Cyber Aid Station - Cyber Security Application

[1]Saurabh Adate, [2]Harsh Chaudhary, [3]Ashish Jaiswal, [4]Yogita Ganage

[1,2,3,4]UG Student, Information Technology, MCT Rajiv Gandhi Institute of Technology Mumbai, India. [1]saurabh.adate@gmail.com, [2]jaiswalashish405@gmail.com, [3]harsh3030.hc@gmail.com, [4]yogita.ganage@mctrgit.ac.in

Abstract: We can see that technology has touched many spheres of our lives in India. There is technology in business, in education, in socializing and retaining human relations, in purchasing, in agriculture, in banking, communication, and nearly each a part of our lives. This intrusion of technology has aided the add of these sections and has proved beneficial, and time and effort-saving. The only major part of our society that remains majorly devoid of this luxury is the Indian Police Department."[6]" Digitization in the Police department is the need of the hour. The conventional approach of visiting a police station for registering a police complaint and getting updates needs to get replaced with an internet process. Hence a cyber station system is being developed which will collect complainant's data through an application, sends the information over to the Police branch on their web portal, and on this manner the whole interplay takes place online, with record exchanges over the software and the web portal". Phishing costs Internet users billions of dollars per annum. It refers to luring techniques employed by identity thieves to fish for private information during a pond of unsuspecting internet users. . The methods used for detection of phishing websites based on lexical features, host properties, and page importance properties. We consider various data mining algorithms for evaluation of the features so as to get a far better understanding of the structure of URLs that spread phishing. The fine-tuned parameters are useful in selecting the apt machine learning algorithm for separating phishing sites from legitimate sites.

Keywords: FIR Registration, Chatbot, Phishing URL Detection, Random Forest Classifier.

I. INTRODUCTION

In today's world as people start using a smartphone, and almost 90% of people in our country connected through the internet. Some people afraid of the police to register a complaint and thousands of people got attacked by hackers. Our aim to protect peoples from cyberattacks as well as to keep minimal distance between citizens and police. Our projects have three main aspects to protect citizens are as follows: -

1. F.I.R. Registration
2. Chatbot (For any queries)
3. Phishing websites Detection

In F.I.R. Registration, users can registered there complain using simple instruction that was given by the bot. As Bot knows all rules, regulations and how police body works, so it gives the right instruction to the user based on the cases.

Chatbot is designed to provide better conversation with the user. AIML Chatbots responses are best used for frequently asked questions or current issues that must be addressed immediately. Artificial Intelligence Markup Language (AIML) is an XML-compliant language that permits knowledge content to be received, processed, and served on the online. Artificial intelligence programming is meant with a series of responses that represent the way a person's being responds to an issue. These responses and their corresponding questions are contained during a simple document, called an AIML file. AIML files are typically used by web-based chat robots (chatbots). AIML provides artificial intelligence functionality that is not available in either XML or HTML[8].

Phishing maybe a common attack on credulous people by making them disclose their unique information using counterfeit websites. The objective of phishing website URLs is to purloin the private information like user name, passwords, and online banking transactions. Phishers use websites that are visually and semantically similar to those of real websites. As technology continues to grow, phishing techniques began to progress rapidly and this must be prevented by using anti- phishing mechanisms to detect phishing. Machine learning maybe a powerful tool used to strive against phishing attacks. This paper surveys the features used for detection and detection techniques using machine learning [11].

II. LITERATURE SURVEY

As mentioned earlier, the 'Cyber Aid Station' is the first of

DOI : 10.35291/2454-9150.2021.0117

its kind. There no exact technical papers published on this topic. However, there were few technical papers related to this topic gave us some ideas.

1. Virtual Policing from Hungarian Perspectives by Edina Kriskó on April, 22.-2016
This technical paper gives me some ideas of all working virtual police stations in the world. So, it helps me a lot to gather information related to our topics. This paper also gives ideas on how to make projects successful.

2. CHATBOT by Akshay Kumar, Pankaj Kumar Meena, Debi Prasanna Panda, Ms. Sangeetha on November, 11-2016
In this project, we have introduced a chatbot that is able to interact with users. This chatbot can answer queries in the textual user input. For this purpose, AIML with program-o has been used. The chatbot can answer only those questions which he has the solution in its AIML dataset.

