Open Social Data Crime Analytics

Ihsan Ullah, Caoilfhionn Lane, Brett Drury, Marc Mellote, and Michael Madden
Insight Centre for Data Analytics, National University of Ireland, Galway, Ireland
ihsan.ullah@insight-centre.org

Abstract

Crime is under-reported. Reporting crime requires an individual to complete a number of administrative obligations. These obligations, as well as the nature of the crime e.g. fraud, may create an inertia that discourages the reporting of the crime such as reputation damage. There may be, however, information leak from compromised organizations via affected individual customers on social media. A big advantage of using social data is that it is often immediate, and can have indications of the nature of a crime such as (1) named entities, for example, Bitcoin or PayPal; (2) geocoding information; and (3) the affected persons. Such signaling of incidents is arguably a better indicator of the extent and effect of cybercrime than traditional reporting methods. Our aim in this work is to use social media platforms e.g. Twitter, Reddit, Facebook, etc. for the detection of cybercrime.

1 Introduction

Advancement in computers and enormous usage of internet has made more people to like and rely on the ease of its provided services. This daily usage of smart phones, phenomena of the internet of things, cashless countries etc. has increased the risk of cybercrime. Cybercrime is the crimes committed or carried out with the help of computers or the internet. However, the majority of these crimes are under-reported due to many reasons e.g. mainly because of lack of knowledge or due to potential reputation damage [Levi *et al.*, 2016; Calnan and Denise, 2016]. In January 2015, Facebook & Google confirmed a phishing scam carried out between 2013 and 2015 in which both the companies lost $100M$dollars. The vendor management team were asked to transfer money based on fake documents. In December 2016, a CEO fraud perpetrated on Meath County Council, Ireland. A cybercriminal impersonated the county council CEO Jackie Maguire and requested the transfer of 4.3m to an account in Hong Kong. According to IBM X-Force threat intelligence index 2017, the five most targeted areas are financial institutions, information and communications, manufacturing, retail, and healthcare [Alvarez *et al.*, 2017]. On 12 May 2017, ransomware attack hit around 150 countries and damaged about 200000 computers [Karla and Adam, 2017]. However, the actual damage is still unknown.

In a recent survey [Levi *et al.*, 2016], a three-month crime related data provided by 'Action Fraud' UK national center for reporting fraud and cybercrime was analyzed. Out of 106,681 reported incidents, 4% of the incidents were related to cybercrime. Further, out of that 4%, 43% were cyber-related, 13% were cyber-enabled, and 29% were cyber-assisted crimes. The banking and credit industries face the highest number of fraud incidents [Levi *et al.*, 2016; Alvarez *et al.*, 2017]. Table.1 shows a list of recent attacks on financial institutions since 2014. When the end customer is affected directly, information leaks e.g. Tesco Bank shown in Fig. 1 and 2. Otherwise, financial institutions do not report due to potential reputational damage e.g. Vietnams or Ecuadorian bank [Spier, 2016]. These incidents show that cybercriminals may attack any weak organization.

2 Indicators of Cybercrime on Social Media and web forums

There are a number of sources of social media, social news, and web forums such as Twitter, Facebook, Instagram, Google+, Reddit, IRC, devRant, etc. One of the most frequent sources quoted in the literature is Twitter (more than 317 million active users who tweet between them 500 million times per day about their feelings, activities, opinions etc.). A number of studies have examined the role of Twitter in relation to criminal acts. A recent survey shows that the best channel to reach the public during a health crisis is Twitter, Facebook and Instagram [Duggan, 2014] that is also true in other scenarios (collecting and analyzing data) as well [Burnap and Williams, 2016]. Using Twitter to detect criminal acts is a recent area of study. Tweets have been used to detect both offline and online criminal acts, for example, to predict hit-and-run crimes from traffic alerts[Wang *et al.*, 2012], detect cyber hate [Burnap and Williams, 2016], detect online tension [Procter *et al.*, 2013; Burnap *et al.*, 2015], and to study online rumors in terms of offline harm [Webb *et al.*, 2015]. Social media is also used in research for health surveillance. Some of the famous indicators to detect cybercrime are Phishing, Fraud, Ransomware, Spam, Rumour, Riot, Bullying, Hate speech, etc.

Table 1: Some of the major cyberattacks/fraud in financial institutions during 2014-2017

No	Name	Location	Year	Target	Damage	Type
1	Many Institutions	~ 150 Countries	2017-05	~ 200K Computers	NA	Ransomware WannaCry
2	Lloyd Banking Group	UK	2017-01	Bank Online Services	NA	DDoS
3	Tesco Bank	UK	2016-11	9000 Customers	$3.10 M	Breached
4	First Bank and other Taiwanese Banks	Taiwan	2016-07	Banks ATMs	$2.2M	Breached
5	Qatar National Bank	Qatar	2016-04	Limited Customers 1.4GB	NA	Breached
6	Bangladesh Central Bank	Bangladesh	2016-02	Bank	$101M	SWIFT
7	HSBC	UK	2016-01	Bank Online Services	NA	DDoS
8	Royal Bank of Scotland and Natwest	UK	2015-07	Bank Online Services	NA	DDoS
9	Vietnams T. P. C. J. Stock Bank	Vietnam	2015-05	Bank	NA	SWIFT
10	Ecuadorian Banco del Austro	Ecuadore	2015-01	Bank	$12M	SWIFT
11	HSBC	Turkey	2014-11	Credit Card Details	NA	Breached
12	FirstBank, Scotiabank, Banco Popolare	British/US Virgin Islands	2014-11	~ 1230 Customers Data	NA	Breached
13	European Central Bank	Germany	2014-07	20K Records	NA	Breached

