I.J. Information Engineering and Electronic Business, **2024, 2, 54-67**
Published Online on April 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijieeb.2024.02.05

Dark Web Monitoring as an Emerging Cybersecurity Strategy for Businesses

Ashwini Dalvi*
Department of Computer Engineering, Veermata Jijabai Technological Institute, India
Email: aadalvi_p19@ce.vjti.ac.in
ORCID iD: https://orcid.org/0000-0001-9015-457X
*Corresponding Author

Sunil Bhirud
Department of Computer Engineering, Veermata Jijabai Technological Institute, India
Email: sgbhirud@ce.vjti.ac.in

Received: 24 June, 2023; Revised: 25 August, 2023; Accepted: 13 October, 2023; Published: 08 April, 2024

Abstract: The increasing frequency and sophistication of cyberattacks targeting institutions have necessitated proactive measures to prevent losses and mitigate damages. One of these measures is to monitor the dark web. The dark web is a complex network of hidden services and encrypted communication protocols, with the primary purpose of providing anonymity to its users. However, criminals use the dark web to sell stolen data, launch zero-day attacks, and distribute malware. Therefore, identifying suspicious activity on the dark web is necessary for businesses to counter these threats.

An analysis of dark web monitoring as an emerging trend in cyber security strategy is presented in this article. The article presents a systematic review of (a) why dark web surveillance enhances businesses' cybersecurity strategies, (b) how advanced tools and technologies are used to monitor dark web data in the commercial sector, (c) the key features of threat monitoring frameworks proposed by researchers, and (d) the limitations and challenges associated with dark web monitoring solutions. In summary, the proposed work involves analyzing various sources of information related to the topic and presenting a thorough assessment of the need and challenges of dark web surveillance to enhance the security measures of businesses.

Index Terms: Dark web, Monitoring Tools, Commercial cybersecurity solutions, Threat Intelligence framework.

1. Introduction

Cybercriminals increasingly target businesses to steal valuable data for financial gain or other malicious purposes. As a result, businesses in the modern age are increasingly facing threats from malicious cybercriminals who are constantly looking for ways to exploit their sensitive data and information. This malicious activity can harm businesses, being incredibly costly and potentially damaging their reputations. As such, it is paramount that businesses remain vigilant and take all necessary steps to ensure their data remains secure. In addition, better strategies can be developed to combat cyber attacks by systematically conceptualizing modern cyber attack platforms [1,2,3].

The term" dark web" refers to a part of the deep web comprising various darknets or private networks. Because they require specialized software (like the TOR browser), one can not access dark websites using standard web browsers. It is also more challenging to browse the contents of the dark web than on the public internet since it is encrypted and can not be indexed by conventional search engines. The dark web is home to various kinds of information, such as passwords, financial data, stolen credit cards, and software that can be used for attacks against vulnerable systems [4,5]. Moreover, the anonymity of this platform makes it the ideal destination for criminals who want to remain undetected while exchanging sensitive data or carrying out their malicious activities without facing any consequences.

Hackers use the dark web to share knowledge on hacking tools and techniques, create malware or viruses, buy or sell stolen credit cards and other personal data, plan cyberattacks against companies and governments, and launch ransomware attacks. The hacker community shares information and learns from each other constantly. For example, discovering cyber threat intelligence through conversations in online forums is possible [6]. Therefore, understanding how this community works can help organizations protect themselves from these threats. In addition, security professionals can gain valuable insights into the latest hacking attacks by learning about the latest hacking techniques,

tools, and vulnerabilities. This information allows defenses to be developed, and the threat landscape can be better understood.

Small and medium-sized enterprises lack the expertise and resources to implement adequate cybersecurity measures. However, these businesses can secure their corporate ecosystems and protect sensitive information by identifying pawned email accounts and other data breaches. Therefore, organizations must be aware of the risks the dark web poses and take proactive steps to monitor their information on it. Dark web monitoring helps organizations detect any threats or breaches associated with their data quickly and take action accordingly to protect their data from being misused or stolen.

Monitoring the dark web is a proactive way to identify if a company's data has been compromised and to prevent it from happening again. In addition, companies can use dark web monitoring tools to monitor the dark web for references to their organization's name or sensitive information, such as login credentials or financial information. The presented research article reviews dark web monitoring as an emerging trend in business cybersecurity strategy to investigate the potential benefits and limitations of using dark web monitoring to improve an organization's cybersecurity posture. This article provides an overview of the dark web and explains why enterprises must monitor it. It also evaluates the efficacy of dark web monitoring to identify potential information security hazards to an organization.

The proposed review examines commercial dark web monitoring solutions and a framework for threat monitoring proposed by researchers. The Systematic Literature Review (SLR) model is a systematic approach to conducting literature reviews, aiming to increase the review process's reliability and transparency. Traditional literature reviews can benefit from incorporating lessons from systematic reviews, such as mitigating bias, increasing transparency and objectivity, and critically appraising the evidence. By applying these principles, traditional reviews can improve the reliability of their findings, even when a full systematic review is not feasible.

The Systematic Literature Review (SLR) based approach is followed to present the review. The steps performed to present SLR are as follows:

- The research objectives the proposed study aims to attempt are the following:

Research Objective 1: "How does dark web surveillance enhance businesses' cybersecurity strategies?"

Research Objective 2: "How does the commercial sector monitor dark web data with advanced tools and technologies?"

Research Objective 3: "To collect, analyze, and share intelligence on criminal activities on the dark web, what are some of the key approaches proposed by researchers?"

Research Objective 4: "What limitations or challenges are associated with dark web monitoring solutions and threat monitoring frameworks, and can these be overcome?"

- Conduct a literature search to identify relevant dark web monitoring studies, reports, and other sources of information [7].

Researching the literature on dark web monitoring is essential for identifying relevant studies, reports, and sources of information. In online databases like Google Scholar, ACM, and IEEE digital libraries, relevant keywords can be searched using relevant terms. A search can also be conducted for publications and reports from reputable organizations and cybersecurity companies.

- Source selection and screening

After identifying sources, the next step is determining their relevance and quality. The inclusion and exclusion criteria should be clearly defined to ensure consistency and reduce bias in the screening process.

- Obtain and analyze data

The next step is to gather and analyze information from the selected sources. Then, thematic analysis approaches can be used to organize the data, and narrative summaries and tables can be used to summarize the results.

- Conclude the findings

At the end of the research process, conclusions are drawn from the findings.

The following paper is organized as follows: Section 2 of the paper outlines the methodology used for conducting this literature review. Then, the research question and key findings are discussed in Section 3. Finally, section 4 concludes the paper by discussing limitations and suggesting future research topics.