So, to extend the knowledge of the chatbot, we can add the APIs of Wikipedia, Weather Forecasting Department, Sports, News, Government and tons more. In such cases, the user will be ready to talk and interact with the chatbot in any quiet domain. Using APIs like Weather, Sports, News and Government Services, the the chatbot is going to be ready to answer the questions outside of its dataset and which are currently happening in the real world. a subsequent step towards building chatbots involve helping people to facilitate their work and interact with computers using natural language or using their set of rules. Future Such chatbots, backed by machine- learning technology, will be ready to remember past conversations and learn from them to answer new ones. The the challenge would be conversing with the varied multiple bot users and multiple users.As future work, we will make a chatbot that's supported AIML and LSA. This technology will enable a client to interact with a chatbot during a more natural fashion. we will enhance the discussion by including and changing patterns and templates for general client queries using AIML and therefore the right responses are given more often than LSA[3].

3. Phishing Website Detection using Machine Learning: A Review by Purvi Pujara and M.B. Chaudhary on September, 09-2018.

Phishing maybe thanks to obtaining user's private information via email or site. Because the usage of the internet is extremely vast, most things are available online now it's either about buying clothes, electronic gadgets, crockery or to payment of mobile, TV & electricity bill. instead of standing call at the line for hours, people are being aware of using the web method. thanks to this phisher feature a wide scope to implement phishing scams[1].

4. E-Police System supported Android Application for Enhancement of Services of Developing Countries by Dr.Ayesha Butalia1, Nilofar M. Shaikh, Avez Quadari, Roshan Undirwade, Nahid Pathan on September, 07-2017

Proposed epolice system for enhancement of e-government services of Bangladesh", during this paper the future vision the investigators and constables will also have mobile workstations which are linked to the digital files also as join to the Interpol and databases in order that the police personnel will immediately get answers from their databases also and also decide on the software solution. Fundamental upgrade of the interior network system within the National Police Agency of Japan., during this paper The National Police Agency of Japan has contributed to the development, maintenance, and management of its info-communications network, and has been devoted to modernizing its police information infrastructure for improving the efficiency and effectiveness of police activities. EPolice Police Record Management System, during this paper Epolice is meant to supply total computerized data system support for the work of the police. Survey on the Police Tracking System, in this paper the system help the surveyor in their work for handling such accounting a part of fine pay and policemen location to understand surveyor.

Survey Paper on Online Virtual police headquarters , during this paper the application helps the general public to report about the crimes to the police with no fear within the correct time. Police. People Friendlier, this paper provides information to the people about the police headquarters daily safeguard mission. People can inform any complaints or accidental complaints via this app so that it directly reaches their station limit in charges, police in charges can take actions immediately[4].

III. PROPOSED SYSTEM

Cyber Aid Station is an android application which aims at three main aspects.

The three main functions of this application consist of
I. FIR Registration
II. Chatbot
III. Phishing URL Detection

I. FIR Registration:

The user can file the complaint directly without registering itself on the application because registration process will take time. By excluding registration process one can simply file the complaint easily whenever the crime happens. Also the user will be able to view the complaints filed from that application by anyone.[4]

DOI : 10.35291/2454-9150.2021.0117

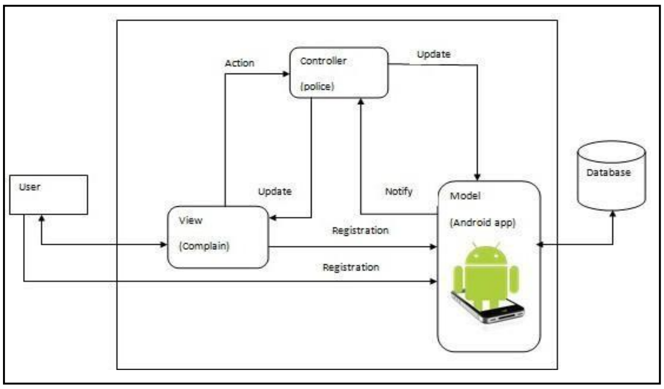


Fig 1: FIR Registration Flowchart

II. Chatbot:

“In this work, we've developed an interactive chatbot using the Flask framework in python, and the workflow of the proposed framework is shown in Fig 2.

This Chatbot uses pattern matching. Pattern matching is the process of checking whether a specific sequence of characters or tokens exists among the given data. Programming languages make use of regular expressions (regex) for pattern matching. Pattern matching is used to determine whether source files of high-level languages are syntactically correct. It is also used to find and respond to a matching pattern in a text or code with another text/code.