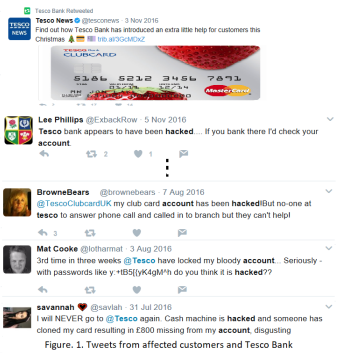

Figure. 1. Tweets from affected customers and Tesco Bank

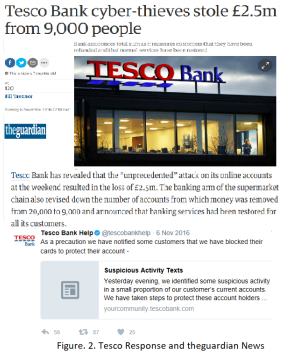

Figure. 2. Tesco Response and theguardian News

Figure 1: Compromised Financial Institutions and its leakage on Twitter

3 Methodologies, Data Sources, Tools, and Supplementary Material

A number of methodologies can be adopted to detect cybercrime in open social data e.g. Text classification, Spam classification, Rumour/Riot modeling, Pandemic disease detection, Event detection/alignment etc.

Data Sources:

We have identified a number (30+) of general cybercrime related open datasets. The majority of them are related to network related e.g. intrusion detection. Second most were related to reviews for spam or negativity detection. Only four of them were related to social media i.e. Twitter. Currently, no suitable dataset available that can be used.

Tools:

Although there is no known dataset that can be directly used for this project, however, there are tools with the help of which we can collect required data. e.g. Twitter, Reddit, In-House Insight API, IBM X-Force Exchange. As a starting point, we can take hints from previously committed crimes, search data related to that. And then use it for Irish financial institutions. e.g. there are only 20 banks in Ireland, so we can manually search data for them.

Supplementary Material:

In addition, to social media data, digital currency can or might also be a helpful tool in detecting suspicious activities. e.g. in recent ransomware attack, the ransom has to be paid in Bitcoins. Bitcoin Stock can be a valuable source in detecting cybercrime. As an example, an increase in Bitcoin purchases relative to the usual baseline number of transactions in Ireland should provide an indicator of how many individuals paid the ransom to ransomware attack.

4 Conclusion

Open social data can be helpful in detecting cybercrime in financial institutions. Analyzing digital currency stock may supplement open social data. Existing methodologies for riot/rumor, sentiment, or pandemic disease etc. can be adopted. Currently, no suitable dataset available that can be used. Certain APIs can be used to collect data. We can use a combination of two platforms to confirm a cyberattack e.g. any tweet on Twitter asking/discussing a compromised account of a specific bank and any discussion in same time period on a blog or forum e.g. on Reddit.

References

[Alvarez *et al.*, 2017] Michelle Alvarez, Nicholas Bradley, Pamela Cobb, Scott Craig, Ralf Iffert, Limor Kessem, Jason Kravitz, Dave McMillen, and Scott Moore. IBM X-Force Threat Intelligence Index 2017 The Year of the Mega Breach. (March):1–30, 2017.

[Burnap and Williams, 2016] Pete Burnap and Matthew L. Williams. Us and them: identifying cyber hate on Twitter across multiple protected characteristics. *EPJ Data Science*, 5(1), 2016.

[Burnap *et al.*, 2015] Pete Burnap, Omer F. Rana, Nick Avis, Matthew Williams, William Housley, Adam Edwards, Jeffrey Morgan, and Luke Sloan. Detecting tension in online communities with computational Twitter analysis. *Technological Forecasting and Social Change*, 95:96–108, 2015.

[Calnan and Denise, 2016] Nicola Anderson Calnan and Denise. Garda Cyber Crime unit head Michael Gubbins: 2016 is the year of ransomware, 2016.

[Duggan, 2014] Maeve Duggan. Social Media Update, 2014.

[Karla and Adam, 2017] Dwoskin Elizabeth Karla and Adam. More than 150 countries affected by massive cyberattack, Europol says, 2017.

[Levi *et al.*, 2016] Michael Levi, Alan Doig, Rajeev Gundur, David Wall, and Matthew Williams. Cyberfraud and the implications for effective risk-based responses: themes from UK research, 2016.

[Procter *et al.*, 2013] Rob Procter, Farida Vis, and Alex Voss. Reading the riots on Twitter: methodological innovation for the analysis of big data. *International Journal of Social Research Methodology*, 16(3):197–214, 2013.

[Spier, 2016] Thomas Spier. Tesco Bank Breach, 2016.

[Wang *et al.*, 2012] Xiaofeng Wang, Matthew S. Gerber, and Donald E. Brown. Automatic crime prediction using events extracted from twitter posts. In *5th International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction*, pages 231–238, 2012.

[Webb *et al.*, 2015] Helena Webb, Marina Jirotka, Rob Procter, Bernd Carsten Stahl, Omer Rana, Pete Burnap, William Housley, Adam Edwards, and Matthew Williams. Digital Wildfires': a challenge to the governance of social media? *ACM Web Science 2015*, 2015.