2. Research Methodology to Conduct a Literature Review

This section describes the Systematic Literature Review (SLR) method. There is literature on the dark web, but none addresses the research questions mentioned in this study. Researchers investigate recent studies involving the analysis of dark web content to gain cyber threat intelligence (CTI), explaining how they used techniques, methods, tools, approaches, and outcomes to analyze this content [8]. Another review comprehensively described dark web crime threats and the technical and forensic challenges [9]. Additionally, researchers reviewed how extremist/terrorist websites on the dark web are identified and handled [10].

The authors searched for studies examining dark web monitoring as a business cybersecurity strategy to address their research questions. The studies included in this review discussed the advantages and challenges of implementing dark web monitoring in business contexts and insights regarding threats on the dark web that can be monitored. As part of the study, the authors also searched for studies that explored how dark web monitoring could be integrated into existing cybersecurity frameworks and evaluated the effectiveness of dark web monitoring in identifying and mitigating cyber threats. Finally, the authors included studies on guiding businesses in selecting and implementing dark web monitoring tools, exploring the legal and ethical considerations associated with monitoring the dark web.

This review focused on dark web monitoring for cybersecurity strategy for businesses using the PRISMA flow diagram to identify and select relevant articles. Systematically and transparently, the PRISMA flow diagram helps identify and select articles relevant to a review.

Figure 1 shows the article searching process. As part of the first stage, searching relevant databases is essential to determine which databases are relevant to the article the researcher is looking for. Table 1 shows part of the research; the authors searched for studies on dark web monitoring as a cybersecurity strategy with the following keyword strings in the second stage:

Fig. 1. Article searching process

Table 1. Search keywords strings

Sr No	Search keywords strings
1	Dark Web monitoring
2	Cybersecurity Strategy
3	Cybersecurity Strategy for Business
4	Dark web data leak
5	Data breach on the dark web
6	Zero-day exploits on the dark web
7	Enterprise security vulnerabilities discussion on a dark web forum
8	Dark web forum for exploits
9	Dark web marketplace
10	Business security breaches on the dark web

Further, articles were screened for relevance to the review based on their titles and abstracts in the third stage. Figure 2 visually represents a set of article titles, where certain words have been emphasized, and emphasized words from article titles highlighted specific topics or themes discussed in the articles. When "cybersecurity" and "dark web" are highlighted in Figure 2, the articles likely focus on those subjects.

Figure 3 gives closer glimpses of selected articles by presenting a word cloud of the mentioned keywords in articles.

In the fourth stage, several criteria were used for selecting literature in this study as part of the inclusion principles focusing on dark web monitoring and business cybersecurity strategy.

Fig. 2. A word cloud of article titles

Fig. 3. A word cloud of keywords mentioned in articles

Table 2. Article Inclusion Principles

Sr No	Inclusion Principle
1	Literature relevant to the cybersecurity strategy of a business
2	Literature that addressed both the benefits and challenges of implementing dark web monitoring in business
3	Literature that discussed tools, technologies, and methodologies for monitoring dark web traffic, providing a comprehensive understanding of how dark web monitoring works and how to implement it
4	Literature demonstrating how businesses have successfully utilized dark web monitoring as part of their cybersecurity strategy
5	Language criteria – English
6	Publication years – 2001-2022
7	Article types – Journal and Conference publication, web articles

Table 3 presents exclusion criteria.

Table 3 Article Exclusion Principles

Sr No	Exclusion Principle
1	Literature not relevant to the topic being researched
2	Based on the source of information, literature from less creditable sources is excluded
3	Language criteria – Other than English
4	Before publication years – 2001-2022
5	Article types – Journal and Conference publication, web articles

Figure 4 depicts the research article screening process. First, articles were identified through search records in Google Scholar and other citation databases and merged into a citation management system. Then, for a comprehensive search, all relevant studies were included. Duplicate articles are removed after the initial search has been conducted. The digital object identifiers were used to remove duplicate records.

The third stage involved reading the remaining articles and determining whether they contributed to answering the research question. To determine whether articles meet the inclusion criteria, the authors screened the titles and abstracts of articles. An article that meets the inclusion criteria is screened for eligibility for the review through full-text screening. Lastly, irrelevant studies were excluded from the systematic review list at the final stage. By eliminating irrelevant studies and considering what was relevant, the authors determined the number of studies to be included in the systematic review.

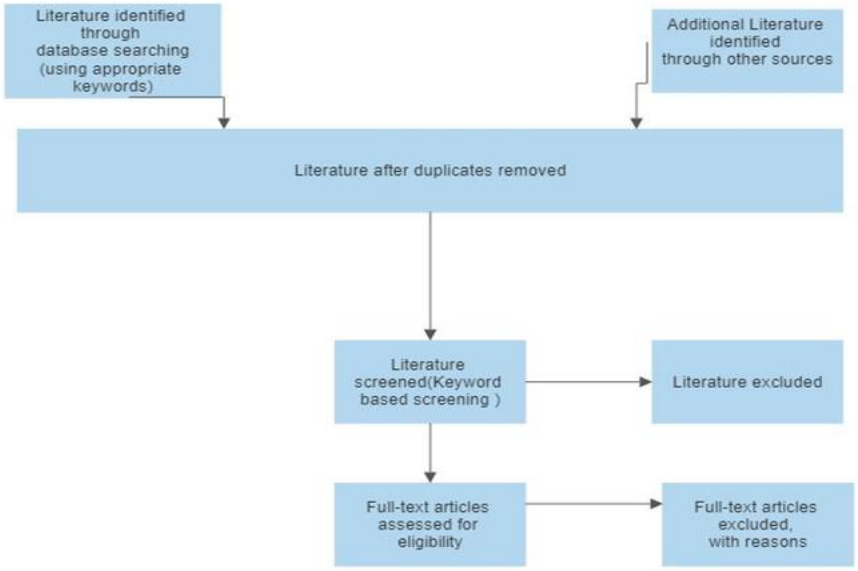

Fig. 4. Literature article screening process

Some significant findings emerged after reviewing the literature on dark web monitoring as an emerging trend in business cybersecurity. In recent years, there have been increasing research articles on this subject. Most publications from the last five years were visualized in a pie diagram, shown in Figure 5, indicating a growing interest in the field.

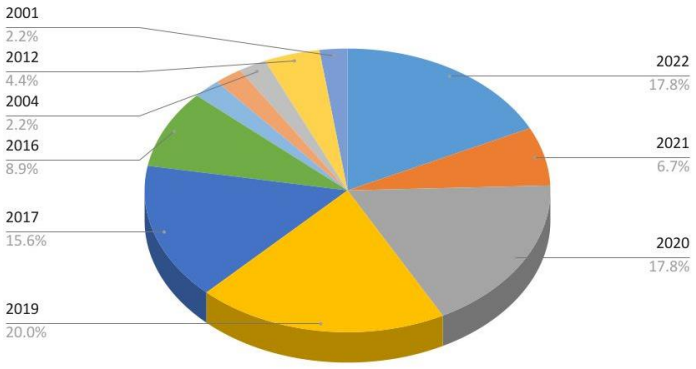

Fig. 5. Distribution of papers based on the year

A wide range of publishers and databases has published research on dark web monitoring. As a result, there is a multidisciplinary approach to understanding and dealing with the cybersecurity risks presented by the dark web, which is evident from the contribution of researchers from various academic backgrounds. As shown in the pie diagram visualization in Figure 6, IEEE, ELSEVIER, and Springer are among the top publishers in this area. Publishers like these are well known for generating high-quality academic publications and are frequently involved in engineering, computer science, and information technology research.

