IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AI BASED SPAM SPOILER FOR PUBLICAND PRIVATE E-MAIL SERVICES

S.S.Vasantha Raja,[1]

A.Vijayanarayanan [2] B.Priya [3] C.Kalaiarasi[4] R.Savithiri[5]AssistantProfessor[1,2,3,4,5]

Department of Computer Science and Engineering

PERIINSTITUTEOFTECHNOLOGY

Abstract-

In recent years, cyber security incidents have occurred frequently. In most of these incidents, attackers have used different type of spam email as a knock-on to successfully invade government systems, well-known companies, and websites of politicians and social organizations in many countries. The detection of spam mail from big email data has been paid public attention. However, the camouflage technology of spam mail is becoming more and more complex, and the existing detection methods are unable to confront with theincreasingly complex deception methods and the growing number of email. In this project, we proposed to design a novel efficient approach named Spam Spoiler for big e-mail data classification into four different classes: Normal, Fraudulent, Harassment, and Suspicious E-mails by using LSTM based GRU. The new method includes two important stages, sample expansion stage and testing stage under sufficient samples. This project The LSTM based GRU efficiently captures meaningful information from E-mails that can be used for forensic analysis as evidence. Experimental results revealed that Spam Spoiler performed better than existing.ML algorithms and achieved a classification accuracy of 98% using the novel technique of LSTM with recurrent gradient units. As different types of topics are discussed in E-mail content analysis. Spam Spoiler effectively outperforms existing methods while keeping the classification process robust and reliable.

1. INTRODUCTION

Email stands for Electronic Mail. It is a method to sends messages from 2one computer to another computer through the internet. It is mostly used in business, education, technical communication, document interactions. It allows communicatingwith people all over the world without bothering them Email messages are conveyed through email servers; it usesmultiple protocols within the TCP/IP suite. For example, SMTP is a protocol, stands for simple mail transfer protocol and used to send messages whereas other protocols IMAP or POP are used to

retrieve messages from a mail server. If you want to login to your mail account, you just need to enter a valid email address, password, and the mail servers used to send and receive messages. Although most of the webmail servers automatically configure your mail account, therefore, you only required to enter your email address and password. However, you may need to manually configure each account if you use an email client like Microsoft Outlook or Apple Mail. In addition, to enter the email address and password, you may also need to enter incoming and outgoing mail servers and the correct port numbers for each one.

Fig 1 : Email Email messages include three components, which are as follows:

- Message envelope: It depicts the email's electronicformat.
- Message header: It contains email subject line andsender/recipient information.
- Message body: It comprises images, text, and other fileattachments.

2 PROBLEMS IDENTIFIED

Many people rely on the Internet for many of their professional, social and personal activities. But there are also people whoattempt to damage our Internet-connected computers, violate our privacy and render inoperable Internet services.Email is a universal service used by over a billion people worldwide. As one of the most popular services, email has becomea major vulnerability to users and organizations. The statistics are astounding. Email remains the number one threat vector for data breaches, the point of entry for ninety-four percent of breaches. There is an attack every 39 seconds. Over 30% of phishing messages get opened, and 12% of users click on malicious links. As cybercrime becomes more advanced and bypasses the legacy controls put in place to defend against it, security must become more advanced too.

Fig 2: PhishingBelow are some of the most common types of Attacks:

1. Phishing:

Phishing is a form of fraud. Cyber criminals use email, instant messaging, or other social media to try to gather information such as login credentials by masquerading as a reputable person. Phishing occurs when a malicious party sends a fraudulent email disguised as being from an authorized, trusted source. The message intent is to trick the recipient into installing malware on his or her device or into sharing personal or financial information. Spear phishing is a highly targeted phishing attack. While phishing and spear-phishing both use emails to reach the victims, spear-phishing sends customized emails to aspecific person. The criminal researches the target’s interests before sending the email.

2. Vishing:

Vishing is phishing using voice communication technology. Criminals can spoof calls from authorized sources usingvoice-over IP technology. Victims may also receive a recorded message that appears authorized. Criminals wantto obtain credit card numbers or other information to steal the victim's identity. Vishing takes advantage of the fact thatpeople trust the telephone network.

