The Social and Technological Incentives for Cybercriminals to Engage in Ransomware Activities

Yichao Wang[1][0000−0002−4633−3690], Sophia Roscoe[1][0009−0008−2029−3601], Budi Arief[1][0000−0002−1830−1587], Lena Connolly[2][0000−0002−7110−9594], Hervé Borrion[3][0000−0003−3624−4763], and Sanaa Kaddoura[2][0000−0002−4384−4364]

[1] University of Kent, UK,
yw300@kent.ac.uk, sophiaroscoe27@gmail.com, b.arief@kent.ac.uk
[2] Zayed University, United Arab Emirates,
alena.connolly@zu.ac.ae, sanaa.kaddoura@zu.ac.ae
[3] University College London, UK,
h.borrion@ucl.ac.uk

Abstract. Ransomware attacks and the use of the dark web forums are two serious contemporary cyber-problems. These two areas have been investigated separately in the past, but there is currently a gap in our understanding with regard to the interactions between them – i.e., dark web forums that can potentially lead to ransomware activities. The rise of Ransomware-as-a-Service (RaaS) exacerbates these problems even further. The aim of this paper is therefore to investigate the social and technological discourse within the dark web forums that may foster or initiate some of the users' pathway towards ransomware-related criminal activities. To this aim, we carried out data collection (crawling) of pertinent posts from the "Dread" dark web forum, based on sixteen keywords commonly associated with ransomware. Our data collection and manual screening processes resulted in the identification of 1,279 posts related to ransomware, with the posting dates between 25 March 2018 and 30 September 2022. Our dataset confirms that ransomware-related posts exist on the Dread dark web forum. We found that these posts can generally be grouped into eight categories: Hacker, Potential Hacker, RaaS Provider, Education, Information, News, Debate and Other. Furthermore, the contents of these posts shed some light on the social and technological incentives that may encourage some actors to get involved in ransomware crimes. In conclusion, such posts pose a threat to cyber security, because they might provide a pathway for wannabe ransomware operators to get in on the act. The findings from our research can serve as a starting point for devising practical countermeasures, for instance by considering how such posts should be handled in the future, or how some follow-up intervention actions can be prepared in anticipation of certain actors getting involved in ransomware as a result of reading posts in such forums.

Keywords: Ransomware · dark web · dark web forum · social interaction · data collection and analysis · crawler.

1 Introduction

Ransomware is one of the most harmful cyber threats to individuals and organisations [25]. Ransomware is a type of malware that locks a computer system or prevents users from accessing their data until a ransom is paid. This is in contrast to other types of malware, which are often aimed at replicating, deleting or overburdening system resources [5]. While cyber extortion is at the heart of recently emerged ransomware variants (including threatening to reveal sensitive data), incredible technological advances (e.g., advanced propagation capabilities and virtually unbreakable cryptography) enable criminals to continue these harmful operations and generate rather lucrative returns. Lately, ransomware turned into a transnational organised crime run by so-called "career criminals" who not only initiate attacks but also run Ransomware-as-a-Service (RaaS) operations. Hence, developing measures to combat this threat is of vital importance.

Research on ransomware has focused on several avenues, including an investigation of dark web forums. These sites are commonly used by cybercriminals and other individuals to socialise, exchange information and sell illegal products and services [34]. Scholars believe that dark web forums help sustain cybercrime ecosystems. Several academic papers have investigated various aspects of interactions within these cybercrime ecosystems [1,30,43], even looking in detail into the actors involved [2,33], and specific types of cybercrime ecosystems [17]. However, ransomware research currently receives less attention than spam emails, spam tweets, hate speech detection and other types of cyberattacks. As such, studying ransomware cybercrime ecosystems further is essential in continuing our fight against cybercrime in a more thorough and balanced manner.

This paper examines ransomware-related posts from a dark web forum called "Dread" [10,15]. Dread is a Reddit-like dark web forum that emerged in 2018 and became popular as a result of Reddit's crackdown on several dark web market discussion communities. While one can find posts on illegal drugs and trades of stolen data, Dread also features professional hacking posts and in-depth guides on hacking. Due to its growth, Dread became a target for frequent distributed denial-of-service (DDoS) attacks. After suffering a prolonged downtime in 2022, in late November 2022 Dread went offline for server upgrades [15]. Recently, the forum returned online, and this allowed us to continue our focus on investigating the social interactions of various stakeholders. The collected dataset provides valuable insights into ransomware activities of various actors. Whilst showing a range of behaviours, there is a strong inclination towards the request and sharing of knowledge.