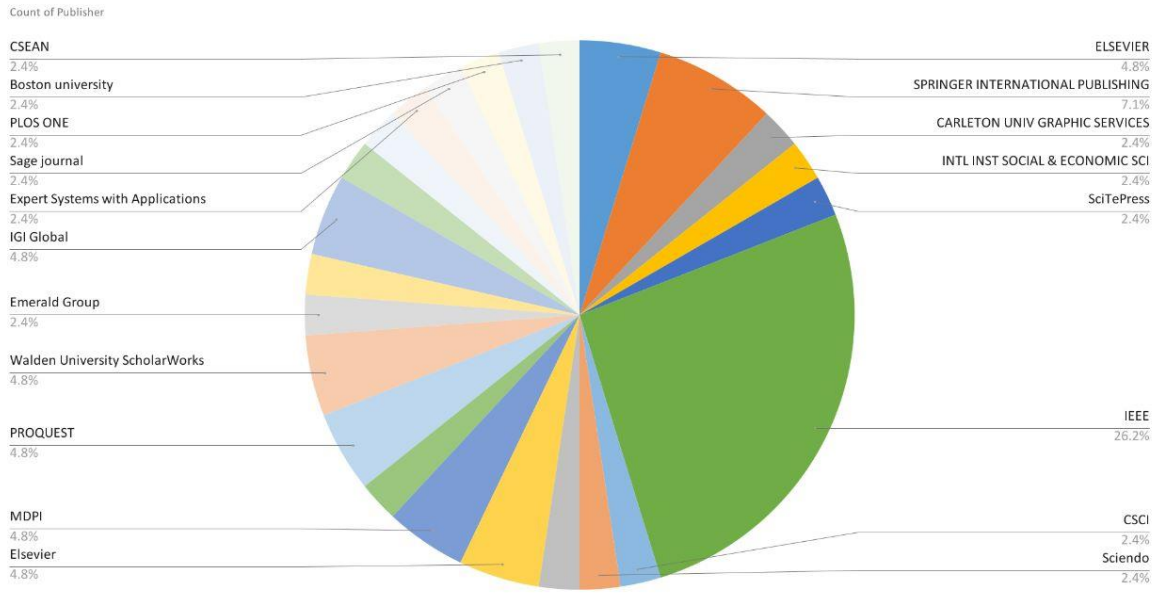

Fig. 6. Publisher-based distribution of papers

SLR analyses are conducted using rigorous and systematic methods, including identifying relevant databases, developing search strategies, screening titles and abstracts, evaluating full-text articles, analyzing data extraction and synthesis, and presenting the results.

3. Discussion on Research Question

A. Research question 1: " How does dark web surveillance enhance businesses' cybersecurity strategies? "

Competition is encouraged in a legal business environment if everyone follows the same governing rules. However, there are no rules on the illegal black market, and anyone can make money however they want. Particular geographic has always been like this; they have always been limited to a specific location. Nevertheless, the black market has expanded globally with the advent of the dark web [11]. Thus, whoever wants to participate in the black market can do so without following the laws. As a result, black markets on the dark web pose significant threats to individuals, governments, and businesses.

Illegal products, services, and content are traded on the dark web[12,13]. For example, drugs, weapons, counterfeit goods, and stolen information are all available for purchase and sale on the dark web. In addition, identities, financial fraud, and other illegal activities can be performed using stolen information, particularly among cybercriminals. Along with ransomware-as-a-service (RaaS), cybercriminals offer malware distribution services. In addition, data breaches or phishing attacks often lead to the sale of personal and financial information, such as credit card numbers and social security numbers.

Cyber attackers often use dark web marketplaces to trade their digital gains. They include vulnerability and exploitation tools, dumps, skimmers, identities, attack tools, mules, credit card information, fake tools, and Bitcoin, among other goods and services. A few marketplaces even allow single-vendor stores, where sellers sell their products on their websites. Buyers can create dark web marketplaces using a framework released as a "platform-as-a-service" in 2017 [14].

"As-a-service" cyber attacks have revolutionized cybercrime by making sophisticated and damaging attacks accessible to inexperienced attackers [15]. Thus, organizations and individuals must counteract the "as-a-service" model for cyber attacks. Furthermore, organizations must take strong security measures and stay informed of the latest threats to prevent cybercrime and protect businesses from its devastating effects. Consequently, monitoring corporate data and financial information on the dark web and keeping updates about recent cyber attacks has become crucial to many cyber threat intelligence operations.

Researchers assessed the ransomware-as-a-service (RaaS) economy in the dark web and found it a dangerous trend [16]. When performed by experienced attackers, ransomware poses a severe threat. This study used observation of forums, interviews, and available data to identify the value chains associated with the dark market. The paper presented a value chain map to understand the underground economy behind ransomware and mitigate its threat. Furthermore, the study highlighted how cryptocurrencies facilitate the development of the dark marketplace.

The financial sector is a prime target for cybercriminals because it contains sensitive data and money that can be stolen from individuals and organizations. With less knowledge on the part of end users, attackers' expertise, and readily

available tools and public data, ensuring the security of all online transactions is essential. Dark web activity and hidden websites often threaten financial institutions with cybercrime [17]. Many businesses are unaware of the wide range of activities on the dark web, including posting sensitive information that can be exploited for financial gain. Therefore, keeping confidentiality and preventing negative publicity are significant challenges for businesses. In addition, it is difficult for companies to track and prevent criminal activities on the dark web because criminals can access it easily and disappear without leaving much trace.

It is common for attackers to make their stolen data public, which helps reveal the current state of cybersecurity and potential threats. According to a surface and dark web dataset study, 71 cyberattacks were generated worldwide, with the US being the top target for state-sponsored hackers [18]. The study proposed the FinFrame framework as a four-step approach to secure financial organizations. Additionally, it suggested educating end users on cybersecurity. Research provided insights into the current financial climate and potential threats.

Organizations can gather insight into cybercriminal tactics, techniques, and procedures by monitoring dark web activities and the types of data and assets the cybercriminals are targeting [19]. It is possible to use this information to develop effective defensive strategies and improve an organization's security posture with the help of this information. A company can, for example, take steps to ensure it is not vulnerable to a particular attack if it sees it occurring more frequently on the dark web.