3. Smishing:

Smishing is phishing using text messaging on mobile phones. Criminals impersonate a legitimate source in anattempt to gain the trust of the victim. For example, a smishing attack might send the victim a website link. Whenthe victim visits the website, malware is installed on the mobile phone.

4. Whaling:

Whaling is a phishing attack that targets high profile targets within an organization such as seniorexecutives. Additionaltargets include politicians or celebrities

5. Pharming:

Pharming is the impersonation of an authorized website in an effort to deceive users into entering their credentialsPharming misdirects users to a fake website that appears to be official. Victims then enter their personal information thinkingthat they are connected to a legitimate site.

6. Spyware:

Spyware is software that enables a criminal to obtain information about a user's computer activities. Spyware often includes activity trackers, keystroke collection, and data capture. In an attempt to overcome security measures, spyware often modifies security settings. Spyware often bundles itself with legit imate software or with Trojan horses. Many shareware websites are full of spyware.

7. Scareware:

Scareware persuades the user to take a specific action based on fear. Scareware forges pop-up windows that resemble operating systemdialogue windows. These windows convey forged messages stating that the system is at risk or needs the execution of a specific programto return to normal operation. In reality, no problems exist, and if the user agrees and allows the mentioned program to execute, malwareinfects his orher system.

8. Adware:

Adware typically displays annoying pop-ups to generate revenue for its authors. The malware may analyse user interests by tracking thewebsites visited. It can then send pop-up advertising relevant to those sites. Some versions of software automatically install Adware.

9. Spam:

Spam (also known as junk mail) is unsolicited email. In most cases, spam is a method of advertising. However, spam can send harmfullinks, malware, or deceptive content. The end goal is to obtain sensitive information such as a social security number or bank accountinformation. Most spam comes from multiple computers on networks infected by a virus or worm. These compromised computers send out as much bulk email as possible.

10. E-Mail Bombing

An email bombing is an attack on your inbox that involves sending massive amounts of messages to your address. Sometimes thesemessages are complete gibberish, but more often they'll be confirmation emails for newsletters

and subscriptions. In the lattercase, the attacker uses a script to search the internet for forums and newsletters and then signs up for an account with your email address.Each will send you a confirmation email asking to confirm your address. This process repeats across as many unprotected sites as thescript can find

3 Existing System

3.1.1. Content-Based Filtering Technique

Algorithms analyse words, the occurrence of words, and the distribution of words and phrases insidethe content of e-mailsand segregate them into spam non-spam categories.

Fig 3: Content-Based Filtering Technique

3.1.2 Case Base Spam Filtering Method

Algorithms trained on well-annotated spam/non-spammarked emails try to classify the incoming mailsinto twocategories.

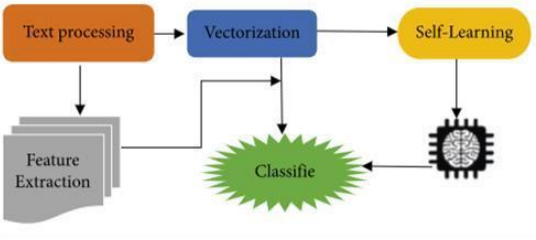

Fig 4 : Case Base Spam Filtering Method

Heuristic or Rule-Based Spam Filtering Technique

Algorithms use pre-defined rules in the form of a regular expression to give a score to the messages present in the e-mails. Based on the scores generated, they segregate emails into spam non-spam categories.

4. The Previous Likeness Based Spam

- Filtering Technique

Algorithms extract the incoming mails' features and create a multi-dimensional space vector and draw points for every newinstance.

- Adaptive Spam Filtering Technique

Algorithms classify the incoming mails in various groups and, based on the comparison scores of every group

with the defined set of groups, spam and non-spam emails got segregated.

- Machine learning classifiers

The machine learning models are selected based on their group, diversity and acceptance in the machine learning community.SVM, Naive Bayes (NB) and DT are from three different groups of classifiers.