The steps for AIML Chatbot are as follows:

1. User discussion, as a rule, begins with the simple welcome or general questions.
2. User inquiries are first taken care of by AIML check, to check whether the entered inquiry is an AIML script or not.
3. AIML is characterized by general inquiries, queries, and welcome which is replied to by utilizing AIML formats.
4. Once the bot-user types within the query within the chatbot, the AIML developed chatbot will identify the category that contains the query pattern.
5. Here the bot-user is predicted to type within the query during a predefined pattern.
6. Once the query the pattern is matched, the template of the category that contains the response is shipped back to the bot user.[8]

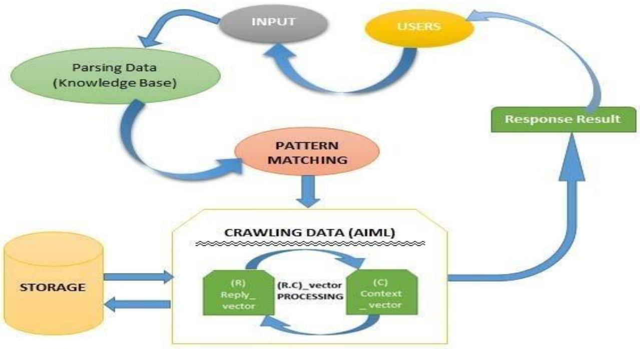


Fig 2: AIML Chatbot Flowchart

III. Phishing URL Detection:

Phishing is a way to obtain user’s private information via email or website. As usage of internet is extremely vast, most things are available online now its either about shopping cloths, electronic gadgets, crockery or to payment of mobile, TV & electricity bill. Due to this phisher has wide scope to implement phishing scam. As technology increases, phishing attackers using new methods day by day. This enables us to seek out effective classifier to detection of phishing. In this paper, we performed detailed literature survey about phishing URL detection. According to this, we can say random forest classifiers in machine learning approach is best suitable than other.[1]

Random Forest Classifier :

Random Forest, as its name implies, contains a large number of individual decision trees that act as a group to decide the output. Each tree in a random forest specifies the class prediction, and the result will be the most predicted class among the decision of trees. The reason for this amazing result from Random Forest is because of the trees protect each other from individual errors. Random Forests achieve a reduction in overfitting by combining many weak learners that underfit because they only utilize a subset of all training samples Random Forests can handle a large number of variables in a data set. Also, during the forest construction process, they make an unbiased estimate of the generalization error. Besides, they can estimate the lost data well.

The Dataset has been used for training and testing purpose is taken from the kaggle site. The dataset that we used in our research was well researched and benchmarked by some researchers.

The features of our dataset are as follows:

1) Having IP Address: If an IP address is used instead of the domain name in the URL
2) URL Length: Phishers can use a long URL to hide the doubtful part in the address bar.
3) Shortening Service: Links to the webpage that has a long URL.
4) Having Sub Domain: Having subdomain in URL.
5) SSL State: Shows that website use SSL
6) Domain Registration Length: Based on the fact that a phishing website lives for a short period.
7) Double Slash Redirection: The existence of // within the URL which means that the user will be redirected to another website
8) Prefix Suffix: Phishers tend to add prefixes or suffixes separated by (-) to the domain name so that users feel that they are dealing with a legitimate webpage.
9) HTTPS token: Having deceiving https token in URL.

DOI : 10.35291/2454-9150.2021.0117

10) Request URL: Request URL examines whether the external objects contained within a webpage such as images, videos, and sounds are loaded from another domain.

11) URL of Anchor: An anchor is an element defined by the < a > tag. This feature is treated exactly as Request URL.

12) Server Form Handler: If the domain name in SFHs is different from the domain name of the webpage.

13) Age of Domain: If the age of the domain is less than a month.

14) DNS Record: Having the DNS record

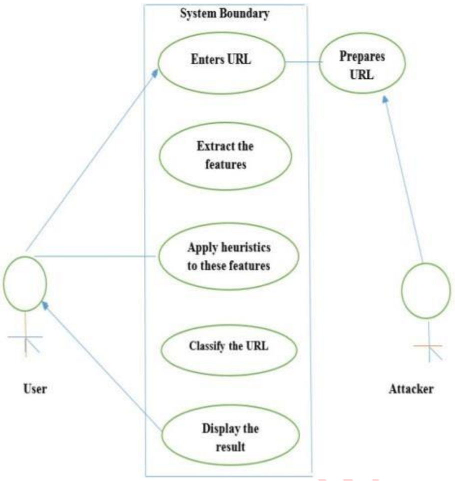


Fig 3: Phishing URL Detection

IV. RESULT

In Figure 4, the user can register complaint by filling up details such as Name, Phone Number, Email ID, Aadhar Number and details related type of complaint in complaint section.

