IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 19, 2020, accepted September 7, 2020, date of publication September 15, 2020,
date of current version September 30, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.3024198*

# Evolution of Dark Web Threat Analysis and Detection: A Systematic Approach

**SAIBA NAZAH[1], SHAMSUL HUDA[1], JEMAL ABAWAJY[1], (Senior Member, IEEE), AND MOHAMMAD MEHEDI HASSAN[2], (Senior Member, IEEE)**

[1]School of Information Technology, Deakin University, Melbourne, Victoria 3125, Australia
[2]College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding author: Saiba Nazah (snazah@deakin.edu.au)

The authors are grateful to King Saud University, Riyadh, Saudi Arabia, for funding this work through Researchers Supporting Project number RSP-2020/18.

**ABSTRACT** Dark Web is one of the most challenging and untraceable mediums adopted by the cyber criminals, terrorists, and state-sponsored spies to fulfil their illicit motives. Cyber-crimes happening inside the Dark Web are alike the real world crimes. However, the sheer size, unpredictable ecosystem and anonymity provided by the Dark Web services are the essential confrontations to trace the criminals. To discover the potential solutions towards cyber-crimes evaluating the sailing Dark Web crime threats is a crucial step. In this paper, we will appraise the Dark Web by analysing the crimes with their consequences and enforced methods as well as future manoeuvres to lessen the crime threats. We used Systematic Literature Review (SLR) method with the aspiration to provide the direction and aspect of emerging crime threats in the Dark Web for the researchers and specialist in Cyber security field. For this SLR 65 most relevant articles from leading electronic databases were selected for data extraction and synthesis to answer our predefined research questions. The result of this systematic literature review provides (i) comprehensive knowledge on the growing crimes proceeding with Dark Web (ii) assessing the social, economic and ethical impacts of the cyber-crimes happening inside the Dark Web and (iii) analysing the challenges, established techniques and methods to locate the criminals and their drawbacks. Our study reveals that more in depth researches are required to identify criminals in the Dark Web with new prominent way, the crypto markets and Dark Web discussion forums analysis is crucial for forensic investigations, the anonymity provided by Dark Web services can be used as a weapon to catch the criminals and digital evidences should be analysed and processed in a way that follows the law enforcement to make the seizure of the criminals and shutting down the illicit sites in the Dark Web.

**INDEX TERMS** Crypto market, cyber crime, dark web, onion router.

## I. INTRODUCTION

The World Wide Web (WWW) is a complex system that consists of unprecedented amount of digital information. The normal Internet used daily is accessible through standard search engines such as Google and Yahoo. However, there are large sections of the Internet that is unindexed and hidden from the normal search engines [141]. These concealed part of the Internet is Deep Web which is estimated to make up about 96 precent of the WWW [64]. Within the Deep Web a subset that is mostly used for illicit purpose is the Dark Web or Dark Net [7]. Criminal activities and illegal contents are used with a percentage of 57% in the Dark Web. These commonly include illegitimate drugs, weapons trafficking, child pornography, stolen financial details, unlawful discussions, fake currency, terrorist communication, and more [37], [105], [157].

In 2013, when the US Federal Bureau of Investigation (FBI) shut down the most infamous marketplace Silk Road operating in the Dark Web, these criminal activities caught the attention of the public [153]. Hidden wiki and Deep search engines are the way to browse malicious intents and illicit contents in the Dark Web. These sites provide links access to many other links in Deep Web [64]. One of the main obstacles the forensic analysts face while investigating the criminals activity in the Dark Web is the anonymity presented in the Dark Web services. The contents and services provided

The associate editor coordinating the review of this manuscript and approving it for publication was Fabrizio Marozzo.

by Dark Web are commonly used by anonymous services such as Tor, Freenet, I2P and JonDonym [110].

Most popular service in the Dark Web is the TOR network which provides the facility for the users to secretly share information anonymously via peer-to-peer connections instead of a centralized computer server [81]. This service was intended to access blocked content, use circumvent censorship and maintain privacy of sensitive communications by U.S. Naval Research Laboratory in 2002 [54]. Monitoring the Dark Web is very challenging due to the anonymous design structure of the TOR network. Criminals utilize the Onion Router (TOR) for navigating the Dark Web because of the untraceable and difficult to shut down infrastructure of TOR [39]. This is one of the reasons for the immense pressure on security agencies and law enforcement for monitoring and tracing the activities in the Dark Web.

Criminals usually setup a relay station in the TOR and conceal their criminal activities in the Dark Web. As a result, only the last TOR exit relay is located by the law enforcement when they link the IP address to the identify the committed crime using the TOR browser [110]. Researchers have developed various strategies and methods to monitor and detect different crimes and criminals in the Deep Web. Memex Project developed and implemented by the United States Defence Advanced Research Projects Agency (DARPA) is one of the successful data mining tool in Dark Web [75]. Some of the techniques to pro-actively monitor the hidden parts of the Internet was discussed in a study that include mapping the hidden services directory, social site monitoring, customer data monitoring, semantic analysis and marketplace profiling [39]. Law enforcement has also applied various methods to locate criminals those include social media, IP addresses, monitor activities of users, monitor Bitcoin accounts [98].

To identify the research questions with the justification of future research in any particular study area a systematic review of scientific literatures have many significances [109]. The systematic literature review (SLR) aims at identifying works of a specific area with systematic study via following several research steps and processes. Although some studies have been conducted by authors to analyse the Dark Web, any systematic literature review on the evaluation of the Dark Web in the context of threats is still insufficiently explored, which has motivated us to present this survey. This study aims to explore the emerging crime threats in the Dark Web such as drug transactions, terrorisms, human trafficking, markets for cybercrime tools and so on with their consequences and the corresponding crime monitoring and locating technologies. For this purpose, we have systematically selected and reviewed 65 articles relevant to our research aim. Thus the contributions of this paper are:

- providing a survey on the emerging crimes happening in the Dark Web
- the consequences of the crimes on social, economic and ethical structures
- challenges and difficulties to trace the criminals
- techniques and methods to locate the criminals and crimes with their limitations

The remainder of this paper is organized as follows: Section II provides the implemented research methodology Section III describes our findings with results IV the demographic information about the selected papers. Section V shows the architectures of the selected papers and finally Section VI concludes our work with a discussion and future direction in this research.

