International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Phishing Email Detection Using Region Based Convolutional Neural Network Model

Vishal Bharati
Computer Science and Engineering
BNCET college of Engineering
Lucknow, India
vishaldlw007@gmail.com

Chandan Kumar
Computer Science and Engineering
BNCET college of Engineering
Lucknow, India
chandangupta9292@gmail.com

Rashmi Verma
Computer Science and Engineering
GOEL Institute of Engineering and Technology,
Lucknow, India
rashmi7897@gmail.com

Prashant Maharishi
Computer Science and Engineering
IET Khandari
Agra, India
prashantbsacet1@gmail.com

***Abstract*— Email which contains suspicious links or malicious files is one of the considerable threats in digital world nowadays. Existing approaches of phishing emails detection are unceasingly being renovated nonetheless the outcomes are not impressive. In our proposed work, the structure of the email is scrutinized and then fed on a recurrent convolution neural network (RCNN). After that RCNN model start processing email header and body of the email for detecting the suspicious or malicious files. If the suspicious links, file extensions are detected then it will automatically transferred to the spam folder. The classification of the spam and non-spam email is done by using the neural network classifier. The phishing emails contains the suspicious links when user click on the link the redirection take place. This way the phishing emails attack take place and due to which financial as well as social loss take place.**

***Keywords*— Support Vector Machine, LSTM, RCNN, ANN, MIME**

I. INTRODUCTION

The instantaneous expansion of digital world has vastly transformed online user‟s involvement instead security threats are also increasing with rapid rate such as different type of novel attacks will take place every day. The current situation signifies that novel attacks will not cause damage only to victims systems rather than that also emphases to steal victims money [1][2]. Along with different type of attacks, phishing email attack is obvious one and it is also illegal action that using social networking technology and platforms for gathering target‟s identity data as well as account information. The phishing email will be identified using some concept such as email filtering[3][4]. Email filtering such as the blacklisting mechanism, visual similarity, heuristic and machine learning methods. These method used only for detecting the phishing email at some extent. They are helpful but fails for the email‟s which contain the suspicious IP based urls and non-matching urls.

The emails which have the urls containing the IP address are more harmful because if the user click on that it will redirect to some other web page which are asking the confidential data from users which causes financial as well as the identity loss take place [5]. Some of the techniques are proposed using machine learning model in which the different classifiers are used for the classification of the emails on the basis of some extracted features. Some features which are used for classifications are disparities between the href,

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

presence of links “click” ,“here" and “click here", presence of JavaScript code, number of dots in domain name, html codes. Other features are number of links, number of linked to domain, urls containing IP address and from body match domain check and also the update, confirm, verify account, restrict, suspend. In paper [6] method is proposed in which based on the concept of bag of words. In these techniques, all words are extracted from the email and the frequency of the particular words is taken as the features for doing the classification of spam email. The APWG provided the data according to which the number of phishing emails is increased from 87390 in 2019 to 266342 in 2020. And the type of phishing email which are totally different kind are approximately 100000 from December 2020 [7][8]. Increase in momentum of phishing email attack shows that loss in the developed countries like Australia, UK and United State are approximately 600 million dollars per year.

The phishing email threat is increasing rapidly due to which awareness about phishing email is more important for user. They should not open the suspicious links present in emails and do not share confidential data to anyone. With the awareness, the accurate method is also required for detection of phishing emails and preventing method which avoid and stop loss occurs due to phishing emails [9][10].Remaining of the paper describe as follows : section II involves the related work, section III involves the method and material, section IV contains result and discussion and section V contains conclusion.

