

A Study on Analytical Visualization of Deep Web

WooHyun Park*

Department of Electrical and Computer Engineering*, Sungkyunkwan University, Suwon, Korea

cgv@skku.edu

Abstract— Nowadays, there is a flood of data such as naked body photos and child pornography, which is making people bloodless. In addition, people also distribute drugs through unknown dark channels.

In particular, most transactions are being made through the Deep Web, the dark path. “Deep Web refers to an encrypted network that is not detected on search engine like Google etc. Users must use Tor to visit sites on the dark web” [4].

In other words, the Dark Web uses Tor's encryption client. Therefore, users can visit multiple sites on the dark Web, but not know the initiator of the site.

In this paper, we propose the key idea based on the current status of such crimes and a crime information visual system for Deep Web has been developed. The status of deep web is analyzed and data is visualized using Java. It is expected that the program will help more efficient management and monitoring of crime in unknown web such as deep web, torrent etc.

Keywords— Deep Web, Crime Trace, Analysis, Visualization

I. INTRODUCTION

Deep web is a kind of dark web that is not searched on a regular portal site.

As Deep Web has the advantage of ensuring anonymity through a search engine that is exposed to everyone, users are on the rise. However, there are many problems of social problem in deep web. It is easy for criminals to purchase goods that will be used for crimes because of their easy access. In addition, it is a stage where people prepare for crimes.

It is necessary to develop security techniques and investigative techniques to detect illegal criminal activities.

Because sellers can sell illegal information without significant risk through Deep Web, they can generate profits.

So, it is necessary to collect crime information profiling through dark web traffic and to require protocol analysis.

To prevent deep web, the government is focusing on recruiting developers such as offering rewards by holding a contest to track Tor browser. In fact, it is said that the government is taking a technology that can track dark web criminal behavior. In this paper, based on the current status of such crimes, a crime information visual system for Deep Web was developed. The state of crime is visualized by pilot analysis using Oracle DB and Java.

In related studies will introduce the programs used to implement the Deep Web Crime Status visualization system

and introduce crime situations that are focused in pilot analysis.

Subsequently, in methodology chapters explain how to implement and what to do, and finally conclude.

II. RELATED WORK

A. Oracle Database

Oracle Database is a relational database management system for private company.

After you create database, new data items can be added to the database without controlling the applications involved. This software system allows all app to share a database to be relatively easy to write and read and use and easy to expand.

The reason why we used to implement the system is that the data stored in the database can be accurately stored and processed to continuously respond to changes. The same data should be available for different purposes at the same time.

Besides, real-time accessibility is good because the processing time for the needs of many users is less than a few seconds

As one of the characteristics of the database, the data records should be referenced according to the value. This enables systematic use of information.

B. Deep Web

The Deep Web is another range of the Internet called the Deep Internet. “Deep Web refers to an encrypted network that is not detected on search engine like Google etc. Users must use Tor to visit sites on the dark web” [4]. In addition, it includes all user databases, and pages that must be passed through payment system. It can remain highly anonymous without being searched in search engines such as Google etc.

If the site's owner blocked the page's exposure, external access could be impossible. “It can only be accessed through a specific program, such as Tor[4].”

It was developed by American computer scientists in 2004.

Tor's user can access system inefficiently. For example, It runs in eight hours, circling a five-minute distance.

III. APPROACH

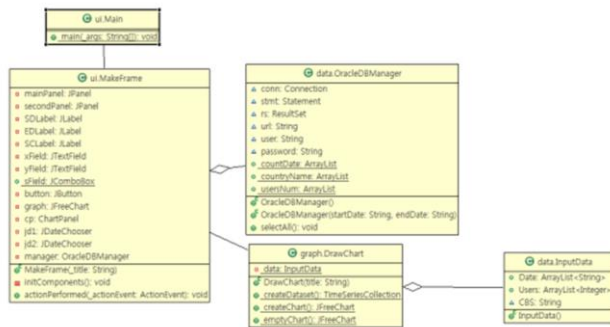


Figure 1. The structure of Implementation

Figure 1 shows the implementation of analytical visualizations through UML (Unified Modeling Language).

A class was marked on a square and the relationship between classes was shown using the arrows. The top is the class name, the middle is the field part, and the bottom is the method part. It is object-oriented and the data in charting class and deep web are interlinked with the OracleManager class and the DrawChart class as InputData class.

Specifically, OracleDBManager receives date and country information users. In InputData, we can choose CBS to load country. Using repeated process, we loaded the correct date and number of users for the country into Java.

In the DrawChart class, the chart was drawn using two parameter values: date and user. In the MakeFrame class, we created a frame and imported a chart here. MakeFrame also sent information to InputData CBS for the selected country. By default, we are preparing to run OracleDBManager to retrieve data in advance. Since there are many queries when data is retrieved from OracleDBManager, attempts to reduce queries used to extract data separately for each country selected in the InputData class.

A. Crime

Figure 2 is a pie graph of the crime in Korea related to the Deep Web.