## II. RESEARCH METHODOLOGY

This section describes the Systematic Literature Review (SLR) method [87] used for conducting this review. We have also considered some recent studies with SLR method to apply in our work [66], [146]. SLR uses systematic methods to define research question, conduct literature search, screen the findings, extract the data from the selected findings, analyse and synthesize the findings qualitatively or quantitatively [11]. The methodology includes defining (i) the research questions (ii) relevant data sources and the search procedures (iii) inclusion and exclusion criteria (iv) extraction of data (v) analysis and synthesis of the data

### A. RESEARCH QUESTIONS

Providing the summary of emerging crimes happening in the Dark Web with their consequences and defence techniques is the main goal of this work. Thus, the following research questions and the motivations are as follows:

- **RQ1: What are the rising threats in the Dark Web crimes?**

Identifying the type of Dark Web threats globally can help to show how the illegal contents and the services are accessed and what their consequences are. This raises the challenges and importance of creating better technologies and law enforcement to trace the criminals.

- **RQ2: What types of techniques are applied to locate the criminals in Dark Web?**

Identify the law enforcement methods, applied and available technologies for tracing and detecting the crimes and criminals in the Dark Web. It gives the future path to apply different strategies with latest technologies along with law enforcement to defeat the cyber-criminals plan.

### B. SEARCH STRATEGY AND SELECTION

We have followed the search strategy guidelines stated in [87] and [163] which is explained below in details. To collect the data for the review papers, electronically-based search was performed from IEEE Xplore, ScieneDirect, Springer, Scopus, ACM Digital Library and Google Scholar.

We included in our search term the Dark Web crimes mentioned in [33] and [153]. We have used the terms from our research questions in Section II.A. Boolean search operation with ANDS and ORs has been implemented to specific phrases. The search terms used for retrieving our relevant articles are presented in Table 1. However, this is to mention that

**TABLE 1.** Search term for selection of literatures.

| Sl | Search Term |
|---|---|
| 1 | "Dark Web" AND " crimes" OR "Dark Net" AND "cyber security" |
| 2 | "Dark Web" AND "threats" OR "attack" AND "crime rates" |
| 3 | "crypto markets" OR "Dark Net marketplaces" OR "bit coin" OR "silk road" OR "TOR" |
| 4 | "illicit" OR " illicit Products" AND "Dark Net" |
| 5 | "techniques" AND " Dark Web" OR "strategies" AND "Dark Web" |
| 6 | "law enforcement" OR "darpa" OR "memex" AND "challenges" AND "Dark Web" |
| 7 | "drugs" OR "human trafficking" OR "fraud" OR "prostitution" OR "terrorism" Or "data breach" AND "Dark Web" |

**TABLE 2.** Inclusion criteria for selection of literatures.

| IC# | Description |
|---|---|
| IC1 | A study that is focused on Dark Web related crimes and demonstrates the crimes, consequences and assess techniques |
| IC2 | A study that is based on the tracing techniques and technologies for locating the criminals in the Dark Web for cyber security |
| IC3 | The search term keywords in Tab.I applies the operators of search syntax OR, AND.<br>AND operator signifies that both keywords must be present in the search queries and OR means that at least one keyword must be present in the queries searched |
| IC4 | Studies published in English language |
| IC5 | Include data from journals, conferences and web articles published between the year 2003 to 2019 |
| IC6 | Include abstract based and full-text studies |

**TABLE 3.** Exclusion criteria for selection of literatures.

| EC# | Description |
|---|---|
| EC1 | Exclude the duplicate articles obtained by authors and/or different libraries |
| EC2 | Exclude articles those specifies Dark Web but do not include crime threats or crime locating techniques |
| EC3 | Exclude cyber security defense articles not related to Dark Web |

different search terms have been used for retrieving relevant publications. Also, a search for additional articles from the references of the relevant articles found was considered.

The matching of the strings in the search terms from the digital libraries are based on the title, abstract and keywords of the papers except the Springer library which does not allow to restrict the search in specific parts of the paper. We did the filtering and screening for selecting the most relevant papers based on the inclusion and exclusion criteria discussed below.

The inclusion and exclusion criteria followed in this survey are described in Table 2 and 3 respectively. After applying these screening steps with the inclusion and exclusion criteria 65 articles were selected for this review paper. The selected publications are listed in Appendix (A), Table 11.