Contributions. The key contributions of our paper are as follows:

- To the best of our knowledge, this is the first academic work that specifically focuses on ransomware posts on a dark web forum;
- The posts were classified into categories, which demonstrated the nature of discussions among individuals who have interest in ransomware. These ranged from educational (e.g., actors who searched/provided advice on vari-

ous ransomware-related subjects) to malevolent (e.g., actors who were interested to buy/sell ransomware). These findings confirm that dark web forums such as Dread facilitate cybercrime;
- A quantitative analysis (i.e., counting posts for each category) indicated the overall "tone" and nature of ransomware discussions. While a majority of the observed Dread users demonstrated curiosity towards a ransomware subject and could not be labeled as malicious with certainty, the minority of users clearly asserted their malicious intents;
- A further examination of each post (i.e., a qualitative approach) confirmed the results of the quantitative phase and provided a deeper understanding of intentions of the users (i.e., from potentially non-malicious to clearly malicious). Such understanding can then be used by security researchers and law enforcement agencies to devise more effective intervention measures.

The remainder of this paper is organised as follows: Section 2 provides an overview of related work. Section 3 details our methodology, including the implementation of our crawling approach. Section 4 outlines the results obtained, while Section 5 discusses the insights gained from our research. Finally, Section 6 summarises the main points of the paper and outlines ideas for future work.

2 Related Work

A great deal of research has been conducted to understand the dynamics of dark web forums, including the structures of forum user networks [2,43], the analysis of the key actors [33] as well as their social dynamics such as how users gain or lose trust [1,30], and how these sites might facilitate various forms of cybercrime [17]. This body of research is vital in our attempt to better understand cybercrime and develop more effective measures against the threat of "internet organised crime" [13]. While there are many ongoing investigations in this area, we are currently not aware of any research that specifically focuses on the threat of ransomware and how the dark web forums may influence it.

The development of ransomware has received a lot of (and an increasing) interest within the security community in recent years. Researchers have studied various technical aspects of ransomware, including its detection [3,20,37], recovery from ransomware incident [8,23], as well as other potential mitigation measures [28,36]. In comparison, there are still limited studies that analyse the social aspects of ransomware – some of them are discussed below. Moreover, ransomware attacks nowadays not only rely on technological aspects, but also on human factors, which involve the spread of ransomware and the negotiation process that differ from other attacks. Therefore, there is a need to conduct more detailed investigations into the incentives of ransomware cybercriminals.

With the rise of cryptocurrencies in recent years, several studies have been conducted to track and analyse their economic impact. Huang et al. [19] conducted an end-to-end measurement of ransomware payments, victims and operators over a two-year period based on ransom wallet addresses. They conservatively estimated that approximately 20,000 victims were extorted during the

two years of the study and that the criminals earned more than $16 million in illegal revenue in the overall ecosystem. Hernandez-Castro et al. [16] carried out an economic analysis of ransomware, predicting that further ransom increases should be "expected". Moreover, with the increased popularity of RaaS in recent years, criminals from non-technical backgrounds are increasingly getting involved in ransomware attacks [29,32].

Connolly and Wall [7] conducted an analysis of 26 ransomware attacks by collecting data via interviews with victims and law enforcement representatives, leading to an interdisciplinary data-driven taxonomy of ransomware countermeasures. Connolly et al. [41] utilised data from 55 ransomware cases to assess factors that influence the severity of ransomware attacks. They found that private organisations and/or organisations that had weak security postures may be more vulnerable and that targeted attacks are often more devastating. Yilmaz et al. [39] conducted a survey to examine the relationship between personality characteristics and ransomware victimisation. They found that there is no clear evidence to indicate that personality traits would influence ransomware victimisation. Lang et al. [25] conducted a qualitative comparative analysis of 39 ransomware attacks based on interviews and secondary sources. They aimed to understand how the COVID-19 pandemic affected the tactics of these ransomware attacks. The results showed that working from home increases the risk of being attacked compared to traditional work patterns, while the laissez-faire attitude of organisations towards such attacks may lead to more serious issues.

Interestingly, we found that most of the research on the social aspects of ransomware have been focused on the victims. A closer look at attackers' activities and interactions can offer valuable insights. Dark web forums – as important places for the exchange of information between cybercriminals – are notorious in facilitating cybercriminal activities [27,34]. They constitute a rich data source to understand the activities and perspectives of cybercriminal actors. By analysing a forum, Pastrana et al. [33] demonstrated how members of this forum, who are interested in technology and games, are gradually transitioning to committing crimes. Yue et al. [40] analysed the discussion of DDoS attacks in forums and discussed the impact of dark web forums on such attacks. Bada and Pete [4] analysed the discussion in the dark web forums around Shodan, which is a search engine that could pose a threat to Internet of Things devices.

The availability of datasets is often a challenge for this type of work. This is usually due to restricted access to dark web forums or technical difficulties [35]. A potential dataset is the CrimeBB [34], collected and maintained by Cambridge Cybercrime Centre. This dataset contains several forums from both dark web and clear web. However, this dataset is not specifically crafted for ransomware-related research (and it was last updated in December 2021). Several other studies [11,18,42] have highlighted the need to collect their own data from various dark web or underground forums, such as Dread. These papers have shown that such an approach is possible and can provide valuable insights, because the collected data will be tailored to the specific research questions or aims. As such, we also decided to collect our own pertinent and more recent data from Dread.