The research examined the extent to which malware is distributed on the dark web, the trends in malware types available, the methods for discovering new exploits, and the methods of discovering malware at the earliest possible stage using darknet analysis [20]. In addition, it discussed banking Trojans, ransomware, and remote access Trojans as examples of the types of malware available on the dark web. It also examined deceptive techniques such as honeypots, which organizations could employ to find new exploits.

Researchers examined the use of the dark web in cybercrime, including stolen datasets, compromised email accounts, and breached credentials, as well as hiring cyberattack botnets [21]. In addition, it discussed how to protect small and medium-sized businesses (SMEs) and microenterprises (MEs) against the threats posed by these activities. As part of this process, businesses must monitor the dark web for any mentions or discussions of their company or industry and stay up-to-date on any new attack methods or tools that cybercriminals use. In addition, according to the paper, businesses should prioritize protecting their email accounts since email is often the primary way cybercriminals gain access to sensitive data.

Because healthcare holds sensitive information, including protected healthcare information (PHI), it is a high-value target for cyberattacks [22]. A breach of patient privacy affects not only the privacy of patients but also the reputation and ability to operate healthcare organizations. Unfortunately, healthcare entities frequently lack the resources and expertise to keep up with the evolving threat landscape, putting them at risk. In addition to underinvesting, malicious actors have been attracted by PHI's high value and have been searching for ways to steal and sell it on the dark web. The loss, theft, or disclosure of sensitive healthcare information can result from insider attacks and external cybercriminals. In the dark web, stolen healthcare data can be sold for hundreds of dollars per patient record, with a complete record fetching thousands of dollars. As a result, cyberattacks are accelerating, and the healthcare sector needs to take proactive and responsive measures to counter them.

Using the dark web, the authors of [23] analyzed the cybersecurity landscape and identified examples of data dumps available for sale or download. The authors proposed a solution that involves monitoring social sites and hidden services on the dark web and analyzing their content semantically. The proposed solution aims to provide a proactive approach to cybersecurity by monitoring the dark web for potential threats. Organizations can gain valuable insights regarding the evolving cybersecurity landscape on the dark web by monitoring social media and hidden services and performing semantic analysis of retrieved content.

Large companies risk data breaches caused by malicious software attacks executed by hackers. Therefore, the visibility of hacker communities and underground markets is critical to improving businesses' cybersecurity. Organizations can stay updated on the latest cybercriminal threats and attacks by monitoring these communities. Furthermore, according to the researchers, companies could use the framework developed in the study to extract assets and appropriate functions from hacker forums [24]. The study claimed a valuable contribution to cybersecurity by using hacker assets as a resource for analysis and would assist in developing more effective defenses for businesses.

To understand the market for crimeware, the researchers identify three essential players [25]. Aside from anonymous e-currencies and anonymity networks, the authors also discuss mobile computing as a facilitator of crimeware marketplaces. Crime marketplaces have grown due to these technologies, which have made it easier for cybercriminals to operate. However, law enforcement agencies face many challenges, and the researchers suggest that private-public cooperation is essential to combat cybercrime. According to the study, adapting to new technologies by law enforcement and businesses is also essential in combating cybercrime.

Researchers explore the dark web's role in providing insight into malicious hacking communities and examine the emerging field of cyber threat intelligence [26]. The researchers analyze dark web data using human and automated techniques. Cybersecurity professionals and researchers could use this information to develop better strategies to protect their networks and solutions. Therefore, dark web forum surveillance provides valuable information about cyber threats and potential attacks, which can be used to enhance businesses' cybersecurity strategies.

Researchers mentioned that cyberattack trends could be understood by monitoring the dark web [27]. Therefore, researchers proposed an AI web-content analyzer to monitor underground marketplaces, especially the dark web, where

items used in cyberattacks are often traded illegally, such as exploits and malware. The researchers recommended a two-step procedure for monitoring the dark web. The first step was to create a Tor crawler to collect product data from AlphaBay, the largest marketplace. Additionally, Latent Dirichlet Allocation (LDA) was used to identify societal trends and the most prevalent outcomes of cyberattacks. According to researchers, Law enforcement could benefit from their work by identifying cyberattack products and tracking their sales. In addition, cybersecurity companies can benefit from their expertise in developing products to protect against cyberattacks.

Cyber threat intelligence (CTI) is essential to identify and mitigate potential cyber threats. Therefore, using dark web monitoring to generate CTI has become an upcoming cybersecurity trend. To provide timely, relevant, and actionable information about emerging threats, cybersecurity companies like Deepwatch, DarkOwl, Norton Dark Web Monitoring, and Ecosec Beacon are developing CTI capabilities with dark web investigation [28]. Using advanced technologies, these companies collect data from various sources, including the dark web, and then analyze and interpret it for intelligence.

As far as cybersecurity is concerned, the dark web presents both opportunities and risks for businesses. This platform can be used for illegal activities, including selling stolen data and counterfeit documents and orchestrating cyber attacks. However, it can also provide valuable threat information to businesses and security professionals [29]. In addition, businesses can identify potential threats and vulnerabilities by monitoring the dark web for emerging threats and vulnerabilities.

B. Research question 2: "How does the commercial sector monitor dark web data with advanced tools and technologies?"

The research for this question is based on web-based articles. Commercial industries offering paid services to monitor specific data on the dark web are on the rise. One of the standard features of dark web monitoring services is searching the dark web and alerting the organization about the spread of data breaches or potential threats by curating intelligence collected from the dark web. Following are a few such commercial solutions that refer to dark web monitoring features:

- DarkOwl

Commercial solutions like DarkOwl Vision offer a proactive strategy of entity searching, monitoring, and tracking to identify and assess the threats in the marketplace or environment to provide additional security measures if required for prevention and cybersecurity defenses [30].

- Ecosec Beacon

It is a web-based threat and risk intelligence solution that rapidly detects critical online content [31]. Echosec Beacon's protection mechanisms are proactive and reactive, depending on the functionality. Reactive measures include asset tracking 24 hours a day, detecting and monitoring emerging threats, and recognizing potential threats. In addition, it assists in comprehending risks, intruder motivations, and risk reduction.

- Alert Logic Dark Web scanner

Alert Logic proactively monitors the dark web for stolen credentials and notifies users when stolen credentials are found so that they can take immediate action [32]. It thus allows users to fight online fraud by informing them when accounts have been compromised and contextualizing attacks in the Alert Logic portal. In addition, the tool shuts down account takeover attacks by taking action when compromised accounts are identified. Part of Alert Logic's comprehensive security solution provides Vulnerability Assessments, Threat Detection and Response, and Web Application Security.

- ACID Cyberintelligence

Real-time monitoring of social networks, criminal sites, and the deep and dark web for suspected threats such as leaked usernames and passwords, emails, and financial information is provided by ACID Cyberintelligence [33]. It also has an easy-to-search interface with unlimited keywords to help find breached information, and the results can be translated into any language. Real-time alerts and a smart and friendly dashboard are the most attractive features of this tool.