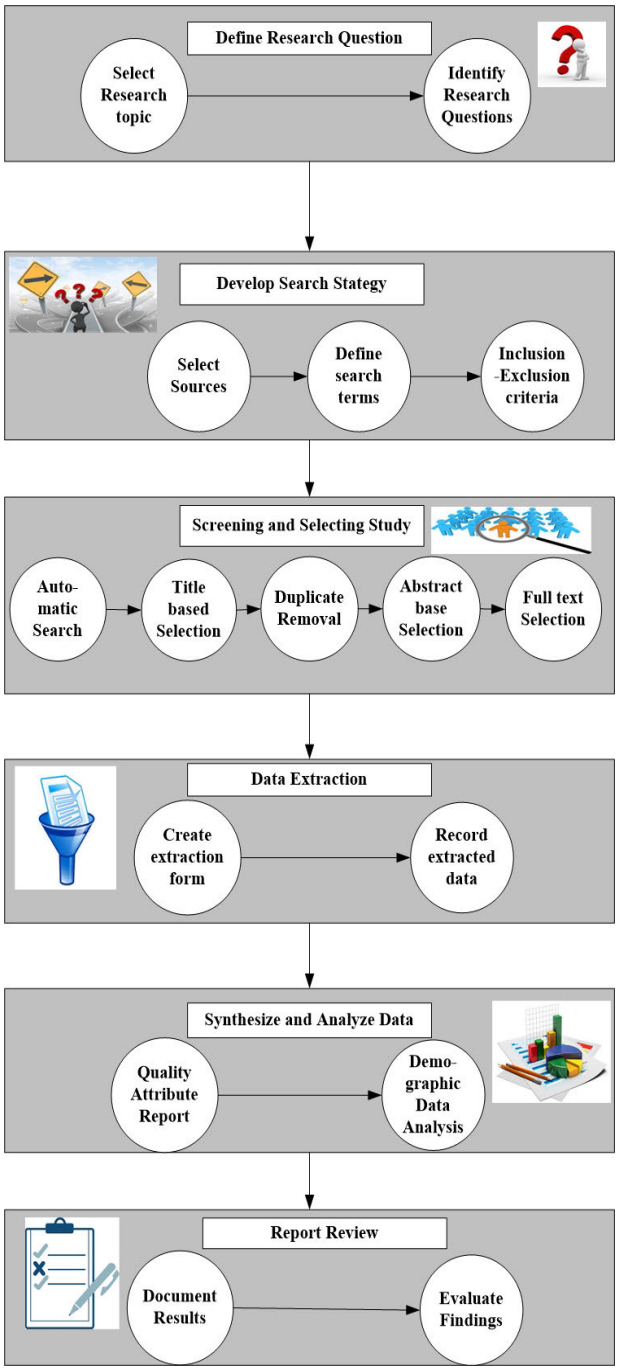


**FIGURE 1.** Applied SLR Methodology.

After applying the inclusive and exclusive criteria in our filtering process we ended up with 65 papers. The overall SLR article selection procedure applied in our survey paper is described in the Figure 1. Some of the selection techniques are described as follows.

•*Automatic search:* Using the search terms on the six database libraries mentioned with automatic search we could collect 1920 papers.

•*Title-based selection:* For a fast article picking method title based selection was performed. We choose the articles

from the title of the paper that are relevant to our SLR. This made the number of papers to 581.

•*Duplication removal:* In this case duplicate papers were removed as some of the database indexes papers are available in the other databases. After removing the duplicates the number of articles reduced to 393.

•*Abstract-based selection:* To check whether the selected 372 papers are related to our SLR, the abstract of the papers were read. The irrelevant abstract articles were disregarded at this stage and 100 papers were selected.

•*Full-text selection:* Each of the 100 papers was completely read through and 65 papers were selected based on this.

Out of 65 papers, 18 were selected from IEEE Xplore, 22 papers were from Google Scholar, 6 articles were chosen from ScieneDirect, 9 papers were selected from Springer, 5 papers were selected from Scopus and the rest 5 papers were extracted from ACM Digital Library.

### C. DATA EXTRACTION AND SYNTHESIS

To answer our research questions in this systematic literature review the procedure of data extraction and analysis of the data extracted from the filtered papers is discussed in this section. The extraction of data from the filtered articles has been done based on the data extraction form that is shown in Table 4. Microsoft Excel spread sheet was used to record the extracted data.

**TABLE 4.** Data extraction form.

| SL | Data Item | Description |
|---|---|---|
| 1 | Year | Publication year of the article |
| 2 | Authors | Authors of the article |
| 3 | Source | Source of the publication article(e.g., google scholar) |
| 4 | Title | Title of the article |
| 5 | Type | Type of the publication article(e.g., journal) |
| 6 | Criteria | Selected criteria of the articles (e.g., cryptomarkets) |
| 7 | RQ1 | Emerging threats in the Dark Web crime |
| 8 | RQ2 | Techniques applied to locate the criminals in Dark Web |

To evaluate article suitability in accordance to answer the research question the quality attribute rules were applied. 6 QARs were identified and each one is worth 1 mark out of 10.The score is as follows "fully answered"=1, "above average"=0.75," average"=.50 and "below average"=0.25,"not answered"=0. The overall score of the article will be the summation of marks obtained from the 6 QARs. If the result was 3 or higher, the article was considered to answer our RQs otherwise was excluded. The QARs are shown in Table 5.

For data synthesis on the extracted data we have used different procedures that aggregate the evidence to answer our RQs. The demographic data of the reviewed articles have been analysed using descriptive statistics.

**TABLE 5.** THE QARS of this SLR.

| QAR# | Description |
|---|---|
| QAR1 | Are the objectives of the research articles clearly defined? |
| QAR2 | Are the Dark Web crime backgrounds addresses properly? |
| QAR3 | Are the Dark Web crimes tracing techniques used clearly defined? |
| QAR4 | Are the articles comprehensive and take into consideration past and current literature? |
| QAR5 | Are the methods used to analyze the results appropriately? |
| QAR6 | Do the articles identify the gap of knowledge? |

The demographic information result analysis is discussed and shown in section IV. The architectures of the literatures obtained from the data analysis process are discussed and shown in section V.

The data analysis process for our RQs is summarized in Figure 2. We have first identified the quality attributes from the selected articles based on the selection criteria for each RQ. Depending on the RQ we extracted the motivations and applied thematic analysis [43] to extract the themes from the articles.

*RQ1:* Depending on the quality attributes and selection criteria we first extracted the motivations from the papers and identified different architectures used in the literatures. After that we applied thematic analysis to extract the themes from the studies. Then we did the qualitative data analysis based on the themes extracted. This gives the architecture frameworks used in the literatures based on the themes. From Appendix A Table 11 ID 10, ID 26, ID 40 are themed as '*Terrorist and ISIS*' study from the obtained motivation of these studies, as the studies says the Terrorism is one of the threatening crimes incurred with the Dark Web.ID 7, ID 53, ID 54, ID 55, ID 56 are themed as '*Drug transactions'* study as the motivations obtained from these studies aim to the study of terrorist organizations on the Deep Web. The summary of the different architecture frameworks obtained from the RQ1 block is the input of the RQ2 block. Thus RQ1 provides important information that goes to RQ2.

