Studying Ransomware Attacks Using Web Search Logs

Chetan Bansal* chetanb@microsoft.com Microsoft Research Redmond, WA, USA
Pantazis Deligiannis* pdeligia@microsoft.com Microsoft Research Redmond, WA, USA
Chandra Maddila* chmaddil@microsoft.com Microsoft Research Redmond, WA, USA
Nikitha Rao* t-nirao@microsoft.com Microsoft Research Bangalore, India

ABSTRACT

Cyber attacks are increasingly becoming prevalent and causing significant damage to individuals, businesses and even countries. In particular, ransomware attacks have grown significantly over the last decade. We do the first study on mining insights about ransomware attacks by analyzing query logs from Bing web search engine. We first extract ransomware related queries and then build a machine learning model to identify queries where users are seeking support for ransomware attacks. We show that user search behavior and characteristics are correlated with ransomware attacks. We also analyse trends in the temporal and geographical space and validate our findings against publicly available information. Lastly, we do a case study on 'Nemty', a popular ransomware, to show that it is possible to derive accurate insights about cyber attacks by query log analysis.

KEYWORDS

ransomware; web search; query logs; web security

ACM Reference Format:
Chetan Bansal*, Pantazis Deligiannis*, Chandra Maddila*, and Nikitha Rao*. 2020. Studying Ransomware Attacks Using Web Search Logs. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), July 25–30, 2020, Virtual Event, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3397271.3401189

1 INTRODUCTION

With internet becoming ubiquitous, cyber attacks have become increasingly prevalent. Cyber attacks have had significant negative implications in major sectors such as healthcare, finance, manufacturing, etc. As per a study by the Internet Society[1], over 5 billion data records were exposed in 2018 and the total cost of these attacks was estimated to be over $45 billion. There are several different types of attacks such as phishing, wiretapping, denial of service, ransomware, etc. In particular, with the advent of cryptocurrencies, which are used for making payments to attackers, ransomware attacks have increased exponentially over the last decade. Ransomware encrypts and locks the victims' files until they pay money to the attackers. The financial damage due to ransomware attacks is estimated to increase from $8 billion in 2018 to $20 billion in 2021 [1]. Several cities and organizations have been impacted by these attacks [4] affecting millions of people in the process.

Given the significance of these attacks, it's critical to understand the impact and scale of these attacks. Most of the information available today is from manually curated public repositories [13] or by cyber security vendors [2]. However, these suffer from fragmentation and delays. So, it's important to explore alternative data sources and methods to improve our understanding of these attacks. Web search query logs offer a unique way to do population-scale analysis [6, 12]. For instance, Paparrizos et al. [15] and Xu et al. [16] have analyzed health related search queries to predict trends for various diseases and epidemics. Chancellor et al. [9] have shown that macro-economic factors like employment demand can be characterized using query logs. In the security domain, Canali et al. [7] used the browsing history of users to predict the risk of visiting malicious websites.

In this work, we do the first study to analyze cyber attacks, specifically ransomware, using query logs from Bing, a major web search engine. We mine ransomware related queries from anonymized query logs and use machine learning models to extract queries where users are seeking support for ransomware attacks. This is critical since we want to analyze queries and sessions where the users were likely attacked rather than those who were just looking for information about these attacks. Next, we do feature correlation analysis to understand if search behavior and user attributes are correlated with attacks. We also report on temporal and geographic trends for users who were seeking support for ransomware attacks. Lastly, we do a case study on the Nemty ransomware [8] and show that just by query log analysis we are able to learn about the origin and the effectiveness of the attack.

2 DATA

We use the anonymized query logs from Bing to perform our analysis which is conducted over a four month time span between July 1st, 2019 and October 31st, 2019. As web search patterns tend to vary significantly based on several factors, we focus this study on queries from US region with English locale. However, the methodology is generic and can be expanded to other regions and locales.

2.1 Terminology

Below, we define some key terms that we use throughout the paper:

(1) **Ransomware Queries** - Ransomware related queries having the keyword 'ransomware' in the query or the clicked URL(s).
(2) **Support Queries** - Ransomware queries indicating that the user is trying to seek solutions for attacks. Sample queries: 'how to recover encrypted files', '.besub ransomware decryption software'.
(3) **Non-support Queries** - Ransomware queries where the intent is not to find support or solution for attacks but for seeking

* These authors contributed equally to the work.