As Figure 2 shows, crime through information sharing topped the list with 22 percent, followed by hacking crime with 21 percent. Thirdly, it is a crime of forgery. Fourth, the crime rate through virtual money transactions was high. We could also see that arms and drugs trade, which we had never thought of in Korea, were happening frequently.

This provides a useful analogy for when bitcoin and virtual money transactions are active and provides a single glimpse of the changes in each crime rate every year.



Figure 2. Status of Dark Web-related Crime in Korea

B. User

date	country	users	lower	upper	frac
2014-09-28	id	18972	15948	23480	70
2014-09-29	id	23652	20542	28168	68
2014-09-30	id	24022	20910	28065	69
2014-10-01	id	24017	21267	27794	68

Figure 3. Data sample collected in study

The data shown in Figure 3 shows user data in Indonesia.

In tor site, statistics can be drawn from which country uses the dark web a lot.

The data statistics were received through the statistics of 2018 Dark Web-related crimes, the National Police Agency and the National. The analysis visualization was conducted as a pilot with data from a total of 11 countries, including France, Germany, India, Indonesia, Netherlands, Russia, the United States, Britain, Ukraine, Iran and South Korea.

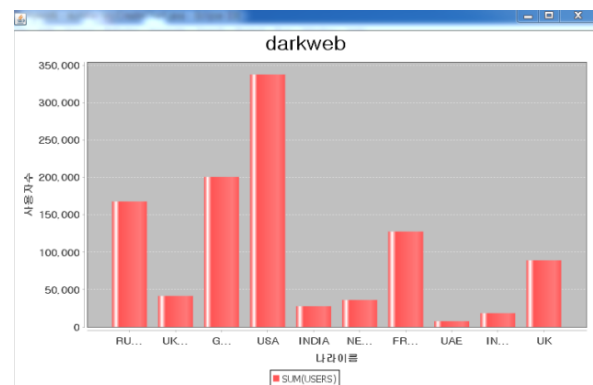


Figure 4. The number of Dark Web User in Countries

Figure 4 shows the number of dark web users in each country. Data from 11 countries were visualized in bar charts, with the largest number of users in the U.S. Second, Germany has many Deep Web users. Third, followed by Russia, then France and Britain and Ukraine, Netherlands, India so on in order.

This seems to be because the CIA has created Thor, a portal for accessing the Deep Web with its own network Web site. It

was initially said that it was made for people to inform the CIA anonymously.

[4] Dec.30,2019 from <https://blog.naver.com/x0472/221507411099/>

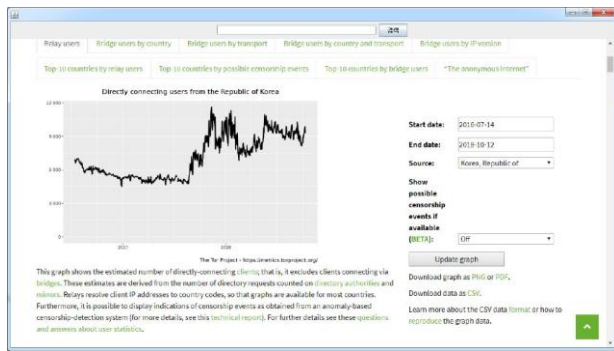


Figure 5. Status of using dark web in Korea

Figure 5 shows a line graph showing users connected to the Deep Web in Korea. Implemented using Google APIs, the section can be visualized by setting the start date and end date effectively.

IV. CONCLUSIONS

In this paper, we studied the analysis visualization through the Deep Web crime data and the number of Deep Web users in each country. Many crimes can happen because people have free access to unknown sites such as deep web and torrent while the site initiator is anonymous.

We took this topic and carried out pilot studies. As a result, it has created the program that can effectively show data on various graphs and can help police track and predict criminals.

It will also help detect criminals in areas where there are many people. Since Korea is still lacking in data, after collecting sufficient data, it will be possible to analyze whether it has been occurred a lot in crimes for clarity.

In future research, we will need to implement a method of accepting real-time users. Based on the status of crime, we will analyze the transactions of dark web and cryptocurrency by developing web robots for the dark web and criminal information retrieval systems.

It is expected to be sold as a security company and a business solution.

ACKNOWLEDGMENT

It is a work that was done jointly with other people in the academy.

REFERENCES

- [1] Chen, Hsinchun, et al. "Uncovering the dark Web: A case study of Jihad on the Web." *Journal of the American society for information science and technology* 59.8 (2008): 1347-1359.
- [2] Abbasi, Ahmed, and Hsinchun Chen. "Affect intensity analysis of dark web forums." 2007 *IEEE Intelligence and Security Informatics*. IEEE, 2007.
- [3] Fu, Tianjun, Ahmed Abbasi, and Hsinchun Chen. "A focused crawler for Dark Web forums." *Journal of the American Society for Information Science and Technology* 61.6 (2010): 1213-1231.



WooHyun Park is now a Doctor Student at Sungkyunkwan University in South Korea. His research interests lie in artificial intelligence