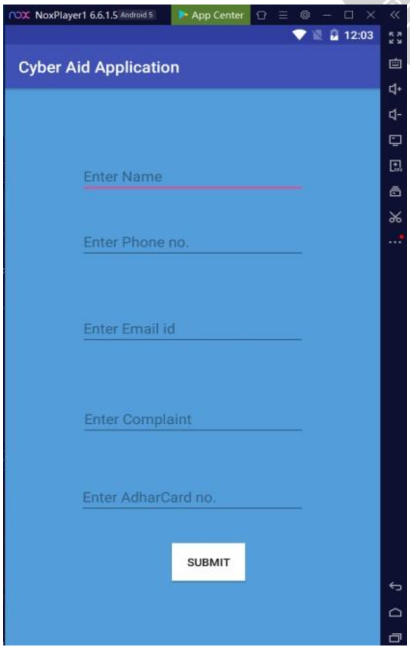
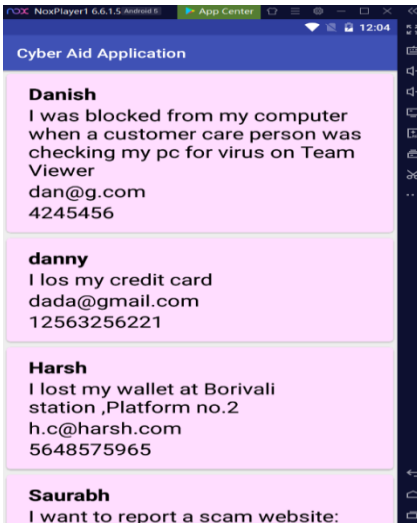


Fig 4: Registering the complaint of the user

In Figure 5, the user can view their complaints and the what is the status of the complaint. By clicking on view complaint tab the user can view the complaints.



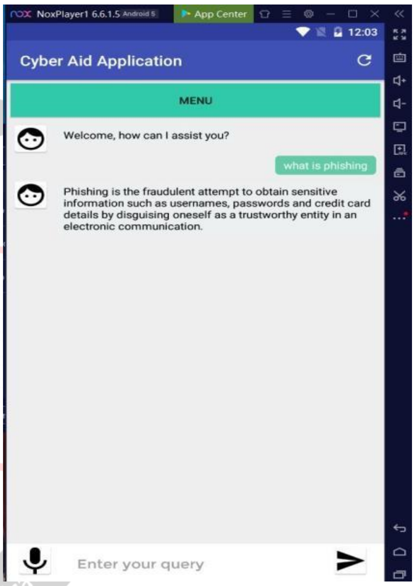
Danish
I was blocked from my computer when a customer care person was checking my pc for virus on Team Viewer
dan@g.com
4245456

danny
I los my credit card
dada@gmail.com
12563256221

Harsh
I lost my wallet at Borivali station ,Platform no.2
h.c@harsh.com
5648575965

Saurabh
I want to report a scam website:

Fig 5: Viewing the status of the complaint

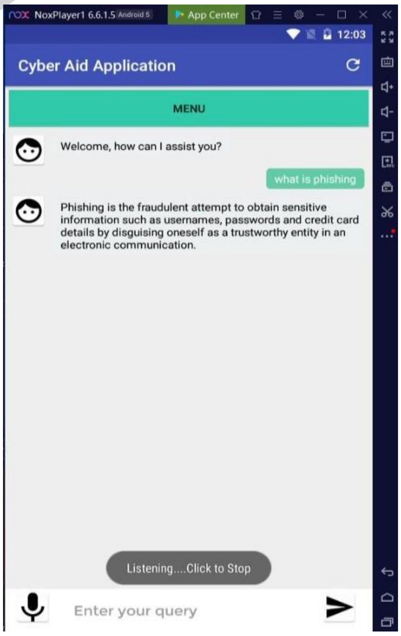


Welcome, how can I assist you?

what is phishing

Phishing is the fraudulent attempt to obtain sensitive information such as usernames, passwords and credit card details by disguising oneself as a trustworthy entity in an electronic communication.

Fig 6: asking query to the chatbot



Welcome, how can I assist you?

what is phishing

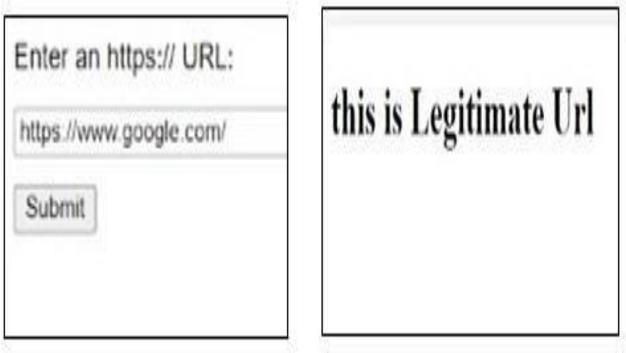
Phishing is the fraudulent attempt to obtain sensitive information such as usernames, passwords and credit card details by disguising oneself as a trustworthy entity in an electronic communication.

