyword extraction technique from ASR output[2]. This technique can be preferred over the other as it maximizes the chances of including the overall potential information that is necessary. It means that we can obtain only the important or necessary information for our model. The server logs that we obtain is a long text (string) from which we need very particular information like IP address, location etc. This keyword extraction model helps us achieve that requirement. After extracting these necessary keywords, these will be clustered to build topically-separated queries and these are finally merged into a set which is ranked orderly.

Keyword Extraction Methods: The dataset for this particular project can be a multi-variate or multi-frequency data. There are many models that can handle such datasets with its semantic representations such as WordNet, LSA, PLSA or LDA[2]. Some supervised machine learning methods are also used for such keyword extraction models.

Diverse Keyword Extraction Algorithm: This paper introduces various modeling techniques and its advantages for representing multi-frequency data. It first fragments the data, select content words as keywords and applies various summarization methods. The important advantage of this model is that the conversation of data fragments here is maximized.

Fig.1 shows keyword extraction rates for variable datasets.

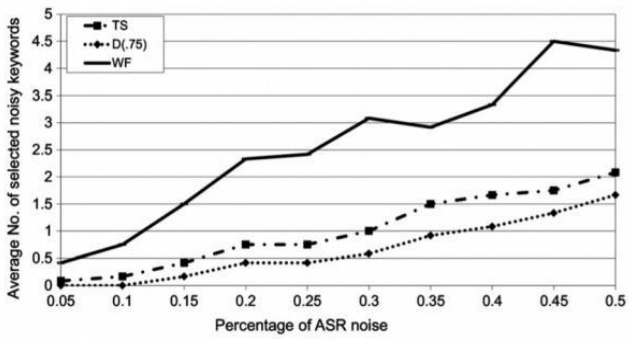

Fig 1: Keyword Extraction Rates For Variable Datasets [2]

Average number of noisy keywords chosen by the algorithms over the 8 conversation fragments of the AMI Corpus, for a varying percentage of artificial ASR (automatic speech recognition) noise from 5% to 50%. The best performing method is D.75 (diverse keyword extraction method).

This paper compares the diverse keyword extraction method with existing methods, based on word frequency or topical similarity, in terms of the representatives of the keywords and the relevance of retrieved documents. These are tested by human ratings obtained from Amazon Mechanical Turk crowd sourcing platform.

C. Tracking User's Email Behavior through Usage Pattern

BDI model is an agent model, which is widely used to simulate user behavior in complex and dynamic environment. When a user clicks on any malicious link sent via email the user host is subjected to virus. This action of the user operation on the mail is called as :user's email behavior. To better secure the user in a way to protect them from the malicious URL links or spam, authentication problems, etc BP-BDI model [3] proposes a machine learning model that simulates user's email behavior accurately and effectively.

When trying to understand a user's behavior there are notably two difficulties: the difference and the complexity of online user behavior. Difference in the user's behavior supposedly refers to the individual behavior pattern in mail activities such as: sending, receiving, deleting the email and so on. Due to the complex and dynamic environments, user online behavior becomes complicated, that is, the complexity of user behavior. Hence, it means that we cannot adopt the traditional procedure-oriented and object-oriented modeling method to track user's email behavior.

The main focus here is to track the user's email usage pattern to better train the model to adapt to individual user's mail server. Therefore, in order to track user's email behavior better, we modify the Belief-Desire-Intention model in two aspects: one where it carries out automatic updates to the belief set, and the other where a behavioral learning module is constructed based on BP(back propagation) neural network. In the BP-BDI model, the learning module mainly studies the behavior of clicking on URL, which makes user's email behavior involves network security is accurately simulated. In

the simulation, we employed the three layer BP neural network to study features of the user's behavior. In a nutshell, the BP-BDI model can improve the accuracy of simulation based on the behavioral learning module which is experimentally found to be 80% accurate. We will be using this model as a reference to build our module which requires similar training to better understand email user's behavior.

This paper proposes BP-BDI model to track user's email behavior by constructing a learning model with BP neural network and has shown an accuracy of 80% while simulating the email user's behavior.

D. True IPv6 Address Access

The secure e-mail system is divided into authentication module, mail proxy module and management module. Authentication module is the main body of mail source address authentication, and the mail source address authentication is divided into inter-domain and intra-domain authentication[5]. Mail proxy module mainly realizes some functions, for example, sending, receiving, forwarding and storing mail.