II. RELATED WORK

The Phishing email is one the subset of the spam attack and the detection of this attack is totally depends upon the classifier used for the classification in the model. The feature which are extracted and used for the training and testing model is important. In paper [12] method proposed in which 10 features are used for representing the phishing email after that with the help of the random forest in place of classifier for creating number of decision tree [6]. The decision tree helps to detect the emails as a phishing emails if the same feature found in any email. The accuracy of this model is close to 96% with false negative rate of 4% and false positive rate of 0.1%. Another technique is proposed which is based on the concept of multi-layer classifier in which one of the three tier classifier technique is deployed for detecting phishing emails [11]. In this, if first two tier classifier does not do good classification then third tier classifier classification taken for making decision that the particular email belongs to phishing email or to legitimate emails. The average accuracy of this model is more than 96% but this consumes more time as well as more memory[12][13]. One approach of phishing email is based on clustering method. This method totally depends on the sharing method between the supervised classification algorithm and unsupervised clustering method [14]. After that training of model with consensus clustering take place with the help of this technique the classification take place easily with improved accuracy then the k-means classification algorithm. One framework of phishing email detection is deployed which uses the concept of the fuzzy logic [15][16]. In this, 21 features are extracted at the stage of the pre-processing and then at next stage the feature reduction take place .the features are remains are used for making the basis rule that these features are used to categorize the email either as a legitimate email or phishing email, the accuracy of this model is approx. 97%[17][8]. One of the model is proposed which uses the concept of the logistic regression as a classifier with the help of which the classification of email take place i.e. the email is either legitimate or non- legitimate. The model is trained using the noisy dataset and with 10 features used in which URLs and disparities of href are included. The accuracy of the model is 95% with false positive rate of 0.03% [18][19]. The Phishing email detection model is proposed in which the support vector machine is used for the classification. The KDDCup99 dataset is used for training and testing the model.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

The model achieve the accuracy up to 87%[7][20].Other proposed model is uses the various features which are extracted from the URLs present in the emails. The URLs are scrutinized and compare with the legal form of urls if anything found as a suspicious then that email directly transferred to spam folder[7][8].

Table.1 Phishing email detection techniques

References	Classifiers Used	Accuracy
[21]	Random Forest	0.956
[22]	SVM + J48	0.937
[5]	Support Vector Machine	0.795
[4]	Random forest and Decision Tree	0.899
[15]	C5.0	0.957
[2]	Bayes Net	0.9711
[23]	Naïve Bayes ,SVM and Logistic Regression,	0.981
Our methodology	Region Based Convolutional Neural Networks	0.979

III. METHOD AND MATERIALS

The proposed method involves the various stages for detecting the email; either it is spam email or non-spam email. The first stage is input data set and at the stage of html parsing the parsing of the html documents takes place. The next stage involves the feature extraction. At this stage the features are extraction take place where some features are extracted with the help of which further analysis and comparison take place with the extracted feature of the particular email. The model is trained with balanced as well as unbalanced data set due to which the probability of the feature recognition is closer to value one which signifies the result become more accurate. After features are extracted the convolution neural network computation of features take place on the basis of which the classification of email take place i.e. the decision are made either the email is spam or non-spam email. The particular email header and body of the email is analyzed at character as well as at word level, if at any stage the email is matched with the features of the spam email either it is with suspicious links or suspicious extensions the immediately it will detected as a spam email and transferred to spam folder. If the spam features of email is not matched with particular email then it will non-spam email and transferred to the legitimate email. The model is trained with balanced as well as unbalanced data set due to which the probability of the feature recognition is closer to value one which signifies the result become more accurate. The output of this stage is transferred to neural network classifier in this the R-CNN is used as a classifier. In R-CNN the email is inputted and split into different segments. Each segment then used for the extraction of features which are relevant in nature. The particular email header and body of the email is analyzed at character as well as at word level ,if at any stage the email is matched with the features of the spam email either it is with suspicious links or suspicious extensions the immediately it will detected as a spam email and transferred to spam folder. If the spam features of email are not matched with particular email then it will legitimate email and transferred to the legitimate email folder.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