5. Proposed System

The proposed approach comprises data collection, pre-processing,feature extraction, parameter tuning, and classification using the LSTM-GRU model. In this project, E-mail datasets are divided into normal, harassing, suspicious, and fraudulent classes. The E-mail is divided into word levels of the E-mail body, and the embedding layer is applied to train and obtain the sequence of vectors.

- **LSTM and GRU**

In Deep learning, Long-Term Short-Term Memory Networks and Gated Recurrent Units, LSTM and GRUsfor short.

- **LSTM – Long Short-Term Memory**

LSTMS are a special kind of RNN which is capable of learning long- term dependencies. LSTMs are designed to dodge long-term dependency problem as they are capable of remembering information for longer periods of time. Long short-term memory (LSTM) units (or blocks) are a building unit for layers of a recurrent neural network (RNN). A RNN composed of LSTM units is often called an LSTM network. A common LSTM unit is composed of a cell, an input gate,an output gate and a forget gate. The cell is responsible for "remembering" values over arbitrary time intervals; hence the word "memory" in LSTM. Each of the three gates can be thought of as a "conventional" artificial neuron, as in a multi-layer (or feedforward) neural network: that is, they compute an activation (using an activationfunction) of a weighted sum. Intuitively, they can be thought as regulators of the flow of values that goes through the connections of the LSTM; hence the denotation "gate". There are connections between these gates and the cell. The expression long short-term refers to the fact that LSTM is a model for the short-term memory which can lastfor a long period of time. An LSTM is well-suited to classify, process and predict time series given time lags of unknown size and duration between important events. LSTMs were developed to deal with the exploding and vanishing gradient problem when training traditional RNNs.

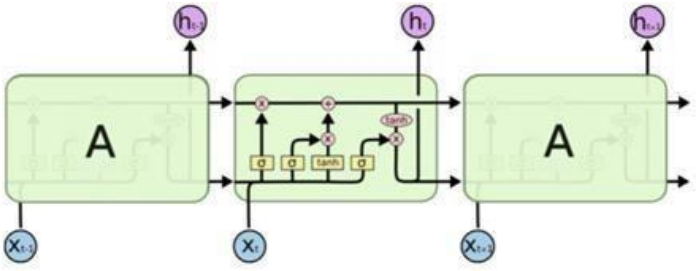

Fig 5: LSTM – Long Short-Term Memory

The popularity of LSTM is due to the Getting mechanism involved with each LSTM cell. In a normalRNN cell, the inputat the time stamp and hidden state from the previous time step is passed through the activation layer to obtain a new state. Whereas in LSTM the process is slightly complex, as you can see in the above architecture at each time it takes input from three different states like the current input state, the short-term memory from the

previous cell and lastly the long-term memory.

These cells use the gates to regulate the information to be kept or discarded at loop operation before passing on the long term and short-term information to the next cell. We can imagine these gates as Filters that remove unwanted selected and irrelevant information. There are a total of three gates that LSTM uses as Input Gate, Forget Gate, and Output Gate.

- Input Gate

The input gate decides what information will be stored in long term memory. It only works with the information from the current input and short-term memory from the previous step. At this gate, it filters out the information from variables that are not useful.

- Forget Gate

The forget decides which information from long term memory be kept or discarded and this is done by multiplying the incoming long-term memory by a forget vector generated by the current input and incoming short memory.

- Output Gate

The output gate will take the current input, the previous short- term memory and newly computed long-term memory to produce new short-term memory which will be passed on to the cell in the next time step. The output of the current time step can also be drawn from this hidden state.