*RQ2:* Our RQ2 data analysis is done based on the output of RQ1 that helps to organize the architecture of different literatures based on the analysis as shown in Figure 2. At the first step we extracted the motivations from the architectures summary to identify the characteristics of the models used in different frameworks of the architectures. These characteristics are the data size used, practical usage and the performance of the models. Then we applied thematic analysis to get the generalized models. This step defines the introduction and motivation of our architecture analysis. Once the generalized models are identified we analysed the proposed models to identify the different tools and techniques used in the models and applied thematic to obtain the generalized methods from the analysis. This step serves as the description and example exploration in the architecture analysis.

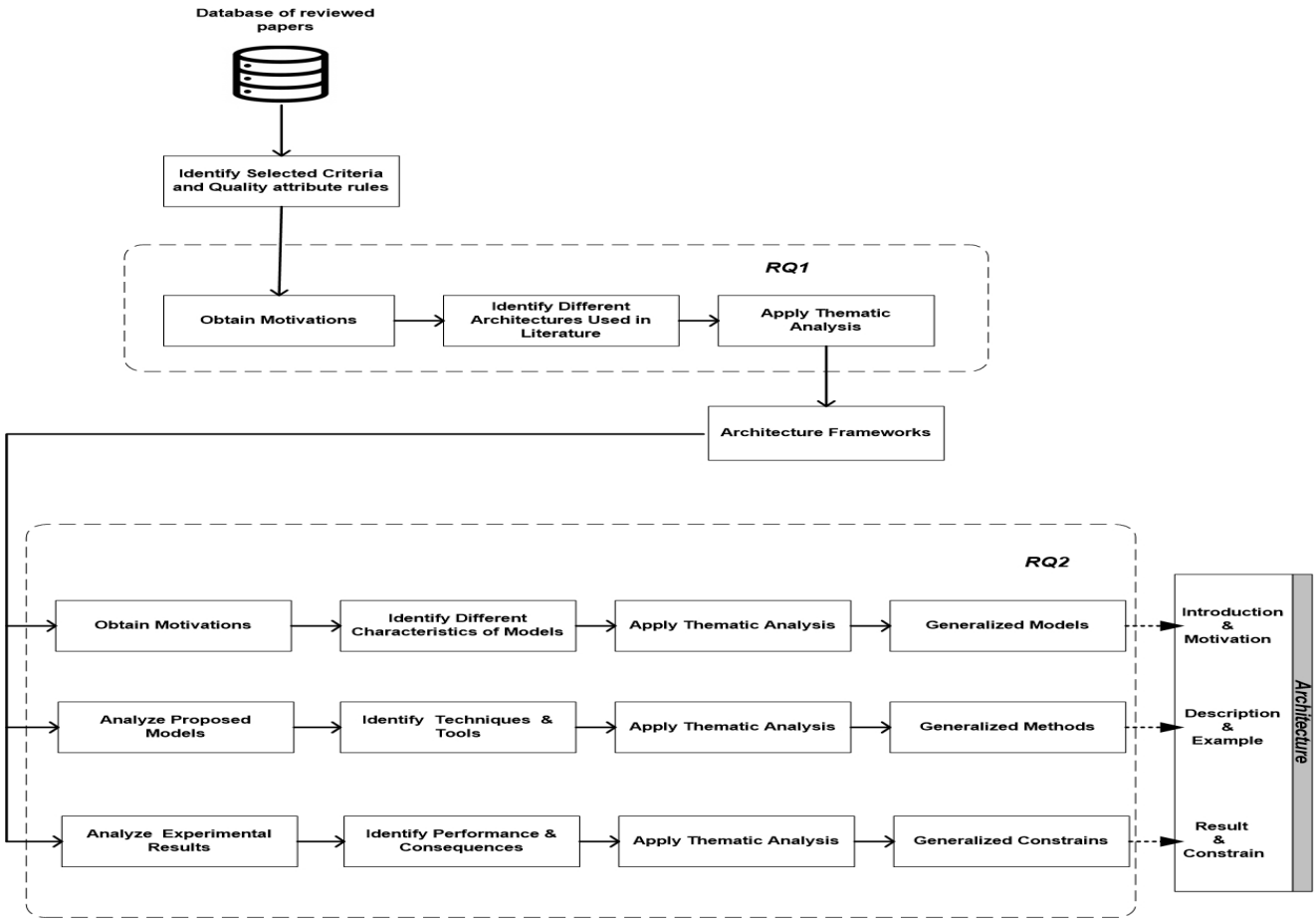


**FIGURE 2.** An Overview of the Data Analysis Process.

To analyse the performance of methods applied in the models in the frameworks, we studied the experimental results and the consequences of the methods then applied thematic to get the final results and constrains from applied methods. Thus we got the result and constrains for our architecture analysis. From the RQ2 we could summarize the obtained analysis results as below.

*Method:* We found 9 generalized implemented crime and criminal detection methods in Dark Web those include Hash value analysis, Sock puppet and key informants, TOR attacks, exit node analysis, Monitoring network traffic, classification of network traffic, marketplace scarping, monitoring Dark Web. For example ID 2,ID 13,ID 18,ID19,ID 24,ID 27,ID 28, ID 29,ID 30,ID 32,ID 41,ID 49,ID50

*Tools:* We found 4 generalized law enforcement tool implemented in Dark Web crime forensics those include MLAT, DARPA, Bit coin flow, social media For example ID 9, ID 21, ID 35, ID 36, ID 65.