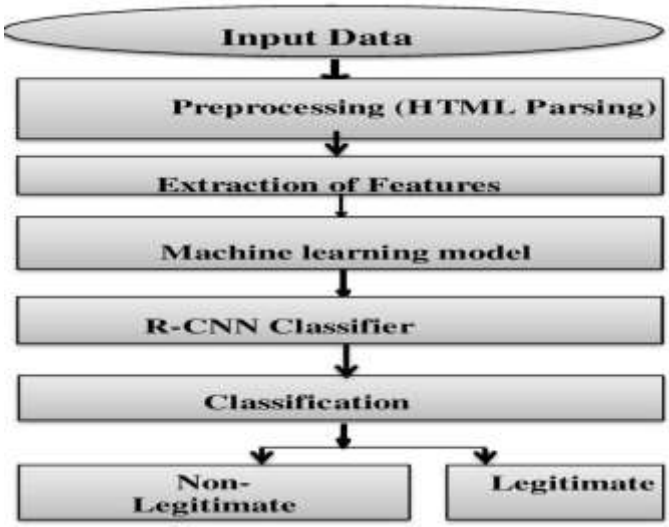

Fig.1 Model of Phishing email detection

IV. Result and Discussion

The testing set made from the dataset into the propositioned model assimilate and add a series of pointers, for evaluating functioning of model pointers are operated precisely. Assume True Negative is a way for indicating the number of authentic email that has been categorized as genuine. Similarly, False Positive is a method of expressing the number of genuine emails that are misclassified as illegal, False Negative is a method of expressing the number of phishing emails that are misclassified as legal, and True Positive is a method of expressing the number of non-legitimate emails that are misclassified as phishing email.

The outcome is compared to the output obtained by using the entire email. In contrast to the identical non-appropriate classified email, the revealing output using simply the email body and email header. However, the issue is that when the entire email is used for processing, the CNN model makes an inaccurate prediction. The email mean that the email having the body and header. The result is dictated by the body of email because body of the email generally contains suspicious contents and files if processing the complete email take place. The email's header is filled with phishing emails that have been verified as phishing emails, yet the email's body is contradictory. Because the final output of the email is the email body, the CNN model is not suitable for accurately detecting phishing emails. The believe in those emails that the CNN model missed and that those emails are highly disguised phished email in which the email body has a huge amount of content with genuine emails body i.e. the header of email still has differences) .The situation in question is one in which the body of the email has a low weight and the header of the email has a high weight, resulting in a more accurate detection of phishing emails. The model is accurately recognized the misclassified emails when using mechanism of the attention within the email header and the body of the email for giving different weightage dynamically, according to different number of experiments, the attention mechanism assigns the email header a superior weight than the email body.

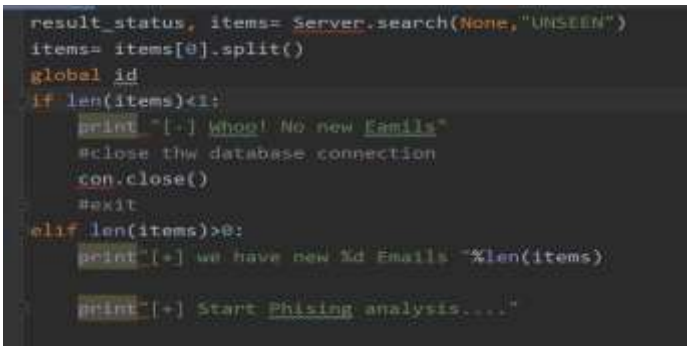

Fig.2 New emails downloading list

Fig.3 Non-Legitimate emails moving into spam

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

```
extension_blacklist = [
    'ade', 'adp', 'bat', 'chm', 'cmd', 'com', 'cpl', 'exe', 'hta', 'ins', 'isp',
    'jse', 'lib', 'mde', 'msc', 'msp', 'mst', 'pif', 'scr', 'sct', 'shb', 'sys',
    'vb', 'vbe', 'vbs', 'vxd', 'wsc', 'wsf', 'wsh', 'htm', 'htm']
```

