CYBER CRIME ATTACK PREDICTION

Twinkle Shah, Sagar Parmar , Kishan Panchal

Under the guidance of Prof. Rashmi Malvankar
Department of Information Technology,
Shah And Anchor Kutchhi Engineering College, Mumbai,Maharashtra,India.

---***---

Abstract - *This paper focuses on to find spatial and temporal criminal hotspots using a set of real-world and past datasets of cyber crimes. We will try to locate crime locations and their frequent occurrence time based on past data. We will also predict what type of crime might occur in future on location with percent of attack. Finally, we intend to provide an analysis study by combining our findings of a particular cyber crimes with its specific information. We will conduct Apriori algorithm to produce interesting frequent patterns for criminal hotspots. We will provide this information in geographical view on map and provide the information with options.*

Key Words: *Cyber crimes, Prediction, Location, Criminal hotspot , Cyber Crime Attack Analysis.*

1.INTRODUCTION

Cyber crimes are a common social problem affecting the quality of life and the economic growth of a society. With the increase of crimes, law enforcement agencies are continuing to demand advanced geographic information systems and new approach approaches to improve crime analytics and better protect their communities.

Although cyber crimes could occur everywhere, it is common that criminals work on crime opportunities they face in most familiar areas for them. By providing a such approach to determine the most criminal hotspots and find the type, location and time of committed crimes, we hope to raise people's awareness regarding the cyber crimes and their effect on the company's legal data. In addition, having this kind of knowledge would help people to take appropriate steps to secure their data. On the other hand, Government can use this solution to increase the level of crime prediction and prevention. Moreover, this would be useful for company resources allocation. It can help in the distribution of company at most likely crime places for any given time, to grant an efficient usage of security resources.By having all of this information available, we hope to secure our data from cyber crimes attacks.

A . PROBLEM STATEMENT

Our study aims to find spatial and temporal criminal hotspots using a set of real-world datasets of cyber-crimes. We will try to locate the crime locations and their frequent occurrence time. In addition, we will predict what type of crime might occur next in a specific location within a particular time and which countries are most likely to get cyber attacks. Finally, we intend to provide an analysis study by combining our findings of a particular cyber-crime's dataset with its demographic information We will also display the live map which we display the number of attacks on a particular country using different colours to easily identify which countries are most likely to get attacks.

II . LITERATURE SURVEY

As survey done on various types of papers there are more surveys done on the physical crimes rather than cyber crimes. There has been countless of work done related to cyber crimes. Large datasets have been reviewed, and information such as location and the type of cyber crimes have been extracted to help people follow law enforcements and create the awareness how dangerous can be cyber attacks. Existing methods have used these databases to identify crime hotspots based on locations. There are several maps and online sites that show the crime location along with the crime type for any given country. Even though crime locations have been identified, there is no information available that includes the crime occurrences, locations, date and time, type of attack along with techniques that can accurately predict what crimes will occur in the future.

Figure 1. Geographical data

On the other hand, the previous related work and their existing methods mainly identify crime hotspots based on the location of high crime density without considering either the crime type or the crime occurrence date and time.

III. METHODOLOGY

We strongly believe that finding relationships between crime elements could highly help in predicting potential dangerous hotspots at a certain time in the future. Therefore, our proposed approach aimed to focus on three main elements of crimes data, which are the type of crime, the occurrence time and the crime location. We tried to extract all possible interesting frequent patterns based on the cyber crime variables. Then, we applied some classification methods in order to predict cyber crime attacks in a specific location within a particular time.

IV. DATA ANALYSIS

As an initial step to analyse and get the big view of our data, we conducted statistical analysis on the attribute values of our datasets. For each country, we started with generating a script to calculate frequencies of distinct values for every attribute. Then, we created a variety of graphs to give better understanding of our data. Each graph came up with the percentage of crime occurrences regarding a particular aspect.
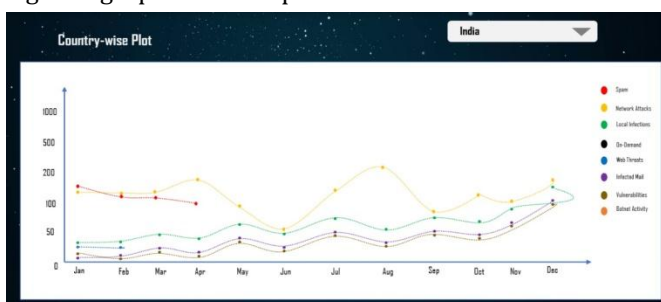

Figure 2. Number of Attacks based on country

Figure 3. Prediction of attack based on country, day and type of attack

There are steps in doing **Crime Analysis**:

1) Data Collection:-

Collection of the past data from various sources that where this attack is already occurred on collect this information that is data in form of attack.

2) Classification:-

Classify that which type of attack is occurred in different locations and gather that data in form of attributes. Different types of attack is classified based on the way they have attacked the systems.

3) Prediction:-

Predict the result on the basis of collected data from different places and resources. Prediction of result is generally based on the frequency of attacks at a particular location.

V. EVALUATION & IMPLEMENTATION

In this section, we evaluate each of the constructed module regarding different aspects. Apriori Algorithm is used to extract frequent crime patterns. The key strength of this model are its readiness and easiness of use and implementation. It takes time to load the webpage, especially with smaller values of the minimum support. We applied the cross validation strategy on both models then compared the prediction accuracy for each country.

We implement Apriori Algorithm to check pattern and extract in different pattern. It gives result in different section and its sub-section. We are providing map to give accurate result through graphical view. It gives result in different color on map that which country have high percent of attack and which country have less percent of attack. We are providing type selection option to check attack as per country wise and attack wise.

VI. CONCLUSIONS

We have gathered information from different sources and putting all information at single place, we made one dataset of past data of Cyber Crime of different countries. We have used Apriori Algorithm to check the pattern in dataset and to perform prediction on dataset. We are giving this information in geographical view of map with time wise prediction, country wise prediction and attack wise prediction, also we are providing information about the attack (description of attack and use of that attack).

We hope by publishing this paper starting a trend of Cyber Crimes Prediction, which can help law enforcements and keep our community safer for everyone.

REFERENCES

[1] https://arxiv.org/ftp/arxiv/papers/1508/1508.02050.pdf.

[2] https://pdfs.semanticscholar.org/2c4f/d3627edaa2e3908dfe9fb23ecf9ef6758d77.pdf.

[3] https://www.researchgate.net/publication/280722606_Crime_Analysis_and_Prediction_Using_Data_Mining