SIGIR '20, July 25–30, 2020, Virtual Event, China

ACM ISBN 978-1-4503-8016-4/20/07…$15.00
https://doi.org/10.1145/3397271.3401189

general information, facts, etc. Sample queries: 'top ransomware attacks', '20 Texas cities attacked with ransomware'.

(4) **Attacked Users** - Users who searched for at least one support query for ransomware attacks.
(5) **Safe Users** - Users who did not search for any support queries for attacks. We have randomly sampled one million safe users for the study.

Limitations: Please note that since the query logs are anonymized, we lack ground truth about individual users to validate our observations. However, in Section 5, we show that the insights from this study are consistent with public information about the attacks.

2.2 Manual Annotation

Owing to the large volume of query logs, manually labeling each query will be a mammoth task. So, we manually label 1000 queries and then train a machine learning model to find the support queries and the attacked users. First, four annotators individually label a random sample of 200 ransomware queries as either *support queries* or *non-support queries*. We then calculate the inter-annotator agreement score using Fleiss kappa [10]. With the resulting score being 94.21, translating to almost perfect agreement, each of the four annotators were asked to label a disjoint set of 200 samples each. Including the initial set of 200 samples, a labeled dataset of 1000 samples was created with support queries being 32.8% of the data.

2.3 Support Query Classification

The labeled data (see Section 2.2) is then processed before we train a binary classification model. We tokenize the query string and the clicked URLs and compute the word embeddings of tokens that are not stopwords using a pretrained Word2Vec model [14]. The individual token embeddings are then aggregated together resulting in a 300 dimension feature vector. Several classification models are trained on the data and the five-fold cross validation scores are reported in Table 1.

Table 1: Comparison of support query classifiers with 5-fold CV.

Classifier	Accuracy	Precision	Recall	F1 Score
Decision Tree	83.6	80	81	81
Random Forest	89.1	94	70	80
LinearSVC	**93.20**	94	**89**	91
Gaussian SVM	92.6	**98**	87	**92**

We observe that LinearSVC is the best performing model with the highest five-fold cross validation accuracy of 93.2% and a F1 score of 91. For the 158, 001 ransomware queries that were found in the four month duration, the trained LinearSVC model was used to derive the inference labels. A total of 12, 357 unique users were identified as attacked users which corresponds to 11.64% of the total users that searched for ransomware queries. The resulting dataset, which is a union of all the queries searched for by the attacked users and the safe users, comprises of 23, 439, 988 queries out of which 8, 957, 348 queries belong to attacked users.

3 USER BEHAVIOR ANALYSIS

The data collected from the previous section is analysed to identify the behavioral differences in attacked users and safe users. To this end, we identify different features and group them into different categories based on the type of behaviour it indicates. The list of categories and the corresponding features are as follows:

- Volume of search - number of queries, number of adult queries, dwell time, clicks, sat clicks (clicks with dwell time > 30s [11]).
- Diversity in searches - unique URL domains.
- Time of search - morning (6AM - 7PM), evening (7PM - 12AM) or night (12AM - 6AM)
- Day of the week - weekday (Monday to Friday) or weekend (Saturday and Sunday).
- Device used - device type, operating system and browser type.

Along with the total counts, we normalize the features at a session level as well as the user level. The feature values are computed for all users in the dataset. We then analyse the differences in distribution of feature values for all attacked users and safe users. Table 2 summarises the percentage difference in mean values of the feature distributions of attacked users and safe users where the feature values are aggregated at a session level. Note that only the features where the percentage difference was higher than 100% are shown in the table.

Table 2: Feature comparison for attacked & safe users.

Feature	Percent Difference
Total Number Of Queries	192.16
Total Number Of Adult Queries	191.91
Total Number Of Clicks	193.25
Total Number Of Unique URL Domains	193.22
Total Number Of Sat Clicks	193.49
Total Dwell Time	193.55
Total Number Of Requests At Morning	184.22
Total Number Of Requests At Evening	196.72
Total Number Of Requests At Night	188.81
Total Number Of Requests On Weekday	189.21
Total Number Of Requests On Weekend	189.13
Mean Total Number Of Queries	107.84
Mean Total Number Of Adult Queries	103.21
Mean Total Number Of Clicks	110.38
Mean Total Number Of Unique URL Domains	110.64
Mean Total Number Of Sat Clicks	127.80
Mean Total Dwell Time	130.40

Following the feature comparison, a feature correlation analysis is carried out using Spearman's correlation coefficient [17] as it is able to capture monotonic relationships between variables without assuming the data to be of normal distribution. The values of the coefficient range from -1 to 1 to denote negative and positive correlations. Once the coefficients are computed, the confidence of the results obtained is tested via a standard significance test. The correlation value of a feature is considered statistically significant if the significance level (or p-value) is less than 0.05 indicating a confidence level of 95%. Figure 1 summarizes the set of features which satisfy this threshold condition. It is evident from the coefficient values that there is very weak or no correlation between the variables and likeliness of being attacked by ransomware.