Fig.4 Malicious file Extensions

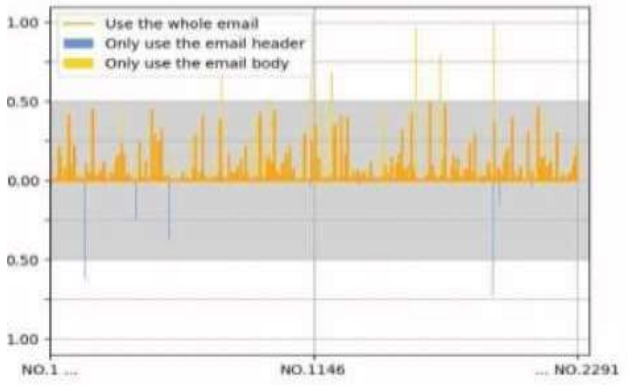

Fig.5 Predicted probability Vs emails

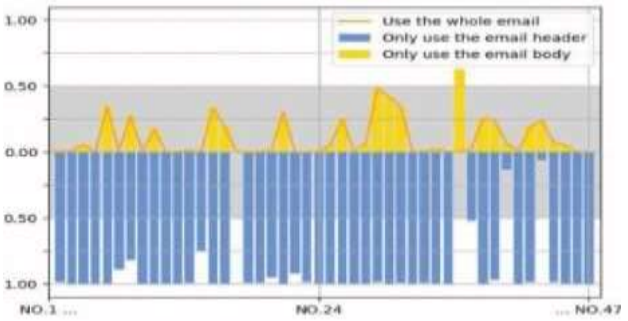

Fig.6 Predicted probability Vs emails

v. Conclusion

The phishing email detection is a process in which the suspicious links, malicious files or image are detected in a particular email which is threat the user‟s confidential data as well as financial threat also. In the model the concept of R-CNN is used with the help of which model does the comparison on word level and character level of the particular email. When the comparison takes place at character level as well as at word level the noise present in the output is minimal. The data set used for training and testing the model are unbalanced in nature due to which the experiment done closer to the real world scenario and model become more accurate. The result obtained from the model by using the concept of the R-CNN is more promising result. With the help of model the phishing email detection will be done by using only the body of email, the header of the email is either present or not the result will same.

References

[1] A. Herzberg and A. Gbara, “Trustbar: Protecting (even naive) web users from spoofing and phishing attacks,”*Comput. Sci. Dep. Bar Ilan …*, no. January, pp. 1–28, 2004.

[2] T. Meyer and B. Whateley, “SpamBayes: Effective open-source, Bayesian based, email classification system,” *Proc. First Conf. Email Anti-Spam*, vol. 98, pp. 1–8, 2004, [Online]. Available:

[3] S. S. M. Motiur Rahman, T. Islam, and M. I. Jabiullah, “PhishStack: Evaluation of Stacked Generalization in Phishing URLs Detection,” *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 2410–2418, 2020, doi: 10.1016/j.procs.2020.03.294.

[4] A. A. Akinyelu and A. O. Adewumi, “Classification of phishing email using random forest machine learning technique,” *J. Appl. Math.*, vol. 2014, 2014, doi: 10.1155/2014/425731.

[5] N. B. Harikrishnan, R. Vinayakumar, and K. P. Soman, “A machine learning approach towards phishing email detection CEN-Security@IWSPA 2018,” *CEUR Workshop Proc.*, vol. 2124, no. March 2020, pp. 21–28, 2018.

[6] Y. Fang, C. Zhang, C. Huang, L. Liu, and Y. Yang, “Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism,” *IEEE Access*, vol. 7, pp. 56329–56340, 2019, doi: 10.1109/ACCESS.2019.2913705.

[7] P. Lawson, C. J. Pearson, A. Crowson, and C. B. Mayhorn, “Email phishing and signal detection: How persuasion principles and personality influence response patterns and accuracy,” *Appl. Ergon.*, vol. 86, no. March, p. 103084, 2020, doi: 10.1016/j.apergo.2020.103084.

[8] E. Zhu, Y. Ju, Z. Chen, F. Liu, and X. Fang, “DTOF-ANN: An Artificial Neural Network phishing detection model based on Decision Tree and Optimal Features,” *Appl. Soft Comput. J.*, vol. 95, p. 106505, 2020, doi: 10.1016/j.asoc.2020.106505.

[9] A. A. Orunsolu, A. S. Sodiya, and A. T. Akinwale, “A predictive model for phishing detection,” *J. King Saud Univ. - Comput. Inf. Sci.*, no. xxxx, 2020, doi: 10.1016/j.jksuci.2019.12.005.

[10] N. A. Azeez, S. Misra, I. A. Margaret, L. Fernandez-Sanz, and S. M. Abdulhamid, “Adopting automated whitelist approach for detecting phishing attacks,” *Comput. Secur.*, vol. 108, p. 102328, 2021, doi: 10.1016/j.cose.2021.102328.

[11] K. Selvan and M. Vanitha, “Detection of Phishing Web Pages Based on Features Vector and Prevention using Multi Layered Authentication,” vol. 119, no. 15, pp. 565–573, 2018.