- Norton Dark Web Monitoring

It searches for breached data on the dark web [34]. If users' relevant information is found, the user is notified to take immediate actions, such as updating passwords for logging in to the account. Norton is an antivirus and antimalware software product with a dark web monitoring feature that looks for breached data to verify whether the user's information is present in the breached data to alert the user. The user should take immediate action in response to the breach of sensitive information, such as a login or bank password. Users can turn to two-factor authentication and build more secure and complex passwords, including lowercase and uppercase alphabets and special characters.

The summary features of the tools mentioned above are presented in the following table 4:

Table 4. Scope of dark web monitoring tools

Sr No	Tool name Mention of dedicated dark	Mention of dedicated dark web/ darknet module	The form of information delivery mode
1	Dark Owl	Yes	Search Engine/ REST AP
2	Echosec Beacon	Yes	Search Engine
3	Alert Logic Dark Web scanner	Yes	Account Takeover solution, managing passwords and access keys
4	ACID Cyberintelligence	Yes	Data Dumps finder
5	Norton Dark Web Monitoring	Yes	Data Dumps finder
6	Matchlight Terbium Labs	Yes	Data Dumps finder

Monitoring the dark web with commercial solutions is on the rise, allowing proactive entity search, monitoring, and threat tracking. It is possible to detect and monitor emerging threats, recognize potential threats, and notify users of compromised credentials with dark web monitoring solutions. Social networks, criminal sites, and the deep and dark web can all be monitored for suspected threats using these tools, which offer easy-to-use interfaces, unlimited keyword searches, and intelligent dashboards. As a result of these solutions, businesses can aim for additional security measures to enhance cybersecurity defenses and prevention measures.

C. Research question 3: " To collect, analyze, and share intelligence on criminal activities on the dark web, what are some of the key approaches proposed by researchers?"

The dark web has become a vast data repository that requires automated content analysis to monitor and analyze. To automate the process of crawling, indexing, and classifying content within the hidden services ecosystem, researchers have developed numerous frameworks and tools. Dark web data is analyzed using advanced machine learning techniques like deep learning and natural language processing. Using these approaches, researchers have identified trends, patterns, and relationships between dark web actors and activities.

Researchers introduced a data analysis framework called ATOL for deriving thematic labels to analyze the content of onions crawled [35]. OnionCrawler is an automated tool for crawling and analyzing content within the Tor hidden service ecosystem. Thousands of pages have been crawled and classified by the system over the two years. In addition, researchers mentioned that A keyword weighting scheme called TFICF combined with supervised machine learning classification algorithms performed 12% better than a keyword-based baseline algorithm used by analysts. Finally, researchers commented that several methods could be utilized to enrich keywords, such as multiclassifier analysis, theme learning, graph analysis, and thematic census mining. The paper makes a valuable contribution to automated content analysis in the Tor hidden service ecosystem and suggests promising directions for future research.

The BlackWidow system is designed to gather cyber security information from the dark web [36]. An analytics framework based on a knowledge graph collects data from various sources and fuses it with a single analytics framework. The system uses a Docker-based microservice architecture, combining preexisting and customized machine learning tools. Within less than two days of monitoring seven popular dark web services, BlackWidow collected years of pertinent cybersecurity and fraud monitoring information. In addition, the system can detect trends in topics related to cybersecurity and identify relationships between authors of posts and forums.

The article describes a research study to determine the causal relationship between street crime and the closure of dark web marketplaces[37]. Researchers use a regression discontinuity design to compare crime rates between days before and after dark web marketplaces were shut down. These findings suggested that shutting down dark web marketplaces does not affect street crime permanently. It was hypothesized that street drug dealers could quickly use another dark web marketplace after one closes. Additionally, researchers suggested that street drug dealers could find other means of obtaining drugs, including buying them from other street dealers or growing them.

According to researchers, ULMFit, Bert, and RoBERTa are pre-trained transferable models that could be used to classify illicit activities on the dark web [38]. In comparison with traditional text classification methods such as LSTM, the researchers found that Bert obtained the best accuracy in classifying dark web drug content and types. The researcher approach is significant since law enforcement agencies can identify and track illicit activities on the dark web without manually labeling data. There are, however, some limitations to the mentioned approach. This study's dataset is limited and may not fully represent all onion sites and dark marketplaces.

Unlike traditional CTI, reactive and data-driven, hackers on the dark web can provide valuable proactive CTI services that alert organizations to unknown threats. In addition, hacker forums on the dark web are rich in metadata, with tens of thousands of freely available tools, techniques, and procedures. Therefore, the studies designed a web crawler to collect hacker exploits continuously.

Researchers designed a crawler that can provide an additional layer of protection against the anti-crawling mechanism to collect hacker exploits on an ongoing basis as part of this exploratory study [39]. To classify exploits automatically into predefined categories on the fly and to develop interactive visualizations for proactive, timely CTI, the study employed a state-of-the-art deep learning approach called Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN). According to the study, system and network exploits were significantly more commonly shared than others.

The article discussed cybercriminal shared exploits and relevant tools through online hacker communities to launch cyber attacks against businesses[40]. Organizations could gain a proactive approach to cyber threat intelligence with the proposed text-mining framework that leverages hacker communication's explicit and implicit features in online hacker communities. Furthermore, utilizing machine learning algorithms, the framework could provide actionable insights regarding hacker expertise to mitigate risks associated with detected cyber threats.

The research described the development of cross-lingual hacker asset detection (CLHAD), a software application that detects hacker assets on dark web platforms that are not in English [41]. A major challenge in analyzing foreign-language dark web content has been the lack of human-labeled training data. However, CLHAD utilizes generative adversarial networks (GANs) to generate multilingual text representations with the aid of Adversarial Deep Representation Learning (ADREL). Researchers evaluated CLHAD on Russian, French, and Italian dark web platforms and demonstrated its utility in hacker asset profiling, which provides managers with operational and strategic insights. According to the study, Several dominant dark web languages are prevalent on the dark web, but cybersecurity managers would benefit from focusing on Russia to identify sophisticated hacking assets.

Researchers proposed an exploit-vulnerability attention model based on deep learning to link exploits from the dark web to vulnerabilities [42]. In addition, a device vulnerability severity metric is being developed to assist cybersecurity professionals in prioritizing their efforts. Based on two case studies of CTI, the proposed model was compared against state-of-the-art non-DL and DL-based methods. Researchers commented that with cyberattacks becoming more costly, solutions like the EVA-DSSM and DVSM have important implications for security operations centers, incident response teams, and cyber vendors.