- **GRU – Gated Recurrent Unit**

Gated recurrent unit (GRU) was introduced by Cho, et al. in 2014 to solve the vanishing gradient problem faced by standardrecurrent neural networks (RNN). GRU shares many properties of long short-term memory (LSTM). Both algorithms use a gating mechanism to control the memorization process. A gated recurrent unit (GRU) is a gating mechanism in recurrent neural networks (RNN) similar to a long short-term memory (LSTM) unit but without an output gate. GRU's try to solve the vanishing gradient problem that can come with standard recurrent neural networks. A GRU can be considered a variation of the long short- term memory (LSTM) unit because both have a similar design and produce equal results in some cases. GRU's are able to solve the vanishing gradient problem by using an update gate and a reset gate. The update gate controls information that flows into memory, and the reset gate controls the information that flows out of memory. The update gate and reset gate are two vectors that decide which information will get passed on to the output. They can be trained to keep information from the past or remove information that is irrelevant to the prediction. A GRU is a very useful mechanism for fixing the vanishing gradient problem in recurrent neural networks. The vanishing gradient problem occurs in machine learning when the gradient becomes vanishingly small, which prevents the weight from changing its value. They also have better performance than LSTM when dealing with smaller datasets.

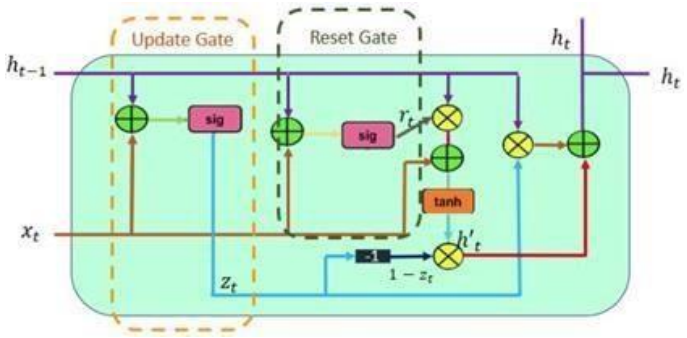

Fig 6: GRU – Gated Recurrent Unit

6. LSTM Based GRU Classification Model

LSTM comprises the classification of E-mails as Normal, Harassing, Suspicious, and Fraudulent. The LSTM and GRU are both based on the gated network architecture, due to which we combined the GRU and LSTM to utilize the gated architecture of both of them.The DL models' layered structure helps in learning without intervention in ML model implementation. Several libraries provide an in-depth learning implementation structure. We split the data into three training,validation, and testing sets with a 65 V 10 V 25 ratio. We extracted the features from textual data of E-mail using the word Embedding

technique. We encode the target values using the one-hot encoding technique into 4-distinct classes. We pass all pre-processed data to the novel architecture of LSTM layers variants for the perfect classification of E- mails. We use the LSTM layers with different GRU and Convo1D layer variants to transform the input textual data into an efficient E-mail classification system.

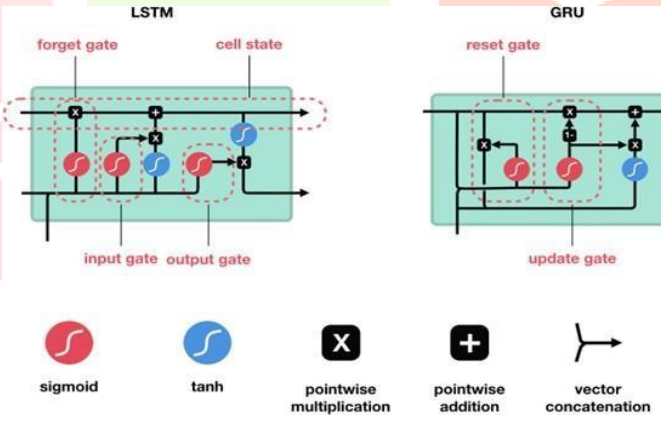

Fig 7 : LSTM Based GRU Classification Model

Textual data needs special attention when feature extraction comes in the proposed methodology. Different feature extractionmethods need to be implemented when solving the Natural language processing problem using DL. The main point is toconvert the textual data into real-valued vectors. There is a unique name for the vector in natural language processing, ``embedding vector". There are multiple ways to generate the embedding vector from the textual data, but famous methods are GLOVE and Word2Vec techniques. Embedding vector dimensions are essential to get all the features extracted from the data. Let us suppose if we have 8 samples of textual data. The data have two distinct classes. Each sample has a maximum of five tokens in it. The vocabulary size will be the unique words in 8 samples, and the vocabulary size needs to be higher than the available unique tokens in the

dataset to avoid collisions with a hash function. In this case, the dimensions of the embedding vector will be 4 x 8. In the case of the classification problem of NLP, we need to encode the target values using the one-hot encoding method.