3 Methodology

One of the main goals of our research is to understand the pathway, motives and facilitating factors that may lead some people to become a ransomware criminal and decide to prepare or carry out ransomware attacks. To achieve this goal, we analysed ransomware-related posts on the Dread forum. This forum was selected as it contains a comprehensive discussion of general matters and strong reputation [15]. Specifically, we expected discussions on Dread to be less technical than on specialist hacker forums. Therefore, the messages on Dread were likely to have been posted by a more diverse group of users. The rest of this section provides an overview of the methodology we followed for data collection, as well as ethical issues that we had to consider.

3.1 Research Design

We used the search function with a list of keywords related to ransomware to identify the initial list of candidate posts. A researcher then manually screened the results, and labelled all threads and posts that were related to ransomware. The filtered results were manually analysed to determine their purpose and the actors involved. Eight categories (themes) were identified in the posts (n=1,279). These categories were further subdivided (were applicable) into sub-categories to better understand their intent. Each post was then assigned to up to two category groups (e.g., a post could fall into both the "Hacker" and "Information" categories). These categories and sub-categories are:

- **Hacker**. The post indicates that its author has performed a ransomware attack. There are two sub-category labels: "group" and "individual".
- **Potential Hacker**. The post indicates that its author plans to perform a ransomware attack. Sub-category labels: "group" and "individual".
- **RaaS Provider**. The post contains a user offering RaaS for sale. Sub-category labels: "group" and "individual".
- **News**. The post refers to ransomware-related real world events (e.g., actual ransomware attacks). No sub-categories were identified for this category.
- **Education**. The post contains explicit educational information about ransomware related subjects. Sub-category labels: "request" and "provider".
- **Information**. The post requests or provides general information that cannot be classified as "Education" or "News". Sub-category labels: "request", "provider", and "moderator".
- **Debate**. The post presents an opinion, often initiating or contributing to a debate. No sub-categories were identified for this category.
- **Other**. Posts that do not fit any of the previous categories. No sub-categories were identified for this category.

These categories also allowed for a quantitative analysis on the frequency of each category as well as what that entails (i.e., a qualitative approach) as discussed in the Results section. Lastly, statistical analysis regarding the frequency of each keyword was performed within the post and thread respectively.

There are sixteen keywords (case insensitive) used to identify posts related to ransomware activities: *Ransomware*, *Extortion*, *Cyber extortion*, *Cyberextortion*, *RaaS*, *REvil*, *Sodinokibi*, *LockBit*, *Avaddon*, *BlackMatter*, *Ransomex*, *DarkSide*, *BlackCat*, *ALPHV*, *Hive*, and *Lockbit Black*. These keywords were provided by two researchers experienced in the field of ransomware. They correspond to terms related to ransomware attacks (e.g., "Ransomware", "Cyberextortion"), or the names of notable ransomware variants and/or groups at the time of the study.

We used a custom crawler to collect ransomware-related posts on 2 November 2022. The initial set of the collected posts (we call them "raw data") contained a sample of 19,109 candidate posts, spanning a period of 1,720 days (16 February 2018 to 1 November 2022). However, there were quite a lot of "false positives" in the raw data, whereby many posts included in this initial dataset were actually not ransomware-related. Therefore, we had to refine the initial dataset to remove any posts that were not ransomware-related. This manual filtering yielded the final dataset of 1,279 posts, covering the period between 25 March 2018 and 30 September 2022. Subsequently, labelling was performed to indicate the purpose and category of each post. The frequencies of the term "ransomware" in the post and thread along with other keywords were also calculated.

To have a better confidence regarding the relevance of the posts, we cross-referenced the ransomware attacks mentioned in some of the posts (e.g., the Colonial Pipeline and the REvil ransomware attacks) to news articles from reliable sources, such as the BBC, Kaspersky, and BleepingComputer.

3.2 Technical Implementation

In this study, we used both The Onion Router (Tor, https://www.torproject.org/) and the Invisible Internet Project (I2P, https://geti2p.net/en/) to access the Dread website. The Tor network experienced widespread DDoS attacks in October 2022, which resulted in reduced accessibility to the Dread forum [9]. Using the I2P network to access and collect data was the only alternative at the time. Technically, both Tor and I2P are decentralisation protocols, and they are just implemented in different ways.

We implemented a customised crawler in Python, based on the Scrapy web-crawling framework [24]. Due to the risk of potential (but unlikely) attacks against us, the crawler run on a virtual machine to avoid compromising the identity of researchers. A VPN tunnel was used to ensure that the geographical location and IP address would not be compromised during data collection. Tor or I2P were employed as a proxy to enable Scrapy to connect to the network. The crawler used the pre-defined keywords to search Dread, and to retrieve relevant posts. It traversed each page to obtain the URLs of all threads, keeping only one URL (if there were duplicates) to minimise the number of requests.