It is crucial to understand dark web marketplaces to understand what kinds of goods and services cyber attackers trade and how they conduct their illegal activities. Therefore, a study was conducted to develop techniques for analyzing images on dark web marketplaces [43]. For dark vendor profiling (DVP), the researchers collected image-based intelligence using various image hashing techniques and metadata. Vendors who conduct business across multiple dark net marketplaces were identified through this profiling, and their aliases were determined. Researchers obtained valuable information by investigating top vendors, markets, and hash analyses. The research investigated how image analysis can provide insight into dark web activity detection and profiling.

D. Research question 4: "What limitations or challenges are associated with dark web monitoring solutions and threat monitoring frameworks?"

There are several reasons why monitoring the dark web is a challenging task. First, anonymity and encryption of the dark web network present a significant challenge to dark web monitoring [44]. By routing internet traffic through multiple servers, Tor and other encryption tools allow users on the dark web to conceal their identities and activities. Despite this, dark web users still have their identities linked to their Bitcoin addresses by postings on forums, tweets, and other online information. The study found that users who had used Bitcoin to transact on the dark web could be retroactively identified using the blockchain and publicly available information [45]. Even though the users had used encryption tools to conceal their identities, their Bitcoin transactions remained vulnerable to identification. Thus, to work within the constraints of the dark web, monitoring tools must be explicitly designed for it.

Monitoring the dark web is also challenging because of the volume of data. In addition, there is a constant flow of new sites and services on the dark web; dark web marketplaces and forums are no exception. As a result, manually monitoring dark websites is nearly impossible, and even with automated tools, it is hard to stay on top of the ever-changing landscape. However, a monitoring process can be automated using artificial intelligence (AI) and machine learning (ML). In addition, dark web data can be analyzed using AI and ML to detect patterns, anomalies, and potential threats. By proactively identifying potential threats, businesses and organizations can mitigate them in advance.

One challenge of using AI and ML for dark web monitoring is the availability of labeled data. A labeled data set is essential for training AI and ML models to detect patterns and anomalies. Here, labeled data refers to data classified and tagged manually by humans. For example, labeled data would be data that has been classified as having cyber threats, such as stolen information, hacking tools, or malware, as part of dark web monitoring.

For content classification, the researchers used legal documents to identify relevant terms [46]. Then, based on these terms, they trained their system to automatically identify and classify dark web content using supervised learning models. The study demonstrated that dark web content could be classified as illegal and harmful. Nevertheless, training a model can still be challenging without predefined terms for identifying content. The terms may not capture variations and nuances in the content. Additionally, illegal activities and content on the dark web are constantly evolving.

Creating and maintaining a dataset to monitor the dark web is intricate. In addition, maintaining an up-to-date dark Web dataset is difficult because the content and structure are constantly changing. Furthermore, it is challenging to identify cyber threats communicated in hacker forums and social media due to some posts' technical language and unclear meanings [47]. For example, the language used on these platforms may be unfamiliar to those unfamiliar with hacker or cybersecurity jargon. Moreover, some posts use coded language or are ambiguous, making it difficult to understand their true intentions. It is, therefore, necessary to have specialized knowledge and expertise in cybersecurity to detect potential cyber threats on these platforms.

A possible solution to this challenge involves manually labeling the data to create a reliable dataset. Researchers manually labeled the data to create a reliable dataset for classifying content. For example, researchers manually labeled the dark web forum dataset [48]. An SVM classifier was used to classify the content using a modified frequency-based

term weighting scheme. However, this approach had limitations since it used existing data for content classification, and manual data labeling was used to train the model.

Researchers gathered a large dataset of 232,792 pages containing illicit and banal contents from the Tor dark web[49]. In particular, the researchers noted that existing conceptual models and labeling schemas were unsuitable for law enforcement purposes. To address this problem, they manually labeled over 4000 unique Tor pages. There are other examples where researcher resorts to annual labeling for dark web content classification [50,51].

There are also challenges associated with the lack of standardization in the content available on the dark web. As a decentralized network, dark web content is not regulated by any governing body. Therefore, dark web monitoring tools should be flexible enough to handle various content types and formats. With a focus on language, the study sought to understand the spread of information on the dark web [52]. Diverse research questions were examined in the study, including the diversity of Tor services and information, the importance of centralizing core services, and the insulation among Tor services.

Researchers have attempted to categorize dark web content using image analysis techniques in addition to text-based analysis. For instance, dark web images might be categorized into five categories using Compass Radius Estimation for Image Classification (CREIC) [53]. In addition, researchers attempted to enhance machine learning algorithms, including image classification on the dark web, by quantum computing [54].

Additionally, dark web monitoring poses significant ethical and legal challenges [55,56]. For example, user privacy is at risk when monitoring the dark web. Therefore, monitoring the dark web must be done carefully and ethically while safeguarding individual rights.

A final challenge is the high cost of dark web monitoring tools. Several factors can influence the cost of dark web monitoring tools, including the type of tool, the level of security, and the amount of data being monitored [57]. In addition, developing and maintaining specialized tools and expertise for dark web monitoring can be expensive. As a result, it may be difficult for small businesses and individuals to invest in these tools, leaving them vulnerable to cyber threats.

4. Conclusion

With the growth of the dark web over the years, cybercriminals can conduct their activities anonymously. In addition, cybersecurity professionals cannot identify and track perpetrators using encryption technologies and hacking codes, facilitating anonymity. The dark web has become harder to control and monitor owing to this. Consequently, cybercriminals increasingly use the dark web to conduct their criminal activities. In addition, illegal activities, such as selling drugs and weapons, are common on the dark web. Data stolen from businesses, such as login credentials, social security numbers, and credit card numbers, is sold, bought, and shared on certain dark websites and forums. Despite government and cybersecurity efforts, taking down the dark web remains challenging.

For businesses, monitoring the dark web is one of the most critical aspects of cybersecurity. Cyberattacks and data breaches are becoming more prevalent, so businesses must stay vigilant in preventing sensitive information from being sold on the dark web. Dark web monitoring services typically scan the dark web for data related to a specific individual or organization using specialized software and tools. For example, it could scan for login credentials, personal identification, and financial information. Companies can use these tools to stay informed about potential threats and take action before a data breach occurs.

Also, monitoring services can track and monitor activity on dark web forums and marketplaces. Businesses can gain valuable insights into cyber threats and vulnerabilities by monitoring dark web forums and marketplaces. Often, cybercriminals use these forums and marketplaces to trade illegal goods and services, such as stolen data, malware, and hacking tools. Therefore, an organization can identify specific threat actors and cybercriminal groups target by monitoring Dark Web forums and marketplaces. As a result of this information, cybersecurity strategies can be improved, such as implementing more robust access controls and security protocols. For example, a business can restrict access to its network or limit its ability to steal data if it can identify a specific group that is targeting it.

Analyzing dark web activity can help businesses develop and implement new security measures ahead of the curve. For example, when a business notices increased sales of a particular malware type on a dark web marketplace, it can prevent that malware from entering its network before it becomes a problem.