## III. RESULT AND DISCUSSION

In this section, we will answer the two research questions and discuss the result obtained from this systematic review of the literature. The first research question (RQ1) is to determine the rising threats in the Dark Web crimes. We will address this at the subsection A. This subsection answers our RQ1 with subsection III.A.1 to III.A.8 giving the overview of the threats in the Dark Web occurred from the criminal activities. Our second research question (RQ2) is to analyse the different techniques applied to locate the criminals in Dark Web. To answer and analyse this we have discussed in the subsection B. The answer of our RQ2 is addressed in subsection III.B.1.a to III.B.2.i with the analyses of the methods and techniques applied against these threats. After this we present the third subsection III.C which evaluates the contents of the extracted 65 studies with their contributions in a summative way.

### A. CRIMINAL ACTIVITY THREATS IN THE DARK WEB(RQ1)

From the selected articles we have determined eight major crime threats in the Dark Web which answers our RQ1. The list of the crime threats follows:

- Human trafficking and sex trafficking
- Pornography industry
- Assassinations and its marketing
- Drug transactions
- Child Pornography
- Terrorist

- Markets for Cybercrime Tools and Stolen Data
- Dark Net currency exchange using bitcoin

#### 1) HUMAN TRAFFICKING AND SEX TRACFFICKING

Human trafficking and sex trafficking are a large part of crimes and has extremely been increased due to online forums, chat services and anonymousness of the Deep Web [49]. Human trafficking is considered a significant human rights challenge. About 2.5 million individuals worldwide are trapped in some form of modern-day slavery according to the United Nations Office on Drugs and Crime [73]. The victims are forced into slavery as beggars, sex workers, and child soldiers, factory workers, domestic workers, and laborers in different commercial industries. To recruit victims for human trafficking and sex trafficking human traffickers do negotiations and contracts.

Trafficking networks can avoid detection, censorship and surveillance system employed by the government and anti-human trafficking organizations as the networks used by traffickers are dynamic [90]. Human trafficking in the field of the hidden nature is difficult and challenging to identify. International Labor Organization (ILO) reported in 2017, Nearly about 40.3 million people were in modern slavery and this number includes 24.9 million in forced labor and 15.4 million in forced marriage in the year 2016. This concludes for every 1,000 people in the world there are 5.4 victims of modern slavery. It is alarming that in every four victims one is a victim of modern slavery. Around 16 million victims of forced labor out of 24.9 million were oppressed to domestic work, construction or agriculture, 4.8 million victims of forced sexual enslaved and the rest 4 million to forced labor. Women and girls are exploited to commercial sex industry and other sectors by 99% and 58% respectively [67].

To see the consequence of human trafficking, the report on human trafficking of can be brought into light where it was reported that in 2014 a federally-funded hotline for trafficking victims received more than 21,000 calls. During that same period, the Department of Justice secured 184 convictions for trafficking, up from 174 in 2013. Among the cases, 157 were victims of sex trafficking and 27 were focused on labor trafficking [125]. Victims of human trafficking are mostly innocent young girls from foreign countries who are manipulated, lied to, and often kidnapped and forced into prostitution [40].

#### 2) PORNOGRAPHY INDUSTRY

Victims mostly exploited by pornography industry are women of human trafficking and sex trafficking [124]. Traffickers force victims by fear of assassination for pornography production once male-female sign agreement for accomplishing the acts. The sex trafficker's records video without the consent of victims and distributes those to interested parties through pornography industry. Traffickers also publishes recordings and photos in their websites [41].

Many web sites related to pornography are hosted in the Dark Web. Similar to human and sex traffickers, pornography industry use Dark Web, social media, online forums to recruit or kidnap victims hiding their identification [40]. New forms of criminal acts as well as extreme levels of violence are present in the online prostitution videos and images [71]. With the extensive use of social media prostitution has also become more sophisticated. From an investigative perspective of interviewing victims, witnesses and tracking behavior thus monitoring the prostitution activity are extremely difficult to conduct effectively [20].

#### 3) ASSASSINS AND MARKETING

Dark Web is used by criminals to sell their assassin skills. MailOnline, White Wolves and C'thuthlu websites provided advertisement for criminals which mentioned the hiring amount of $10,000 in US and $12,000 in European [37]. For a police officer to high rankling politician this price ranges from 40 thousands to 15 million of hire price [45]. Hidden Wiki and Deep search engines are the most common ways to explore the Deep Web as these contains dozens of links to illicit onion sites [64].

Thirteen different illicit onion sites were listed by the researchers under the category of Commercial Services. Some most wanted findings were Black Market Reloaded forum, Sheep Market forum and the Silk Road forum [13]. From year 2011 to 2013 the Silk Road was the most prohibited Deep Web market for buying drugs, discussing same minded political beliefs with individual freedom, free markets, and limited government. Although the original Silk Road website was active for less than three years, during this small period it brought a huge impact. Silk Road users considered themselves entirely anonymous for the communications on because it was operated as a Tor hidden service. Furthermore, transactions on Silk Road could only be made using bitcoin transactions [95]. Bitcoin is the most common currency employed in all Tor hidden-services trade. Although Bitcoin transactions can be monitored but not easily de-anonymised [112].

However, the popularity of Silk Road came to an abrupt end on October 1, 2013, when the FBI, IRS, and DEA seized the Silk Road market and arrested Ross Ulbricht the founder of Silk Road. He was accused in separate complaints of paying for the attempted murders of two business associates who he believed had crossed him. With an estimated value of $3.6 million, the Justice Department shut down the Silk Road website and seized Bitcoins [120].

#### 4) DRUG TRANSACTIONS

Two types of drug markets within the Deep Web are usually found. These include the markets that are dedicated to one specific type of drug such as. heroin. Due to the product expertise and vendor customer relationship these type is much popular. The second type of drug markets is general shop for buyers where all types of illicit products are offered such as weapons, pornography, stolen jewelry, black-market cigarettes and credit cards. The vast assortment of narcotics,

to include drug hardware and drug manufacturing chemicals are the most common items [38].