Finally, the crawler accessed each thread's URL and extracted all the necessary data points and features based on the web page structure. We pre-defined the following features for each post in the raw dataset: post ID, content, creator, whether the post was original or part of a thread, time of post, subdread (like subreddit), thread URL, thread title, number of users involved, number of posts

Table 1. The numeric breakdown of the posts among the eight categories

	Hacker		**Potential Hacker**		**RaaS Provider**		**Education**		**Information**			**News**	**Debate**	**Other**
	Group	Individual	Group	Individual	Group	Individual	Request	Provider	Request	Provider	Moderator	\	\	\
	22	6	22	99	26	44	76	89	216	370	5	\	\	\
Total	28		121		70		165		591			161	265	63
Percentage	2.19%		9.46%		5.47%		12.90%		46.21%			12.59%	20.72%	4.93%

in the thread, the time difference between the previous post in the same thread, the time difference between the last and the original post in the same thread.

3.3 Ethical Considerations

As the dark web is mostly uncensored, there was a risk that the researchers conducting the search could be exposed to detailed information about a wide range of criminal activities. For this reason, data collection was performed automatically and *only textual data* was collected. The data was saved into a comma-separated values (CSV) file on an offline external hard drive to avoid leakage. Access to the file was restricted to the researchers involved in this project.

Due to the anonymous nature of the dark web, we were unable (nor interested) to collect personal information of users, or track their real identities. Nonetheless, we still had to anonymise the usernames of the Dread forum users, because it might be possible to use these usernames to connect back to the their real identities. At the same time, it would still be valuable to be able to link various posts to each entity. As such, when referring to statements from a particular user, we used a pseudonym (e.g., "User 1"), which would still allow some data linking to be performed, while protecting the privacy of the users involved. The ethics of this study has been reviewed and approved by Zayed University Ethics Committee (Ref: ZU22_033_F).

4 Results

In this section, we aim to further examine the reasoning and behaviours of (potential) criminals who engage in ransomware. To achieve this, we have used a combination of figures, tables, and analytical tools to gain insights. This enabled us to find and better understand significant trends and patterns of ransomware activities. In total, eight main categories were identified in the posts (n=1,279), as already outlined in Section 3.

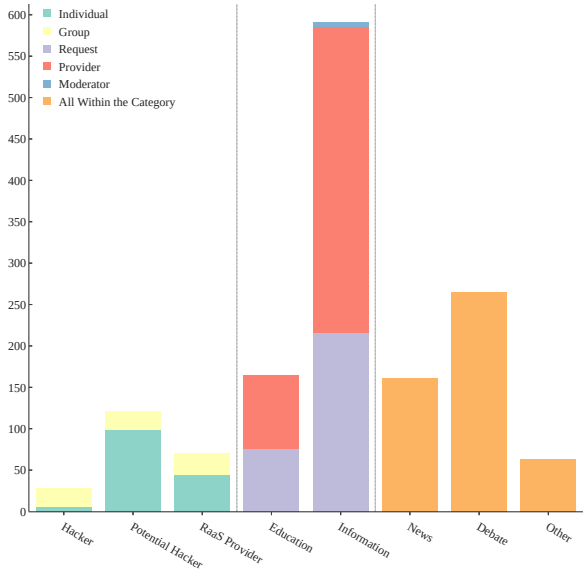

Fig. 1. A stacked bar chart showing the distribution of the eight categories of the ransomware-related posts in the Dread forum

Table 1 provides a detailed breakdown on how the 1,279 posts referencing ransomware topics are being grouped into the eight categories. In order to highlight the popularity of each of the categories more clearly, these 1,279 posts are also represented as a stacked bar chart in Figure 1. The posts for each category are also colour-coded to indicate their additional sub-groups, in order to provide further distinctions between the types of posts on the site. The goal was to document all our information in one structured location in order to allow easy modification, sharing, and analysis. Due to the large size of the dataset, we cannot show everything in this paper. However, interested readers can view a small snapshot, as well as the full set of the raw data at https://github.com/SocialSec2023-Paper-23/SocialSec-2023-Paper-23-Additional-Information.

5 Analysis and Discussion

Our research has shown that the acquisition and use of large dataset is extremely useful for behavioural science investigations. The diagram shown in Figure 1 provides us with a clear visual cue regarding the trend and the intents of the discussions around ransomware on the Dread forum, even though there are some variations in the posts collected (e.g., a post can fall into at most two categories).

5.1 Mapping the posts to categories

These 1,279 ransomware-related posts were split into the eight categories (as outlined in Section 3). In order to better understand their intent, five categories

(Hacker, Potential Hacker, RaaS Provider, Education and Information) were further arranged into up to three sub-categories, as summarised in Table 1. Sub-categories were not identified for the remaining three categories (News, Debate, Other) due to their nature of being too broad.

These sub-categories allow us to better understand the context of an otherwise broad label, and proceed with further analysis regarding the intent of each post. One example of this is "Education" which contains the sub-categories of "request" and "provider", indicating the user's intent to request or provide educational resources respectively. Furthermore, "Hacker" "Potential Hacker", and "RaaS Provider" contain the sub-categories "group" and "individual" to distinguish whether the user is alone or part of a group. This information gives us a deeper understanding of the nature of discussions on the forum.