Fig 7: Using the speech to text feature to convert voice

DOI : 10.35291/2454-9150.2021.0117

of the user to speech

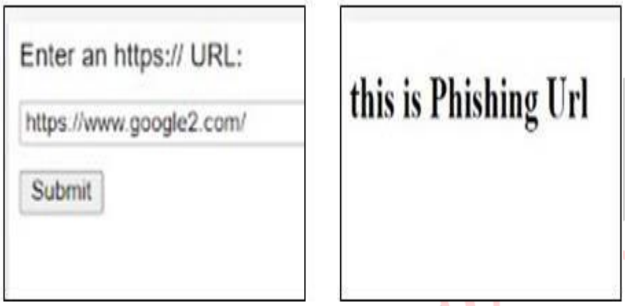
In Figure 6 and 7, aiml chatbot is used to assist the user and to solve all their queries related to the application. Speech to text feature is also given, so user feel comfortable and easy to use application.



https://www.google.com/

this is Legitimate Url

Fig 8: Result after processing URL to check the site is phishing or legitimate.



https://www.google2.com/

this is Phishing Url

Fig 9: Result after processing URL to check the Site is phishing or legitimate.

In Fig 8 and 9, a text field is given where user can enter a URL to check whether the site is phishing or it is a legitimate. It helps user to protect their private information from the phishing sites. It is fastest to determine and give the result.

V. CONCLUSION

Cyber Security Applications must be able to inform and educate the users about cyber threats and how to act when under a cyber threat. Phishing is a way to obtain user's private information via email or website. As usage of internet is very vast, almost all things are available online now it is either about shopping cloths, electronic gadgets, crockery or to payment of mobile, TV & electricity bill. Rather than standing out in line for hours, people are being aware of using online method. Due to this phisher has wide scope to implement phishing scam. As there is lot of research work done in this area, there is not any single technique, which is enough to detect all types of phishing attack.

This paper also discusses E-police reporting and administration framework which is effectively available to people in general and help them navigate through their queries easily through a conversation stimulator in the form of chatbot.

In this paper, we performed detailed literature survey about phishing website detection, implemented a model to detect phishing sites, created a complaint registering system and a system to view its status. We have also implemented a pattern matching chatbot using IBM Watson web services and speech to text web services.

This paper discusses of an application that combines all these different functions into a single application to help the needs of the general public while also educating them about cyber crime and helping them prevent cyber crime.

This paper also helps the user take the appropriate next steps when they are under a cyber crime attack.

VI. REFERENCES

[1] Phishing Website Detection using Machine Learning: A Review by Purvi Pujara and M.B.Chaudhary on September, 09-2018.

[2] Virtual Policing from Hungarian Perspectives by Edina Kriskó on April, 22.-2016.

[3] CHATBOT by Akshay Kumar, Pankaj Kumar Meena, Debi Prasanna Panda, Ms. Sangeetha on November, 11-2016.

[4] E-Police System based on Android Application for Enhancement of Services of Developing Countries by Dr.Ayesha Butalial, Nilofar M. Shaikh, Avez Quadari, Roshan Undirwade, Nahid Pathan on September, 07-2017.

[5] Federico Neri, Paolo Geraci "Online Police Station a state-of-the-art Italian Semantic Technology against cybercrime.", Advances in Social Network Analysis and Mining, 2009.

[6] www.ijret.net/archives/V3/i4/IRJET-V3I4235.pdf

[7] www.ijret.net/archives/V7/i7/IRJET-V7I710.pdf

[8] docs.bmc.com/docs/livechatcurrent/aiml- introduction-887398743.html

[9] Hossein Shirazi, Kyle Haefner, Indrakshi Ray: Fresh-Phish: A Framework for Auto-Detection of Phishing Websites: In (International Conference on Information Reuse and Integration (IRI)) IEEE,2017.

[10] Bhagyashree E. Sananse, Tanuja K. Sarode: Phishing URL Detection: A Machine Learning and Web Mining-based Approach : In International Journal of Computer Applications,2015.

[11] www.ijrte.org/wp-content/uploads/papers/v8i2S11/B10180982S 1119.pdf

[12] M. Goodman, Crime and Policing in virtual Worlds. On the EverEvolving Nature of Cybercrime, Freedom from Fear Magazine, Issue 7, July 2010. pp. 52-58, Avialable at: http://f3magazine.unicri.it/?p=360, accessed: 14th March, 2016.

DOI : 10.35291/2454-9150.2021.0117