An interesting observation made was that attacked users generally had a much higher search volume compared to safe users which implies that the more the users searches the web, the more likely they are to be attacked. There was also significant positive

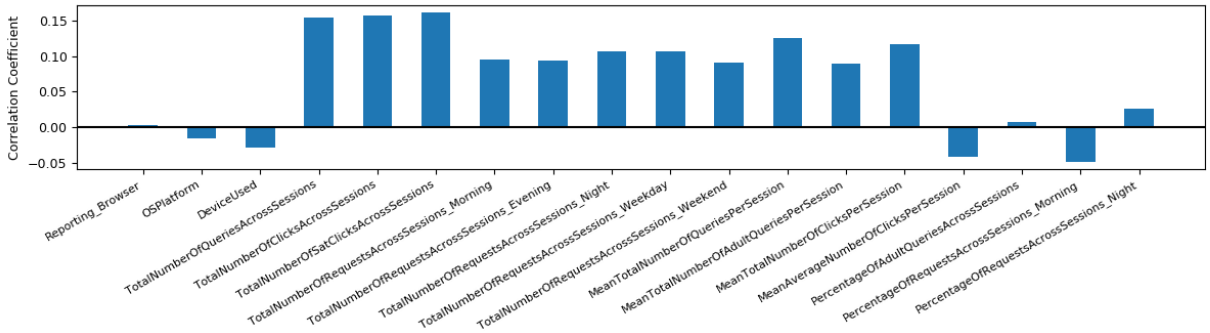

Figure 1: Spearman's correlation coefficients between features and being attacked by ransomware.

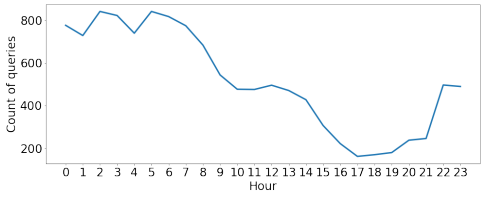

Figure 2: Hourly distribution of support queries.

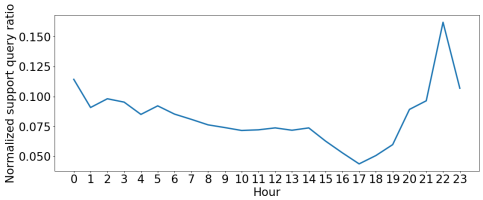

Figure 3: Hourly distribution of normalized support query ratio.

correlation between the percentage of queries searched at night time and a negative correlation for percentage of queries searched in the morning indicating that users are more likely to get attacked at night time. Another interesting behavior seen was that attacked users had higher positive correlations with adult queries.

4 TREND ANALYSIS

4.1 Hourly Trends

We analyzed how the behaviour of attacked users seeking solutions to ransomware attacks changes at each hour of the day by plotting hourly trends emerging from our dataset. Figure 2 shows that users were searching for solutions mostly during non-working hours (outside of the 9AM - 5PM window). This makes sense as users who are really determined to mitigate ransomware attacks on their own, by leveraging web search, are sparing some focused time outside of their regular working hours to find solutions. However, it could also be that the regular web search volume is very high during those hours. To better understand the trend, we plotted a graph shown in Figure 3 that shows the normalized distribution of ratio of support queries and how it is varying at different hours of the day. This bolsters our earlier finding (i.e., users spend more time searching for solutions outside of working hours) and highlights the fact that a lot of search activity to find solutions to ransomware attacks happens between 6PM - 11PM, which makes sense as this is the time window where users usually spare more focused time for finding solutions to their non-work problems.