5.2 Qualitative analysis of select categories

This analysis is useful as it directly allows us to see the most discussed themes during the aforementioned time frame. One immediate insight is the high number of posts where "Information" is requested, with the majority being part of the "provider" sub-category. This is expected, as information is arguably one of the most important tools for those interested in any cybercrime, including ransomware. Posts in this category range from requesting links to leak sites to various ransomware groups (see Quote 1), asking for information on how to access/spread ransomware – RaaS or otherwise (see Quote 2), and discussions around relevant topics at the time (see Quote 3). The quotes are shown below:

"anyone happen to have the onion link to the recent babuk ransomware breach?" [posted by: User 1] (Quote 1)

"can you tell me some gangs that offer raas services and how to contact them, please?" [posted by: User 2] (Quote 2)

"curious if anyone has any information on if revil was paid by kaseya or if they simply shut down to evade le. read today that kaseya was able to obtain a decryptor key from a third party, any thoughts?" [posted by: User 1] (Quote 3)

These posts fit the mindset of a person interested in ransomware – if we can assume that their aim is to become more educated and involved in this environment. This is supported by 22 posts being in both the "Education" and "Information" category. Understanding ransomware (and how to operate it) is not necessarily something that is simple to achieve, and hence people turn to forums such as Dread in order to research this further. In this case, they believe these forums are a place where they can gain this knowledge. Dread is considered easier to access, in comparison to other more specialised forums such as Exploit or Russian Anonymous Marketplace (RAMP). The lack of equivalent information on the clear web only fuels this movement to more specialised channels. This leads to an eager but primarily less knowledgeable group, forming an environment of like-minded individuals that highly encourages a large amount of questions and requests for information.

Notably, a large number of posts requesting information are those asking for RaaS. With RaaS – such as LockBit [12,22] and DarkSide [31] – being on the rise in recent years, it is expected that many conversations would be around such ransomware groups. Since RaaS removed the most technically challenging part in ransomware operation – namely writing the ransomware code – the existence of RaaS increases the accessibility of ransomware software to almost anyone. This ability for RaaS to be used by a large number of people makes it appealing to those who are interested to engage in cybercrime activities. One example of this is a post from a user (known as User 3) who was "*looking for the lockbit2.0 on dread*". Finally, the user friendly and easy-to-set-up nature of RaaS makes it ideal for newbies to pick up, especially when there are some support communities to learn how to use RaaS via discussion forums such as Dread. This popularity is reinforced by the number of posts occurring even after the soft ban on ransomware discussions following the Colonial Pipeline attack [6].

The category "Debate" also contains a significant number of the posts related to ransomware. One example is discussions regarding best practices; these posts often occur after an information or education request. We consider this category to be significant as it is important for users to share information in order to stay relevant and effective in their aim, especially due to the illegal nature of ransomware. The allure of a supportive community can attract individuals to a forum like Dread where they can connect with others who share common interests, especially on sensitive topics and potentially criminal subjects (in this case, ransomware). This is because these topics are usually not allowed on clear web forums. Regardless of whether a users' post is critical or constructive, it is more likely to be accepted on dark web forums. Furthermore, this debate could contribute to the building of trust between potential criminals and the emergence of more private criminal communities, as well as to encourage further learning. Overall, such debating interactions provide a social incentive to continue being a part of the ransomware community.

A large number of users debate the morality of using ransomware on anyone other then large corporations. This is shown in the forum with users calling those who disagree with this principle as "thieves" (and using other expletives). These companies concisely fit the criteria for being justifiable to become a target of ransomware, as described by the idea of Routine Activity Theory [26] (sensitive information, ability to pay high ransom etc.) in addition to being viewed as immoral or corrupt by many in the community. Because of this, many see it as their "duty" to attack these companies as a form of vigilante justice. This drives them to increase their skills and continue attacks. It is important to note that some in Dread disagree with this viewpoint. During the COVID-19 pandemic, many condemned the attacks on healthcare facilities, citing the "impact of it on people's lives" (posted by User 4) despite the valuable information that could be gained.

Posts related to "News" were also prevalent with 161 posts fitting this category. A majority of these messages (54.66%) were posted by a single user, showing a consistent news-like outlet. This access to real world events (which

Table 2. The frequencies and percentages of the keywords being found in ransomware-related posts

Keywords	Ransomware	RaaS	REvil	Extortion	DarkSide	LockBit	Sodinokibi	BlackCat	Hive	Avaddon	BlackMatter
Frequency	618	58	44	39	32	10	7	3	2	1	1
Percentage	75.83%	7.12%	5.40%	4.79%	3.93%	1.23%	0.86%	0.37%	0.25%	0.12%	0.12%

may not be presented as prominently on mainstream news sites) can embolden others, especially if attacks were successful. One prominent example of this is the Colonial Pipeline attack by DarkSide [14,31].