In conclusion, by monitoring the dark web, businesses can improve their cybersecurity strategies, and commercial industries can help with this by offering advanced tools and technologies. In addition, the review examined some approaches researchers have proposed to collect, analyze, and share information about dark web activities. Nevertheless, dark web monitoring solutions and threat monitoring frameworks face some limitations and challenges. Businesses need to monitor the dark web as part of their cybersecurity strategies, but it must be done using the right tools and techniques.

References

[1] Huang, Keman, Michael Siegel, and Stuart Madnick. "Systematically understanding the cyber attack business: A survey." ACM Computing Surveys (CSUR) 51, no. 4 (2018): 1-36.

[2] Leszczyna, Rafał. "Review of cybersecurity assessment methods: Applicability perspective." Computers & Security 108 (2021): 102376.

[3] Cascavilla, Giuseppe, Damian A. Tamburri, and Willem-Jan Van Den Heuvel. "Cybercrime threat intelligence: A systematic multi-vocal literature review." Computers & Security 105 (2021): 102258.
[4] Kaur, Shubhdeep, and Sukhchandan Randhawa. "Dark web: A web of crimes." Wireless Personal Communications 112 (2020): 2131-2158.
[5] Zenebe, Azene, Mufaro Shumba, Andrei Carillo, and Sofia Cuenca. "Cyber threat discovery from dark web." EPiC Series in Computing 64 (2019): 174-183.
[6] Sarkar, Soumajyoti, Mohammad Almukaynizi, Jana Shakarian, and Paulo Shakarian. "Predicting enterprise cyber incidents using social network analysis on dark web hacker forums." The Cyber Defense Review (2019): 87-102.
[7] Rowley, Jennifer, and Frances Slack. "Conducting a literature review." Management research news 27, no. 6 (2004): 31-39.
[8] Basheer, Randa, and Bassel Alkhatib. "Threats from the dark: a review over dark web investigation research for cyber threat intelligence." Journal of Computer Networks and Communications 2021 (2021): 1-21.
[9] Nazah, Saiba, Shamsul Huda, Jemal Abawajy, and Mohammad Mehedi Hassan. "Evolution of dark web threat analysis and detection: A systematic approach." IEEE Access 8 (2020): 171796-171819.
[10] Alghamdi, Hanan, and Ali Selamat. "Techniques to detect terrorists/extremists on the dark web: a review." Data Technologies and Applications (2022).
[11] Mazi, Hilary & Mailewa, Akalanka & Ngniteyo, Arsene "The Influence of Black Market Activities through Dark Web on the Economy: A Survey.". (2020).
[12] Easttom, Chuck. "Conducting investigations on the Dark Web." Journal of Information Warfare 17, no. 4 (2018): 26-37.
[13] Hurlburt, George. "Shining light on the dark web." Computer 50, no. 04 (2017): 100-105.
[14] DEEPWEBADMIN. 2017. Build a black market in dark web only for $4500; Cybercrime goes PAAS. Visited on 30th March 2023 https://www.deepweb-sites.com/build-black-market-dark-web-4500-cybercrime-goes-paas/.
[15] Huang, Keman, Michael Siegel, and Stuart Madnick. Cybercrime-as-a-service: identifying control points to disrupt. Vol. 17. (2017).
[16] Meland, Per Håkon, Yara Fareed Fahmy Bayoumy, and Guttorm Sindre. "The Ransomware-as-a-Service economy within the darknet." Computers & Security 92 (2020): 101762.
[17] Camillo, Mark. "Cybersecurity: Risks and management of risks for global banks and financial institutions." Journal of Risk Management in Financial Institutions 10, no. 2 (2017): 196-200.
[18] Hossain, Md Jafrin, Umme Nusrat Jahan, Rejuan Haque Rifat, Annajiat Alim Rasel, and Muhammad Abdur Rahman. "Classifying Cyberattacks on Financial Organizations Based on Publicly Available Deep Web Dataset." In 2023 International Conference On Cyber Management And Engineering (CyMaEn),. IEEE, (2023): 108-116.
[19] Mador, Ziv. "Keep the dark web close and your cyber security tighter." Computer Fraud & Security 2021, no. 1 (2021): 6-8.
[20] Burgess, Jonah. "Malware and Exploits on the Dark Web." arXiv preprint arXiv:2211.15405 (2022).
[21] Pantelis, George, Petros Petrou, Sophia Karagiorgou, and Dimitrios Alexandrou. "On Strengthening SMEs and MEs Threat Intelligence and Awareness by Identifying Data Breaches, Stolen Credentials and Illegal Activities on the Dark Web." In Proceedings of the 16th International Conference on Availability, Reliability and Security, (2021): 1-7.
[22] Swasey, Katelyn. "Insufficient healthcare cybersecurity invites ransomware attacks and sale of phi on the dark web." Center for Anticipatory Intelligence Student Research Reports (2020).
[23] Akintaro, Mojolaoluwa, Teddy Pare, and Akalanka Mailewa Dissanayaka. "Darknet and black market activities against the cybersecurity: a survey." In The Midwest Instruction and Computing Symposium. (MICS), North Dakota State University, Fargo, ND. (2019).
[24] Samtani, Sagar, Ryan Chinn, and Hsinchun Chen. "Exploring hacker assets in underground forums." In 2015 IEEE international conference on intelligence and security informatics (ISI). IEEE, (2015): 31-36
[25] Gad, Mahmoud. "Crimeware marketplaces and their facilitating technologies." Technology Innovation Management Review 4, no. 11 (2014).
[26] Godawatte, Kithmini, Mansoor Raza, Mohsin Murtaza, and Ather Saeed. "Dark web along with the dark web marketing and surveillance." In 2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), IEEE, (2019): 483-485.
[27] Gupta, Abhineet, Sean B. Maynard, and Atif Ahmad. "The dark web phenomenon: A review and research agenda." (2019).
[28] Dalvi, Ashwini, Gunjan Patil, and S. G. Bhirud. "Dark Web Marketplace Monitoring-The Emerging Business Trend of Cybersecurity." In 2022 International Conference on Trends in Quantum Computing and Emerging Business Technologies (TQCEBT), IEEE, (2022): 1-6.
[29] Miloshevska, Tanja. "Dark web as a contemporary challenge to cyber security." Kriminalističke teme–Časopis za kriminalistiku, kriminologiju i sigurnosne studije 19, no. 5 (2019): 117-128.
[30] Dark owl monitoring. https://www.darkowl.com/ Accessed on 30th March 2023.
[31] Echosec Beacon https://flashpoint.io/platform/echosec/ Accessed on 30th March 2023.
[32] Alert Logic Dark Web scanner. https://www.alertlogic.com/resources/webinars/addressing-vulnerabilities-and-threats-from-dark-web-attacks/ Accessed on 30th March 2023.
[33] Acid Cyberintelligence. https://www.acid-tech.com/ Accessed on 30th March 2023.
[34] Norton DRK Web Monitoring https://us.norton.com/feature/dark-web-monitoring Accessed on 30th March 2023.
[35] Ghosh, Shalini, Ariyam Das, Phil Porras, Vinod Yegneswaran, and Ashish Gehani. "Automated categorization of onion sites for analyzing the darkweb ecosystem." In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2017): 1793-1802
[36] Schäfer, Matthias, Markus Fuchs, Martin Strohmeier, Markus Engel, Marc Liechti, and Vincent Lenders. "BlackWidow: Monitoring the dark web for cyber security information." In 2019 11th International Conference on Cyber Conflict (CyCon), vol. 900, IEEE, (2019): 1-21.
[37] Zambiasi, Diego. "Drugs on the web, crime in the streets. the impact of shutdowns of dark net marketplaces on street crime." Journal of Economic Behavior & Organization 202 (2022): 274-306.
[38] Catolino, Gemma, "Illicit Darkweb Classification via Natural-language Processing: Classifying Illicit Content of Webpages based on Textual Information", ISBN 978-989-758-590-6, (2022)