Management module mainly realizes functions such as: User information management, Domain keys private key management, Domain information management, Mailbox management, System configuration done by administrator.

Inter-domain authentication is used to authenticate the mail sender's domain. It is a combination of three authentications:

- Domain keys based on the encryption: It is done, if authentication is successful, then "source address authentication is successful" and the information that Domain keys authentication is successful are inserted into the mail' header, authentication is over.
- DNS blacklist based on the trust and prestige: It is employed to do prestige authentication, if sender's domain is in the blacklist, then it no longer continues to be authenticated, and mail's sending request is directly rejected, authentication is over
- SPF based on the path: It is done, if authentication is successful, then "source address authentication is successful" and the information that SPF authentication is successful are inserted into the mail's header, authentication is over

Intra-domain authentication module is used to verify user identity and his IP address, which mainly implemented by two modules[5]. They are username and password authentication module and IPv6 address authentication module. Intra-domain authentication does a username and password 2182 authentication for user. If user passes the username and password authentication. Then it does an IPv6 address authentication.

This paper tries to provide a secure email system with backbone support of true IPv6 address access whilst researching email source address authentication mechanism. Inter-domain authentication is inculcated as a combination of path-based, encryption-based and trust-based authentication, whereas intra-domain authentication employs the method of binding mail address with IP address. By running and testing on the CERNET2, this system is found to be feasible and effective.

E. Building a Model Using Clustering

A new approach to UFL using a simple, fast yet effective training by using Fuzzy ART is brought to light in this paper. In this paper, the approach works to decrease the gap in numerical and categorical features and work under the clustering domain which happens to be an unsupervised learning problem[4]. The ability of ART is seen in data fusion by mapping multi-modal features in an incremental manner.

Let X={x1,…xN}be a set of N samples in the given dataset, where xi=[x1,i,…xd,i]. T is a sample belonging to d-dimensional space Rd. Fuzzy ART contains two layers of neurons: the input layer F1 and the clustering representation layer F2.

Contrary to ART 1, where there are bottom-up and top-down weight vectors, Fuzzy ART has only one weight vector wj for each category j, which is initialized to wj,1=wj,2=?=1 when the category is uncommitted.

Before putting the sample into ART, it has to be normalized to [0, 1] and enhanced with complement coding to prevent category proliferation problems[4]. The clusters are formed in layer F2. When an input x is presented to layer F1, both committed neurons and one uncommitted neuron compete in a winner take all manner to select the one with maximum activation considering to the formula below:

$$Tj=|x?wj|a+|wj| \qquad (1)$$

where 'a' is a small real number used to break the tie while the fuzzy AND operator '?' is defined by:

$$(x?y)i=min(xi,yi) \qquad (2)$$

and where the norm |.| is defined by:

$$|x|=Sdi=1|xi| \qquad (3)$$

The winning neuron J, which is argmaxjTj gets activated. If neuron J passes the vigilance criterion that is:

$$\rho<|x,wJ||x| \qquad (4)$$

then the weight adaption occurs:

$$wJ(new)=(x,wJ(old)+(1-x)wJ(old) \quad (5)\ .$$

The benefit of UFL using Fuzzy ART is the dynamics as it creates multiple prototypes used for learning features only by increasing the vigilance threshold value.

This paper introduced us to the UFLA model and its advantages over the other models. This model can learn the features even if the amount of data is minimal and the learned features can remove the distinction in treating categorical and numeric features and hence results to a better clustering model.

III. COMPARISON OF RESULTS

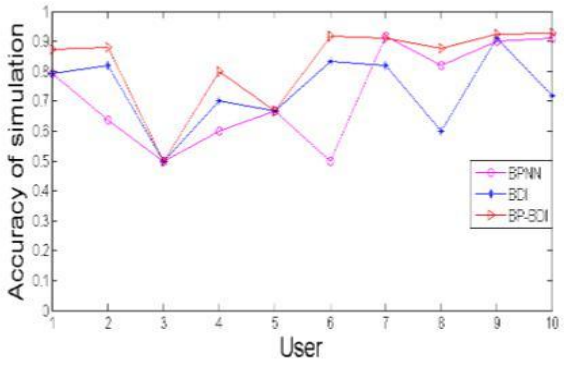

Fig 2: Accuracy Of Users Email Behavior Simulation in BPNN , BDI , BP-BDI Models [3]

The BP - BDI is a model used to track user's email behavior in a dynamic environment by constructing a learning model with BP neural network , the above graph shows an accuracy of 80% while simulating the email user's behavior.