Finally, we would like to note that information sharing in Dread predominately follows the "horizontal communication" model [38]. Horizontal communication is when information is shared between people of the same level in a group. This system of communication works well in a public forum dedicated to similar topics. In addition, due to the illegal nature of ransomware and Dread as a forum, the necessity of protecting it from law enforcement (referred to as "le" in some posts) is paramount. Because of this "us vs. them" mentality, there is an incentive to share knowledge while trying to elude law enforcement and prevent exposure. Vertical communication does exist in the forum too. It involves a communication between "superiors" and "subordinates" and provides the forum with structure. Vertical communication was found within a small group of respected and knowledgeable individuals, with one example being a user who runs an extensive education course called "Hacktown" with many being "very impressed" with its contents. One considers the inspiration and specialised knowledge they and other notable users provide.

5.3 Analysis on the keywords

Table 2 presents descriptive statistics and frequencies of the keywords in the posts. The term "Ransomware" is the most frequent with 618 total hits (75.83% of the total keywords found). This was expected due to the purpose of this investigation. The same applies to the keywords "RaaS" and "Extortion" which were found 58 (7.12%) and 39 (4.79%) times respectively.

Keywords related to notable ransomware groups were also reasonably prevalent with "REvil" having 44 hits (5.40%) and "DarkSide" 32 hits (3.93%). These figures indicate these groups are being discussed more frequently. This lines up with notable attacks from these groups [21,31], which would encourage this discussion. In comparison, groups with less prominence – such as "Avaadon" (1 reference), "BlackMatter" (1) and "BlackCat" (3) – appeared less often.

Other keywords did not receive any references including "Cyberextortion" and "Cyber extortion". One reason for this may be that this term is more closely

associated with DDoS attacks, and thus is preferred not to be used in the context of ransomware. Another keyword ("Lockbit Black") should have already been covered by "LockBit". In addition, the keyword "Hive" (2 hits) – despite being linked to the name of a notable ransomware group – has strong connections to the drug market and "hive-mind" conspiracies. However, this is not confirmed. Refining our keywords will provide us with more accurate information.

5.4 Challenges and Limitations

The number of posts analysed is rather limited and future research should focus on collecting a greater amount of data and from a wider range of forums. One immediate limitation was due to the Dread forum being unavailable from 30 November 2022, caused by DDoS attacks against it. This made extraction of new posts through our web crawler impossible. The reduced time-frame led to a smaller dataset. Nonetheless, we managed to collect more than four years' worth of data, providing a good starting point to reveal some interesting insights into ransomware discussions on the dark web.

Because Dread is not a ransomware-specific forum, it does not attract many ransomware-experienced users, leading to the collection of a relatively small number of ransomware-related posts. This results in a loss of insights from those with more knowledge on the subject. Subsequently, we could use more specialised forums – e.g, RAMP – to further analyse why people engage in ransomware activities in the long term, allowing us to compare these "experts" against those less experienced users. For instance, User 6 mentioned that "*most of the active ransomware gangs now, conti, avos, pysa, grief, lockbit, sugar you can contact only on ramp*", which suggests potentially more revealing insights from RAMP. However, due to the secrecy of these forums, gaining access to them may be difficult. Furthermore, exploring other forums will bring its own challenges. Several forums (e.g., XSS, Exploit and RAID) have banned ransomware topics due to the increased surveillance from law enforcement after certain notable events, such as the Colonial Pipeline attack [6]. Whilst not fully enforceable, this ban may decrease the number of conversations related to ransomware in the future, limiting our dataset. This makes Dread one of the best options at this time.

Finally, despite using a wide range of techniques to achieve the large batch of information we have, this approach is still prone to potential faults. The ad-hoc crawler built for this project is in its early stages of development and therefore requires some refinement. For example, the crawler found 32 instances of the keyword "Darkside" whilst Microsoft Excel formulas found 44. As such, accuracy will need to be improved. Furthermore, the manual filtering of a large number of ransomware posts and categories by a single person did leave room for human errors and biases. This makes categories which have similarities – such as "Education" and "Information" – difficult to objectively separate. To deal with this issue, each category has been given a clear definition to ensure its consistent meaning and help with separation. Nonetheless, further improvement will be beneficiary, for example by employing automation.

6 Conclusion

We present the findings from a study in which ransomware-related discussions posted on a dark web forum called Dread were collected and analysed. Sixteen keywords were used to search for the pertinent ransomware-related posts, leading to eight main themes being identified in these posts: Hacker, Potential Hacker, RaaS Provider, Education, Information, News, Debate and Other.

Our analysis contributed to the growing body of evidence showing that ransomware is a topic of discussion on dark web forums. Our dataset covers a period of more than four years, providing useful social and technological insights into the prevalence and trends of ransomware-related discussions over time. On top of the quantitative indicators, the classification of the posts into four categories (Education, Information, News, and Debate) sheds further light into the nature of the interactions between dark web forum users. Further analysis could be conducted to infer the possible roles, status and influence of their authors.

For future work, the dataset can be expanded by including more keywords and more variations of ransomware terms, such as misspelling. Moreover, both clear web and other dark web forums – such as Russian Anonymous Marketplace (RAMP) and XSS – could be crawled to generate more data. In addition to the descriptive analysis done in this work, machine learning techniques can be employed to carry out predictive analysis. This dataset will be utilised as input for a machine learning-based system to create a model for classifying ransomware posts. This will contribute to automatic detection of such posts and could be used to prevent them from being posted on (legitimate) social networks.