[39] Williams, Ryan, Sagar Samtani, Mark Patton, and Hsinchun Chen. "Incremental hacker forum exploit collection and classification for proactive cyber threat intelligence: An exploratory study." In 2018 IEEE International Conference on Intelligence and Security Informatics (ISI), IEEE (2018): 94-99.
[40] Biswas, Baidyanath, Arunabha Mukhopadhyay, Sudip Bhattacharjee, Ajay Kumar, and Dursun Delen. "A text-mining based cyber-risk assessment and mitigation framework for critical analysis of online hacker forums." Decision Support Systems 152 (2022): 113651.
[41] Ebrahimi, Mohammadreza, Yidong Chai, Sagar Samtani, and Hsinchun Chen. "Cross-Lingual Cybersecurity Analytics in the International Dark Web with Adversarial Deep Representation Learning MIS quarterly 46, no. 2 (2022).
[42] Samtani, Sagar, Yidong Chai, and Hsinchun Chen. "Linking exploits from the dark web to known vulnerabilities for proactive cyber threat intelligence: An attention-based deep structured semantic model." MIS quarterly 46, no. 2 (2022): 911-946.
[43] Jeziorowski, Susan, Muhammad Ismail, and Ambareen Siraj. "Towards image-based dark vendor profiling: An analysis of image metadata and image hashing in dark web marketplaces." In Proceedings of the Sixth International Workshop on Security and Privacy Analytics, (2020): 15-22.
[44] Kavallieros, Dimitrios, Dimitrios Myttas, Emmanouil Kermitsis, Euthimios Lissaris, Georgios Giataganas, and Eleni Darra. "Understanding the dark web." Dark Web Investigation (2021): 3-26.
[45] Al Jawaheri, Husam, Mashael Al Sabah, Yazan Boshmaf, and Aiman Erbad. "Deanonymizing Tor hidden service users through Bitcoin transactions analysis." Computers & Security 89 (2020): 101684.
[46] He, Siyu, Yongzhong He, and Mingzhe Li. "Classification of illegal activities on the dark web." In Proceedings of the 2nd International Conference on Information Science and Systems, (2019): 73-78.
[47] Queiroz, Andrei Lima, Susan Mckeever, and Brian Keegan. "Detecting Hacker Threats: Performance of Word and Sentence Embedding Models in Identifying Hacker Communications." In AICS, pp. 116-127. 2019.
[48] Sabbah, Thabit, and Ali Selamat. "Modified frequency-based term weighting scheme for accurate dark web content classification." In Information Retrieval Technology: 10th Asia Information Retrieval Societies Conference, AIRS 2014, Kuching, Malaysia, Springer International Publishing, (2014): 184-196.
[49] Dalins, Janis, Campbell Wilson, and Mark Carman. "Criminal motivation on the dark web: A categorisation model for law enforcement." Digital Investigation 24 (2018): 62-71.
[50] Avarikioti, Georgia, Roman Brunner, Aggelos Kiayias, Roger Wattenhofer, and Dionysis Zindros. "Structure and content of the visible Darknet." arXiv preprint arXiv:1811.01348 (2018).
[51] Marin, Ericsson, Ahmad Diab, and Paulo Shakarian. "Product offerings in malicious hacker markets." In 2016 IEEE conference on intelligence and security informatics (ISI), IEEE, (2016): 187-189.
[52] Zabihimayvan, Mahdieh, Reza Sadeghi, Derek Doran, and Mehdi Allahyari. "A broad evaluation of the tor english content ecosystem." In Proceedings of the 10th ACM Conference on Web Science, pp. 333-342. 2019.
[53] Fidalgo, Eduardo, Enrique Alegre, Victor González-Castro, and Laura Fernández-Robles. "Illegal activity categorisation in DarkNet based on image classification using CREIC method." In International Joint Conference SOCO'17-CISIS'17-ICEUTE'17 León, Spain, Springer International Publishing, (2018): 600-609.
[54] Dalvi, Ashwini, Soham Bhoir, Irfan Siddavatam, and S. G. Bhirud. "Dark Web Image Classification Using Quantum Convolutional Neural Network." In 2022 International Conference on Trends in Quantum Computing and Emerging Business Technologies (TQCEBT), IEEE, (2022): 1-5.
[55] Gercke, Marco. "Ethical and Societal Issues of Automated Dark Web Investigation: Part 4." Dark Web Investigation (2021): 169-187.
[56] Chertoff, Michael. "A public policy perspective of the Dark Web." Journal of Cyber Policy 2, no. 1 (2017): 26-38.
[57] Best Dark Web Monitoring Tools https://sourceforge.net/software/dark-web-monitoring/ Accessed on 30th March 2023.

Authors' Profiles

Ashwini Dalvi received the M.Tech. Degree in Computer Engineering from the Veermata Jijabai Technological Institute, Mumbai, India, where she is currently pursuing a Ph.D. degree in Computer Engineering. She is an Assistant Professor at K J Somaiya College of Engineering. Her research interests include Dark web Monitoring, Cyber Security, and Intelligent applications.

Sunil Bhirud obtained his B.E., M.E., and Ph.D. in 1987, 1995, and 2001, respectively, from SGGS College of Engineering & Technology, Nanded, India. He worked with Bush India Ltd. and SGGS College of Engineering & Technology, Nanded. He is a Professor in the Department of Computer Engineering & I.T. at VJTI Mumbai. His areas of interest include Signal & Image Processing, Soft Computing, Data Mining, and Machine learning.

How to cite this paper: Ashwini Dalvi, Sunil Bhirud, "Dark Web Monitoring as an Emerging Cybersecurity Strategy for Businesses", International Journal of Information Engineering and Electronic Business(IJIEEB), Vol.16, No.2, pp. 54-67, 2024. DOI:10.5815/ijieeb.2024.02.05