Anonymity of Deep Web has helped significant rise in drug transaction from the Deep Web which has created digital black market for drug. As no face to face communication is required for drug dealing, illicit vendors use Dark Web for buying and selling drugs. Silk Road was one of the examples in Dark Web marketplace that sold the drug over a billion dollar and posted drugs by DHL or drop shipping [8], [9], [17], [103], [148].

After Silk Road was shut down in October 2013, several other crypto markets rose on the Dark Net [30], [152]. Evolution, one of the most popular crypto markets active from January 2014 to March 2015 was studied in a research from Swiss point of view and found more than 48,000 listings and around 2700 vendors claiming to send illicit drug products from 70 countries. They identified three sellers located in Switzerland. The purchases were carried out to confront digital information such as shipping country, type of drug and so on. The illicit drugs purity was found different from the information in respective listings although digital information such as concealment methods and shipping country were accurate [126].

Another Dark Net marketplace for illegal drugs like marijuana, cocaine, and many other legal and illegal goods was Mr. Nice Guy. The security of the site is fairly good as well as the registration has a better level of protection than a standard site [50].

#### 5) CHILD ABUSE

Children are using social media and many applications such as Omegle, Ask.fm that hide identity of users for communicating [42]. Pedophiles use the benefit of these applications to interact with the children. Dark Web is massively used by Pedophiles and related criminals for child pornography, sharing photos and pons. Freedom hosting used 550 servers all over the Europe that offered space to anyone for hosting porn form children [27], [64].

In operation pacifier the FBI arrested several hundreds of pedophiles from US and international territories which involves 2000,000 users, 23,000 explicit images and 9000 video files with explicit sexuality [62]. The suicide case of Amanda Todd, a 15-year-old Canadian girl, brought the global attention to online children abuse. She posted a YouTube video few weeks before her suicide recounting her ordeal in 2012 where she used a series of flash cards to tell her experience of being blackmailed into exposing her breasts via webcam, being bullied and physically assaulted [2].

Webcam child prostitution has become a growing threat of online child sexual abuse where the victim simply sells his/her live sexual images through Voice-over-IP (VoIP) applications. Using the video streaming feature of VoIP applications live child abuse images are produced and sold for profit [2], [123].

In 2011, freedom hosting firm was identified to host 95 percent of child porn on the TOR network. The site hosted around 100 child porn sites with thousands of users and was shut down in 2013 [121]. Freedom Hosting II was identified in 2017 by hackers that followed the previous one. Hackers claimed that over 50% of the content on Freedom Hosting was related to child pornography and the data are dumped but they can identify the users of these sites [144].

The largest child sexual exploitation site Welcome to Video operated by 23-year-old Jong Woo Son of South Korea was started on 2015. This operator has been charged in 2018 and the site has been tracked down. Officials around the world were arrested with a total of 337 Welcome to Video users in 23 US states, Washington DC, and in 11 other countries. This resulted the rescue of at least 23 abused children victims from United States, Spain and the United Kingdom by the site participants [115].

#### 6) TERRORISM

Terrorism and terrorist organizations on the Deep Web are dangerous threat to national security. Terrorist organizations such as al-Qaeda/ISIS and ISIL ISIS have utilized the benefits of Dark Web to fulfill their negative motives and spread propaganda [18], [104], [157], [160].

The Islamic State in Iraq and Syria (ISIS) uses the Dark Web to solicit money to help, support their cause and as a means of passing information throughout the change of command. A technology editor for Defense One, stated that there is substantiated data that supports ISIS or ISIS supported groups are pursuing the Dark Web for a different reason other than passing false information for the purposes of indoctrination and self-aggrandizement. ISIS uses Bitcoins for services provided on the Dark Web. The U.S. Military monitors the Dark Web trying to track ISIS but the problem is that law enforcement and the military have not found a way to track ISIS on the Dark Web without infringing on privacy rights [149].

ISIS uses Dark Web as a weapon for terrorism where they provide live streaming and recording of death sentences of prisoners. They use Dark Web as broadcasting media where they upload small video clips of their inhuman activities. For recruiting the soldiers around the world they also use the Dark Web [137]. To protect the identities and get the safeguard from hackers, ISIS turned to the Dark Net. They spread news and propaganda from the attacks in Paris in November 2015 using Dark Net sites and online platforms.

ISIS's media outlet, Al-Hayat Media Center, posted a link and explanations on how to get to their new Dark Net site on a forum associated with ISIS. The message was sent by Telegram, an encrypted texting program used by ISIS for smartphones and Windows through a TOR browser line. The developers of Telegram were too confident to offer a cash reward who could break the security and solve the encryption [157].

ISIS coordinates by planning operations and discussing the command and control using a public forum, which is encoded. Applications like Skype and WhatsApp are used by them to send messages across the battlefield and have

been able to use small drones to collect real time data to use as propaganda. ISIS recruits by using online magazines with slick production values to communicate their views as well as how to construct weapons for terrorist attacks [137]. Five categories of terroristic activities on the Internet are identified in various studies. Those include propaganda, recruitment and training, fundraising, communications and targeting [98], [168].

#### 7) MARKETS FOR CYBERCRIME TOOLS AND STOLEN DATA

The anonymous marketplaces are one of the main sources for the cyber criminals to sell cybercrime tools and leaked or stolen data in the Dark Web. Black Market Reloaded forum, Sheep Market forum and the Silk Road were the most known and rising Dark Net market places those have been disappeared due to arrest [72] and voluntarily [13].