References

1. Afroz, S., Garg, V., McCoy, D., Greenstadt, R.: Honor Among Thieves: A Common's Analysis of Cybercrime Economies. In: 2013 APWG eCrime Researchers Summit. pp. 1–11. IEEE (2013)
2. Afroz, S., Islam, A.C., Stolerman, A., Greenstadt, R., McCoy, D.: Doppelgänger Finder: Taking Stylometry to the Underground. In: 2014 IEEE Symposium on Security and Privacy. pp. 212–226. IEEE (2014)
3. Aslan, Ö.A., Samet, R.: A Comprehensive Review on Malware Detection Approaches. IEEE Access **8**, 6249–6271 (2020)
4. Bada, M., Pete, I.: An Exploration of the Cybercrime Ecosystem Around Shodan. In: 2020 7th international conference on internet of things: Systems, management and security (IOTSMS). pp. 1–8. IEEE (2020)
5. Bekkers, L., van't Hoff-de Goede, S., Misana-ter Huurne, E., et al.: Protecting your business against ransomware attacks? explaining the motivations of entrepreneurs to take future protective measures against cybercrimes using an extended protection motivation theory model. Computers & Security **127**, 103099 (2023)
6. Cimpanu, C.: Three Major Hacking Forums Ban Ransomware Ads as Some Ransomware Gangs Shut Down (2021), https://therecord.media/three-major-hacking-forums-ban-ransomware-ads-as-some-ransomware-gangs-shut-down
7. Connolly, L.Y., Wall, D.S.: The Rise of Crypto-Ransomware in a Changing Cybercrime Landscape: Taxonomising Countermeasures. Computers & Security **87**, 101568 (2019)

8. Continella, A., Guagnelli, A., Zingaro, G., De Pasquale, G., Barenghi, A., Zanero, S., Maggi, F.: ShieldFS: A Self-Healing, Ransomware-Aware Filesystem. In: Procs. 32nd Annual Conference on Computer Security Applications. pp. 336–347 (2016)
9. DarknetOnions: Dread DDOS Attack Continues, Onion Site Goes Offline (2022), https://darknetone.com/dread-ddos-attack-continues-onion-site-goes-offline/
10. Dread: Dread (2023), http://dreadytofatroptsdj6io7l3xptbet6onoyno2yv7jicoxknyazubrad.onion/
11. Du, P.Y., Zhang, N., Ebrahimi, M., et al.: Identifying, Collecting, and Presenting Hacker Community Data: Forums, IRC, Carding Shops, and DNMs. In: 2018 IEEE Int'l Conf. on Intelligence and Security Informatics (ISI). pp. 70–75. IEEE (2018)
12. Eliando, E., Purnomo, Y.: LockBit 2.0 Ransomware: Analysis of Infection, Persistence, Prevention Mechanism. CogITo Smart Journal **8**(1), 232–243 (2022)
13. Europol: Internet Organised Crime Threat Assessment (IOCTA) (2021), https://www.europol.europa.eu/cms/sites/default/files/documents/internet_organised_crime_threat_assessment_iocta_2021.pdf
14. FBI Press: FBI Statement on Network Disruption at Colonial Pipeline (2021), https://www.fbi.gov/news/press-releases/fbi-statement-on-network-disruption-at-colonial-pipeline
15. Flashpoint: Give Me Libre or Give Me Dread: The Fleeting Promise of Centralized Illicit Communities (2023), https://flashpoint.io/blog/libre-forum-centralized-illicit-communities/
16. Hernandez-Castro, J., Cartwright, A., Cartwright, E.: An Economic Analysis of Ransomware and Its Welfare Consequences. Royal Society open science **7**(3), 190023 (2020)
17. Holz, T., Engelberth, M., Freiling, F.: Learning More about the Underground Economy: A Case-Study of Keyloggers and Dropzones. In: 14th European Symp. on Research in Computer Security (ESORICS). Procs. 14. pp. 1–18. Springer (2009)
18. Huang, C., Guo, Y., Guo, W., Li, Y.: HackerRank: Identifying Key Hackers in Underground Forums. International Journal of Distributed Sensor Networks **17**(5), 15501477211015145 (2021)
19. Huang, D.Y., Aliapoulios, M.M., Li, V.G., et al.: Tracking Ransomware End-to-end. In: 2018 IEEE Symposium on Security and Privacy (SP). pp. 618–631 (2018). https://doi.org/10.1109/SP.2018.00047
20. Hull, G., John, H., Arief, B.: Ransomware Deployment Methods and Analysis: Views From a Predictive Model and Human Responses. Crime Science **8**, 1–22 (2019)
21. Kaseya Press: Kaseya Responds Swiftly to Sophisticated Cyberattack (2022), https://www.kaseya.com/press-release/kaseya-responds-swiftly-to-sophisticated-cyberattack-mitigating-global-disruption-to-customers/
22. Kaspersky: LockBit Ransomware — What You Need to Know (2022), https://www.kaspersky.com/resource-center/threats/lockbit-ransomware
23. Kolodenker, E., Koch, W., Stringhini, G., Egele, M.: PayBreak: Defense Against Cryptographic Ransomware. In: Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security. pp. 599–611 (2017)
24. Kouzis-Loukas, D.: Learning Scrapy. Packt Publishing Ltd (2016)
25. Lang, M., Connolly, L.Y., Taylor, P., Corner, P.J.: The Evolving Menace of Ransomware: A Comparative Analysis of Pre-pandemic and Mid-pandemic Attacks. ACM Digital Threats: Research and Practice (2022)
26. Leukfeldt, E.R., Yar, M.: Applying Routine Activity Theory to Cybercrime: A Theoretical and Empirical Analysis. Deviant Behavior **37**(3), 263–280 (2016)