4.2 Geographical Trends

In this study we focused on understanding how the search trends vary across different states in the US. In Figure 4 (left) we plotted a heat map of how the volume of support queries vary across different states in the US. Noticeably, states that are large (in size or population or internet penetration), like California, Texas, New York, is where a lot of activity is seen. However, this could also

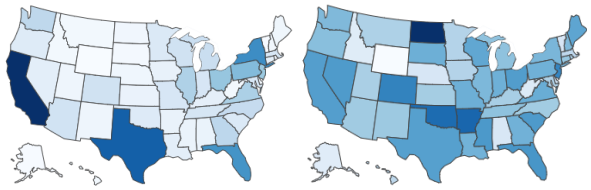

Figure 4: Distribution of support seeking query searches by US state (absolute numbers on the left and normalized ratios on the right).

be because the total search volume in these states is generally higher. To better understand which states record higher volumes of ransomware queries, we normalized the data by calculating the ratio of support queries compared to non-support queries in that state. This yielded some interesting insights, as seen in Figure 4 (right): states like North Dakota, Arkansas, Oklahoma is where the ratio of support queries is high though the overall search volume is low (compared to states like California or Texas). This can also be intuitively correlated to the massive ransomware attacks that were seen in various schools and public offices in states like North Dakota and Arkansas in the year 2019 [3, 5], which could have caused users in these states to record a higher normalized support query ratio.

5 CASE STUDY

We present a case study to see if we can learn insights about specific attacks using query log analysis . We looked at all recent (second half of 2019) ransomware attacks with significant impact listed by the NJ Cybersecurity and Communications Integration Cell (NJCCIC)[1]. For this study, we focus on the Nemty [8] ransomware, however our technique generalizes to any attack.

Nemty is a ransomware that infects Windows OS users, encrypts their files, searches and deletes any shadow copies of these files, and finally asks victims to pay a ransom for restoring their data. Nemty started affecting users end of August 2019, and spread worldwide

[1]https://www.cyber.nj.gov/threat-profiles/ransomware-variants/

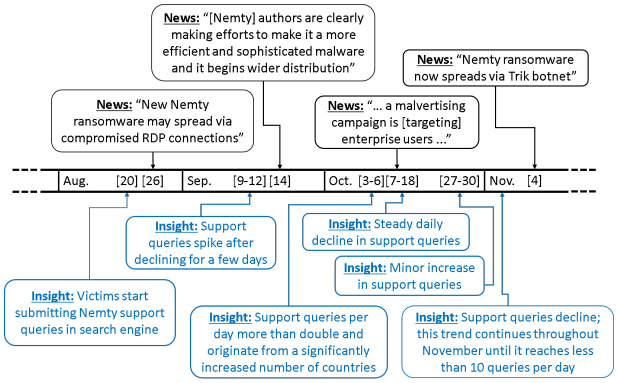

Figure 5: Timeline of news and query logs insights about Nemty.

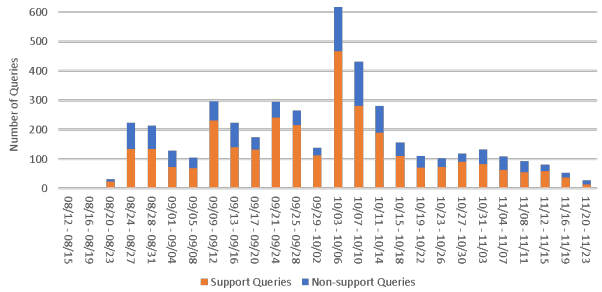

Figure 6: Volume of queries about Nemty per 4-day periods.

through distribution campaigns during September, October and November 2019, as seen in the timeline of Figure 5 (black boxes above the timeline show published news about Nemty [8]).

We gathered attack-related search engine query logs (with English locale) about Nemty between the start of August to end of November 2019 from all countries. We then classified the queries as support and non-support using our ML model (see Section 2.3). Finally, we analyzed the results to gain insights about Nemty, such as when the ransomware started infecting users and how its distribution evolved over time (see blue boxes below the timeline in Figure 5). Our insights found trends related to the distribution of Nemty that start days before they are reported in the news [8]. This result shows that our query log analysis technique could be used to timely learn about the origin and effectiveness of a ransomware attack, even in the early days of its distribution.

In Figure 6, we show the number of such queries per 4-day periods between August and November 2019. We see that until 08/19, there were no queries about Nemty. However, users started submitting support queries on 08/20, which is likely when the ransomware started first spreading. Indeed, on 08/26 the first news about Nemty are published [8] (see Figure 5). On 09/14, an article [8] discusses how the Nemty authors are enhancing the ransomware with the goal of achieving a wider distribution. Our analysis indeed found that although the support queries started declining early September, on 09/09 there was a spike in such queries (see Figure 6). The insight we gained from this corresponds to the news about the ransomware becoming more efficient and sophisticated.