CONCLUSION

With the growing trend of cybercrime and accidents resulting from vulnerabilities, proactive monitoring and post-incident

analysis of email data is crucial for organizations. Cybercrimes like hacking, spoofing, phishing, E-mail bombing, whaling, and spamming are being performed through E-mails. The existing email classification approaches lead towards irrelevant E-mails and/or loss of valuable information. Keeping in sight these limitations, we designed a novel efficient approach named E- Mail Sink AI for E-mail classification into fourdifferent classes:

Normal, Fraudulent, Threatening, and Suspicious E-mails by using LSTM based GRU that not only deals with short sequences as well long dependencies of 1000C characters. We evaluated the proposed E-Mail Sink AI model using evaluation metrics such as precision, recall, accuracy, and f-score. Experimental results revealed that E-Mail Sink AI performed better than existing ML algorithms and achieved a classification accuracy of 95% using the novel technique of LSTM with recurrent gradient units.

FUTURE ENHANCEMENT

For now, we are considering e-mail classes such as normal, harassment, fraudulent, and suspicious; however, many otherclasses can be added to this work in the presence of the massive amount of e-mail data.

REFERENCE

1. S. Sinha, I. Ghosh, and S. C. Satapathy, ``A study for ANN model for spamclassification,"inIntelligentDataEngineering and Analytics. Singapore: Springer, 2021,pp. 331-343.

2. Q. Li, M. Cheng, J. Wang, and B. Sun, ``LSTM basedphishing detection for big email data," IEEE Trans. BigData,earlyaccess,Mar.12,2020,doi: 10.11.09/TBDATA.2020.2978915.

3. Salman Ali, Adnan Ashraf, Saad Bin Qaisar "A WIRELESS SENSOR NETWORK MONITORING PLATFORM FOR OIL AND GAS PIPELINES", in IEEE Systems Journal, DOI: 10.1109/JSYST.2016.2597171.[2018]

4. Shinian Wang, Yi Chai, ZhiminYang"GAS SOURCE LOCALIZATION BASED ON MAXIMUM LIKELIHOOD WITH ARBITRARY DEPLOYMENT WSN", in IEEE Xplore DOI: 10.1109/ChiCC.2014.6896647.[2013]

5. Guijie Wang, A. Heidary "SMART TEMPERATURE SENSOR", in IEEE Journal, DOI:10.1016/B978-0-08-102055- 5.00003-6[2018] Fabian Chraim, Yusuf BufraErol, and Kris Pister "Wireless gas leakage detection and Junru Lin, Baohui Zhu, Peng Zeng, Wei Liang "MONITORING POWER TRANSMISSION LINES USING A WIRELESS SENSOR NETWORK", in IEEE Xplorehttps://doi.org/10.1002/wcm.2458 [2018] Ruping Liu, Liang He, Marijuana Cao "FLEXIBLE

TEMPERATURE SENSORS",in IEEE Systems Journals, DOI:10.1021/acsami.6b1456[2015]

6. Zhou Y.M, Yang X.L "WIRELESS SENSOR NETWORK", in IEEE Systems Journal, DOI: 10.4236/cn.2009.12015 [2021]

7. Wang and Roach "GASLEAKAGE MONITORING USING MOBILE NETWORK" inIEEE Systems Journal, DOI:10.1088/1742-596/1717/1/012068 License CC BY 3.0 JUNE[2022]

8. H.M. Salam, Daniel Cobos Munoz "A systems approach to assessing complexity in interventions: an effectiveness decay model for integrated community case management" in IEEE System Journal,doi:0.1080/16549716.2020.1794106 [2020]