TABLE 1: Comparision Between TS , WF , D(.75) Extraction Methods [2]

Compared methods (m_1 vs. m_2)	Relevance (%)	
	m_1	m_2
WF vs. TS	**54**	46
D(.75) vs. WF	**58**	42
D(.75) vs. TS	**70**	30

Comparative relevance scores of document result lists using single queries obtained from three keyword extraction methods on the elea corpus. The following ranking can be inferred: d(.75)>wf> ts.

This table shows result of the models presented (UFL Fuzzy ART, k-prototype, k-medoids, Fuzzy ART) on different datasets that include mixed type features. Clearly, UFL has superior performance in clustering compared to the other with accurate rate of above 86%. The reason for this better performance is that the UFL features have removed the gap between the categorical and the numerical features like in the other models.

Table 2 gives a clear comparison among the various clustering models that are experimented on various datasets.

TABLE 2: Clustering Results For Multi Feature Variable Datasets [4]

	UFL Fuzzy ART		K-prototype		K-medoids		Fuzzy ART	
	Acc	Rand	Acc	Rand	Acc	Rand	Acc	Rand
Heart disease	81.5	69.7	80.0	63.8	76.5	61.1	46.6	50.4
Teaching assistant	52.2	59.1	40.2	55.3	46.1	53.1	44.2	50.9
Credit assignment	86.0	75.0	79	67.1	75.0	62.5	70	50.0

CONCLUSION

We have seen how the extent of security of a mail server has remained narrow and restricted to classic methods of security and authentication. Spam control is one such aspect of a Mail Server that has shown stronger intent and efforts towards detecting a abused user over a small number of mail instances[1].Inheriting from the ideas of spam detection to detect a abused user through his logs[2], we can make use of models like BP-BDI [3] and Ipv6 addressing [5] to draw a line between a abused and a legitimate user , and feed such processed data to a unsupervised self learning machine learning model that can train itself. A stronger, safer, adaptive, and a much independent security model can be established for a mail server that would secure it beyond the already existing password based authentication.

REFERENCES

[1] Schäfer, C. (2017, April). Detection of compromised email accounts used for spamming in correlation with origin-destination delivery notification extracted from metadata. In *2017 5th International Symposium on Digital Forensic and Security (ISDFS)* (pp. 1-6). IEEE.

[2] Habibi, M., & Popescu-Belis, A. (2015). Keyword extraction and clustering for document recommendation in conversations. *IEEE/ACM Transactions on audio, speech, and language processing*, *23*(4), 746-759.

Sheng, Y., Rong, J., & Xiang, W. (2015, September). Simulation of the Users' Email Behavior Based on BP-BDI Model. In *2015 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery* (pp. 16-22). IEEE.

[4] Lam, D., Wei, M., & Wunsch, D. (2015). Clustering data of mixed categorical and numerical type with unsupervised feature learning. *IEEE Access*, *3*, 1605-1613.

[5] Meng, S., & Xingwei, W. (2015, August). Secure email system based on true IPv6 address access. In *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)* (pp. 2180-2184). IEEE.

[6] Patra, S., Naveen, N. C., & Prabhakar, O. (2016, May). An automated approach for mitigating server security issues. In *2016 IEEE International Conference on Recent*

[7] Yuan, H., Maple, C., Chen, C., & Watson, T. (2018). Cross-device tracking through identification of user typing behaviours. *Electronics Letters*, *54*(15), 957-959.

[8] Khanji, S., Jabir, R., Ahmad, L., Alfandi, O., & Said, H. (2016, April). Evaluation of Linux SMTP server security aspects—A case study. In *2016 7th International Conference on Information and Communication Systems (ICICS)* (pp. 252-257). IEEE

[9] Ganiger, S., & Rajashekharaiah, K. M. M. (2018, June). Comparative Study on Keyword Extraction Algorithms for Single Extractive Document. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1284-1287). IEEE.

[10] Huo, B., Long, Y., & Wu, J. (2017, December). A Secure Web Email System Based on IBC. In *2017 13th International Conference on Computational Intelligence and Security (CIS)* (pp. 578-581). IEEE.