Following the path of the infamous Silk Road numerous crypto markets evolved. A large portion of these marketplace ecosystems has been analysed for the year 2013 to 2015 in [139]. These marketplaces include Agora, Atlantis, Black Flag, Black Market Reloaded, Tor Bazaar, Cloud 9, Deep Bay, Evolution, Flo Market, Hydra, The Marketplace, Pandora, Sheep Marketplace, Silk Road, Silk Road 2.0, and Utopia. Some of these marketplaces were seized by police, had voluntary shutdowns, and suspected fraudulent closures.

Alongside the selling of drugs pornography and weapon these online anonymous marketplaces are the main source of selling stolen or leaked information those include financial statements, credit card details, cloned pins and sensitive information [88]. Dark net marketplaces selling products and services on malicious hacking were investigated with various developed systems [12], [117], [128], [134]. Information of malicious tools and services that facilitate the cyber carding crime are exchanged by the data thieves and cyber criminals via underground black markets and forums were studied in [97].

Some of the studies have been undertaken on finding the vendors origin and the challenges on locating the information of the online anonymous marketplaces [76], [77], [78]. Until this paper writing date according to DNStats the rising and active marketplaces are Empire Market, TOCHKA Market, SLILPP, UAS Service RDP, Hydra, Wannabuy RDP, UNICC Shop, Cannazon, Monopoly Market, Tochka Market, Empire Market [52].

#### 8) DARK NET CURRENCY EXCHANGE USING BITCOIN

Bit coin is a crypto currency that allows the Dark Net marketplace members to stay anonymous while transaction [74]. The Dark Net marketplaces often only accepts bit coin as a currency.Silk Road was one of the successful black marketplace that could earn over US$1.2 billion through this [95]. Bitcoin's legitimate use has been controversial but supported as a currency which rises and conceals the money laundering [25].

Money laundering is a three step process of making illegally gained dirty money. These are defined by the U.S. Treasury's Financial Crimes Enforcement Network. The steps include placing the illegal money in the legitimate financial system, layering it within additional transactions and integrating it into the financial system with more transactions so the funds appear licit [60]. Elliptic's forensic analysis tool was used to identify bitcoins moving from illicit entities for the transaction data between 2013 and 2016 [60]. Because of the weakness with publicly available information stored in blockchain of Bitcoin this is vulnerable to analysis but criminals follows the Escrow system that ensures preventing scamming [3].

### B. TECHNIQUES TO LOCATE CRIMINALS IN DARK WEB(RQ2)

Cybercrimes in the Dark Web and crimes in the real world are quite similar except that it is hard to track virtual crime using the Dark Web by the law enforcement. The anonymity provided by Dark Web services is one of the main problems some forensic analysts may face while trying to investigate criminal activity. Thus, this is hindering the forensic investigation of criminal activity. Many crime detection studies have been done on the Dark Web to locate the crimes or criminals. We will see the detection techniques and law enforcement methods applied and initiated for this purpose in the sub sections under law enforcement and detection method. This section answers our RQ2.

#### 1) LAW ENFORCEMENTS

Smaller law enforcement agencies do not have the technical expertise to fight back specific crimes as cybercriminals have increased their abilities. There are several types of laws pertaining to criminal activity on the Dark Web including criminal law, civil law, and regulatory law. Criminal law relates to crimes at the government level of local, state, and federal. The type of penalty could range - from a fine to life in prison. Depending on the state the crime occurred in, punishment could be death. Civil law relates to a person or organization that has been held responsible and instructed to pay a fine or required to complete a service as part of the punishment. In regulatory law, the agency within a jurisdiction has the right to issue fines as punishment for activities. Regulatory agencies have the right to cease all business operation of individuals or company that are not in compliance [136].

#### 2) SOCIAL MEDIA

To identify criminal activities in the Dark Web social media and Deep Web can be jointly used. Different media such as Twitter, Youtube and Facebook are used to identify suspects [142]. According to RSA, to communicate and sell stolen identities, credit card numbers and other information, cybercriminals rely heavily on social media platforms such as Facebook, Snapchat Instagram, WhatsApp, Telegram and other social media platforms [24].

The reason behind the usage of these platforms by cyber criminals is its quite easy to share and pass on anything including malware in the social sites. On average these

platforms have additional techniques to scam such as adverts, sharing buttons and plug-ins compared to other websites. Moreover, hundreds to thousands people connected on these platforms creates convenience for the criminals to spread malware in a wider range [21]. A six-month global study by criminology expert at the University of Surrey in the UK on social media cybercrime in last year, found annual earnings of cybercriminals exploiting popular social platforms is nearly $3.25 billion [99].

On the positive side, social networking is providing pathways to solve crimes for the law enforcement officers. As there is large number of users in the social platforms, tips can be gathered by their online presences as well as observation of the crimes committed within their communities [129], [148]. Many of the successful track downs by police with the help of social media was the arrest of Raderius Glenn Collins a Florida burglar who posted a seven minute Facebook video bragging about a $500,000 jewelry heist that got 3,000 views; Derek Medina a 33-year old man was sentenced to life in prison for second-degree murder of his wife. The criminal posted the picture of his wife after killing her on Facebook and wrote about murdering her; Maxwell Marion Morton was charged with first-degree murder after he posted via SnapChat a selfie of a classmate who has been shot in the face; 71 people were arrested by Police in Cincinnati after a 9-month investigation used with the help of social media to identify key gang members [55].

Information found on social media with police records and reports were combined by University of Cincinnati's Institute of Crime Science to establish a link between suspects that helped the Police to arrest the gang [55]. As reported by LexisNexis, social media is used as an investigation tool by more than 80% of police departments [96]. The usage of social media in investigation in the year 2012 and 2014 is described by LexisNexis.

#### 3) DARPA AND MEMEX

The Defense Advanced Projects Research Agency (DARPA) identified several tools on the market available to law enforcement to help with locating individuals that are using the Dark Web for criminal activities. In the past, the FBI used the Metasploit Decloaking Engine to help with their investigation of the Dark Web. Metasploit Decloaking Engine and Memex systems are used by US law enforcement agencies through indexing the Deep websites in an intelligent way to identify the criminals particularly the human traffickers in the Deep Web [47].