27. McAlaney, J., Hambidge, S., Kimpton, E., Thackray, H.: Knowledge Is Power: An Analysis of Discussions on Hacking Forums. In: 2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). pp. 477–483. IEEE (2020)
28. McIntosh, T., Kayes, A., Chen, Y.P.P., Ng, A., Watters, P.: Ransomware Mitigation in the Modern Era: A Comprehensive Review, Research Challenges, and Future Directions. ACM Computing Surveys (CSUR) **54**(9), 1–36 (2021)
29. Meland, P.H., Bayoumy, Y.F.F., Sindre, G.: The Ransomware-as-a-Service Economy within the Darknet. Computers & Security **92**, 101762 (2020). https://doi.org/https://doi.org/10.1016/j.cose.2020.101762
30. Motoyama, M., McCoy, D., Levchenko, K., Savage, S., Voelker, G.M.: An Analysis of Underground Forums. In: Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. pp. 71–80 (2011)
31. Nuce, J., Kennelly, J., Goody, K., et al.: Shining a Light on DARKSIDE Ransomware Operations. Tech. rep., Mandiant (2021), https://www.mandiant.com/resources/blog/shining-a-light-on-darkside-ransomware-operations
32. O'Kane, P., Sezer, S., Carlin, D.: Evolution of Ransomware. Iet Networks **7**(5), 321–327 (2018)
33. Pastrana, S., Hutchings, A., Caines, A., Buttery, P.: Characterizing Eve: Analysing Cybercrime Actors in a Large Underground Forum. In: 21st International Symposium on Research in Attacks, Intrusions, and Defenses (RAID), Procs. 21. pp. 207–227. Springer (2018)
34. Pastrana, S., Thomas, D.R., Hutchings, A., Clayton, R.: CrimeBB: Enabling Cybercrime Research on Underground Forums at Scale. In: Proceedings of the 2018 World Wide Web Conference. pp. 1845–1854 (2018)
35. Pete, I., Hughes, J., Caines, A., Vu, A.V., Gupta, H., Hutchings, A., Anderson, R., Buttery, P.: POSTCOG: A Tool for Interdisciplinary Research into Underground Forums at Scale. In: 2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). pp. 93–104. IEEE (2022)
36. Pont, J., Abu Oun, O., Brierley, C., Arief, B., Hernandez-Castro, J.: A Roadmap for Improving the Impact of Anti-Ransomware Research. In: 24th Nordic Conference (NordSec). pp. 137–154. Springer (2019)
37. Pont, J., Arief, B., Hernandez-Castro, J.: Why Current Statistical Approaches to Ransomware Detection Fail. In: Information Security: 23rd International Conference (ISC 2020), Procs. 23. pp. 199–216. Springer (2020)
38. Ratten, V.: The Effect of Cybercrime on Open Innovation Policies in Technology Firms. Information Technology & People (2019)
39. Yilmaz, Y., Cetin, O., Grigore, C., Arief, B., Hernandez-Castro, J.: Personality Types and Ransomware Victimisation. ACM Digital Threats: Research and Practice (2022)
40. Yue, W.T., Wang, Q.H., Hui, K.L.: See No Evil, Hear No Evil? Dissecting the Impact of Online Hacker Forums. Mis Quarterly **43**(1), 73 (2019)
41. Yuryna Connolly, L., Wall, D.S., Lang, M., Oddson, B.: An Empirical Study of Ransomware Attacks on Organizations: An Assessment of Severity and Salient Factors Affecting Vulnerability. Journal of Cybersecurity **6**(1), tyaa023 (2020)
42. Zhang, Y., Fan, Y., Hou, S., Liu, J., Ye, Y., Bourlai, T.: iDetector: Automate Underground Forum Analysis Based on Heterogeneous Information Network. In: 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). pp. 1071–1078. IEEE (2018)
43. Zhao, Z., Ahn, G.J., Hu, H., Mahi, D.: SocialImpact: Systematic Analysis of Underground Social Dynamics. In: 17th European Symposium on Research in Computer Security (ESORICS). Procs. 17. pp. 877–894. Springer (2012)