There was a tremendous increase in support and non-support queries about Nemty between 10/03 and 10/06 (see Figure 6). Published news confirm this finding, as there was a new distribution campaign during October that was targeting enterprise users [8]. We found that Nemty was now spreading worldwide, as support queries started being submitted by an increasing number of countries. After this period of time, the support queries started decreasing, most likely because more people became aware of the ransomware and how to protect from it. In early November there was a minor increase in the queries (see Figure 6), which corresponds to a new distribution method of Nemty via Trik botnet. However, this method was not very efficient, because after this minor increase in queries, the query volume continued to decline.

6 CONCLUSION

In this work, we did the first study to find insights about ransomware attacks using web search logs. We analyzed query logs from a major web search engine using a machine learning classifier to extract support queries by users who were attacked by ransomware. We did a correlation analysis and found that certain features such as query volume and click counts are correlated with attacks. Further, we analyzed geographical and temporal trends and validated our findings from publicly available information. Lastly, we did a case study on the Nemty ransomware and showed that with query log analysis, it is possible to mine key insights about the origin and spread of specific attacks.

REFERENCES

[1] 2019. 2018 Cyber Incident & Breach Trends Report. https://www.internetsociety.org/resources/ota/2019/2018-cyber-incident-breach-trends-report/
[2] 2019. McAfee Labs Threats Report August 2019. https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-aug-2019.pdf
[3] 2019. The Number of Cyber Attacks Against the North Dakota Government Triples in a Year. https://www.kxnet.com/news/local-news/the-number-of-cyber-attacks-against-the-north-dakota-government-triples/
[4] 2020. Ransomware Attacks Grow, Crippling Cities and Businesses. https://www.nytimes.com/2020/02/09/technology/ransomware-attacks.html
[5] 2020. Ransomware attacks hit schools, but experts question intent, whether trend affects Arkansas. https://www.arkansasonline.com/news/2019/jul/22/ransomware-attacks-hit-schools-20190722/
[6] Chetan Bansal, Thomas Zimmermann, Ahmed Hassan Awadallah, and Nachiappan Nagappan. 2019. The Usage of Web Search for Software Engineering. *arXiv preprint arXiv:1912.09519* (2019).
[7] Davide Canali, Leyla Bilge, and Davide Balzarotti. 2014. On the effectiveness of risk prediction based on users browsing behavior. In *Proceedings of the 9th ACM symposium on Information, computer and communications security*. 171–182.
[8] NJ Cybersecurity & Communications Integration Cell. 2019. Nemty Ransomware. https://www.cyber.nj.gov/threat-profiles/ransomware-variants/nemty
[9] Stevie Chancellor and Scott Counts. 2018. Measuring employment demand using internet search data. In *Proceedings of the 2018 CHI*. 1–14.
[10] Joseph L Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin* 76, 5 (1971), 378.
[11] Steve Fox, Kuldeep Karnawat, Mark Mydland, Susan Dumais, and Thomas White. 2005. Evaluating implicit measures to improve web search. *ACM TOIS* 23, 2 (2005), 147–168.
[12] Shagun Jhaver, Justin Cranshaw, and Scott Counts. 2019. Measuring professional skill development in US cities using internet search queries. In *Proceedings of the International AAAI CWSM*, Vol. 13. 267–277.
[13] Amin Kharraz, William Robertson, Davide Balzarotti, Leyla Bilge, and Engin Kirda. 2015. Cutting the gordian knot: A look under the hood of ransomware attacks. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*. Springer, 3–24.
[14] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed Representations of Words and Phrases and Their Compositionality. In *NIPS'13 - Proceedings(Volume 2)*. USA.
[15] John Paparrizos, Ryen W White, and Eric Horvitz. 2016. Detecting devastating diseases in search logs. In *Proceedings of the 22nd ACM SIGKDD*. 559–568.
[16] Danqing Xu, Yiqun Liu, Min Zhang, Shaoping Ma, Anqi Cui, and Liyun Ru. 2011. Predicting epidemic tendency through search behavior analysis. In *22nd IJCAI*.
[17] G. U. Yule and K. M. Kendall. 1968. An introduction to the theory of Statistics. 14 (1968).