[12] A. Bergholz, G. Paaß, F. Reichartz, S. Strobel, and J. H. Chang, “Improved phishing detection using model- based features,” *5th Conf. Email Anti-Spam, CEAS 2008*, no. January 2008, 2008.

[13] K. A. Molinaro and M. L. Bolton, “Evaluating the applicability of the double system lens model to the analysis of phishing email judgments,”

Comput. Secur., vol. 77, pp. 128–137, 2018, doi: 10.1016/j.cose.2018.03.012.

[14] D. Ranganayakulu and C. C., "Detecting Malicious URLs in E-mail – An Implementation," *AASRI Procedia*, vol. 4, pp. 125–131, 2013, doi: 10.1016/j.aasri.2013.10.020.

[15] O. Doctor, "Dynamic Evolving Neural Fuzzy Framework for Phishing E-Mail Detection," no. March, pp. 3960– 3964, 2013.

[16] F. A. Aloul, "The Need for Effective Information Security Awareness," *J. Adv. Inf. Technol.*, vol. 3, no. 3, pp. 176–183, 2012, doi: 10.4304/jait.3.3.176-183.

[17] A. Vazhayil, N. B. Harikrishnan, R. Vinayakumar, and K. P. Soman, "PED-ML: Phishing email detection using classical machine learning techniques CENSec@Amrita," *CEUR Workshop Proc.*, vol. 2124, pp. 69–76, 2018.

[18] S. Smadi, N. Aslam, and L. Zhang, "Detection of online phishing email using dynamic evolving neural network based on reinforcement learning," *Decis. Support Syst.*, vol. 107, pp. 88–102, 2018, doi: 10.1016/j.dss.2018.01.001

[19] M. Volkamer, K. Renaud, B. Reinheimer, and A. Kunz, "User experiences of TORPEDO: TOoltip-poweRed Phishing Email DetectiOn," *Comput. Secur.*, vol. 71, pp. 100–113, 2017, doi: 10.1016/j.cose.2017.02.004.

[20] L. Gallo, A. Maiello, A. Botta, and G. Ventre, "2 Years in the anti-phishing group of a large company," *Comput. Secur.*, vol. 105, p. 102259, 2021, doi: 10.1016/j.cose.2021.102259.

[21] E. G. Dada and S. B. Joseph, "Random Forests Machine Learning Technique for Email Spam Filtering," vol. 9, no. 1, pp. 29–36, 2018.

[22] H. Mi, Z. Wang, and A. Ittycheriah, "Supervised attentions for neural machine translation," *EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc.*, no. 4, pp. 2283–2288, 2016, doi: 10.18653/v1/d16-1249.

[23] J. Drew and T. Moore, "Automatic identification of replicated criminal websites using combined clustering,"*Proc. - IEEE Symp. Secur. Priv.*, vol. 2014-Janua, pp. 116–123, 2014, doi: 10.1109/SPW.2014.26.

[24] Bashir, T. Agbata B.C, E. Ogala, and W. Obeng-Denteh, "The Fuzzy Experiment Approach for Detection and Prevention of Phishing attacks in online Domain," *East African Sch. J. Eng. Comput. Sci.*, vol. 3, no. 10, pp. 205–215, 2020, doi: 10.36349/easjecs.2020.v03i10.001.

[25] T. H. Nguyen and R. Grishman, "Modeling skip-grams for event detection with convolutional neural networks," *EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc.*, pp. 886–891, 2016, doi: 10.18653/v1/d16- 1085.

[26] M. Nguyen, T. Nguyen, and T. H. Nguyen, "A deep learning model with hierarchical LSTMs and supervised attention for anti-phishing," *CEUR Workshop Proc.*, vol. 2124, no. Iwspa, pp. 29–38, 2018.

[27] X. Glorot, A. Bordes, and Y. Bengio, "Domain adaptation for large-scale sentiment classification: A deep learning approach," *Proc. 28th Int. Conf. Mach. Learn. ICML 2011*, no. 1, pp. 513–520, 2011.

[28] P. Sabeeha, MD; Karimullah, SK; Babu, "Detection of Malicious URLs by Correlating the Chains of Redirection in an Online Social Network (Twitter)," *Int. J. Res. Stud. Comput. Sci. Eng.*, vol. 1, no. 3, pp. 33–38, 2014.