The process of finding criminal operating on the Dark Web has been made easy for the law enforcement by DARPA (Defense Advanced Research Projects Agency) by developing a suite of tools, collectively known as Memex [56], [107]. These tools are primarily written in Python and were developed in collaboration with various universities [47]. In 2012, the Drug Enforcement Agency (DEA) completed a two-year investigation on Dark Web based marketplace called the Farmer's Market, which distributed drugs to over 3,000 customers in dozens of countries and all 50 states in the U.S. According the DEA's press release, the TOR network was used to facilitate the operation. TOR masked the IP addresses by allowing the routing of transactions and email to be diffused throughout various relays all over the world. Farmer's Market used many forms of payment for the illicit goods including electronic methods. During the roundup, federal agents and local law enforcement also seized hashish, LSD and MDMA, as well as an indoor psychotropic mushroom grow and three indoor marijuana grows [80].

#### 4) BIT COIN FLOW

Bit coins are the virtual money used for transaction in the Deep Web [69]. The bit coin flow in the Dark Web is used by the law and policing agencies to locate criminals. The criminal's activity can be monitored by the law enforcement agencies by analyzing the bit coin flow. One of the successful examples of identifying criminals from Deep Web is the Silk Road server [95]. FBI identified the location of server through bit coin transaction flow in Iceland data center. Silk Road was operated using TOR which is an anonymous network but the site was identified due to a misconfiguration in Silk Road's login page which described the IP addresses and physical location of the server [73].

FBI also has been successfully able to hack the underground discussion board darkode where criminals join every day and share information, discusses for criminal activities [61]. The massive Dark Web child-pornography site named Welcome to Video that was launched on June 2015 has been shut down in 2018 March [115].

The arrest and surrender of the site's operator, Jong Woo Son has been announced in a joint press conference by the U.S. Attorney's Office for the District of Columbia, the Justice Department's Criminal Division, the IRS Criminal Investigation (IRS-CI), and U.S. Immigration and Customs Enforcement's Homeland Security Investigations (HSI), together with the National Crime Agency of the United Kingdom and Republic of Korea. This takedown has been possible by tracing bitcoin transactions other than any hacking or encrypted communication [57]. The site used to charge fees in bitcoin from every user and give unique bitcoin wallet address to the user account. IRS-CI traced the bitcoin transactions on the site by the people from all over the world who downloaded and uploaded materials to the site. They were also succeeded in identifying the location of the site administrator. The agents of IRS-CI analyzed the blockchain and deanonymized bitcoin transaction which resulted the identification of hundreds of the predators from all around the world [57].

#### 5) MLAT

The Mutual Legal Assistance Treaty (MLAT) helps aid law enforcement in support of their investigations. To assist law enforcement, the U.S. conducts a joint investigation that permits more than one country to be involved in the criminal investigation. The analysts can initiate an official event

notifying the Office of International Affairs (OIA) and generating MLAT protocols [68].

MLAT is one of the established methods for law enforcement sharing information across borders. If a state seeks access to digital evidence that is outside the border of that state, a formal request is required to be fired The aim of MLAT is protecting the legal rights of people those are questionable of offshore offending [156]. However, The MLAT process is slow as it often takes months, the structure is opaque and due to too many requests under stress and thus inconvenient. In 2013, a search warrant was issued by the US government to Microsoft for seizing the information of a specific email account. But Microsoft could oppose it as the emails were on the server storage located in Ireland [143].

#### 6) CRIME DETECTION

To fight back the emerging crimes growing in the anonymous Dark Web the importance of detecting the perpetrators and the associated crimes is inimitable. Although this is a very challenging task to analyze the untraceable anonymous networks used by the cyber criminals, many techniques and methods have been developed that can be implemented in this area.

##### *a: HASH VALUE ANALYSI*

Identification and collection of digital evidence is one of the main criteria to put criminals under law enforcement in the cyber-crime as these types of crimes are performed in computer based systems [154]. Hash functions play powerful role in cryptography to prove any evidence is authentic during the investigation. Hash functions produce has values that represent the original message from which they have been computed [46]. Some popular hash algorithms are MD5,SHA-1,SHA-256,SHA-512 [84], [155].

TOR has a complicated structure with thousands of internal nodes and hash value computations that is untraceable but the exit node can be analyzed. [51], [110]. Hash value analysis at the exit node layer of the onion routing can be implemented to find the destination of the connecting server [10], [119]. Many researches have been done applying the hash value analysis for the detection of crime and digital forensics. These studies include analysis of crimes [82], approximate matching for digital forensic [22], [23], [31], [91], [130], analysis of malware [85], steganography software detection [92], fraud detection and financial crime detection [86].

##### *b: SOCK PUPPETS AND INFORMANT ANALYSIS*

Sock puppet is a false online identity or the same person having multiple usernames to communicate in the online [122]. This method is mostly utilized by the cyber criminals to do ID theft, sell fake products, terrorist activities on the Dark Web which make it difficult for their seizure. Thus, sock puppet detection is very crucial in cyber intelligence operation in order to forensic accounting, extrapolating information about the criminals, monitoring the communication, and screening the Dark Web for tracking the terrorists.

Methods to detect sock puppets using authorship identification has been adopted in [26], [113], [140]. In the discussion forums and social site sock puppet detection method was proposed in [63], [83], [93], [101], [148], [167]. Other studies on the Wikipedia sock puppet detection [138] and detecting authorship for online message using writing style [166] were done.

Another detection method of crimes can be Key informants which is a source of experts' information. Researchers use key informant interviews over an extensive period of time with participants who have unique knowledge of a topic. Key informant analysis has played vital role in the research studies on illegal drug reporting, money laundry for criminal finance and terrorist organization structure [44], [104].

##### *c: NETWORK ANALYSIS METHODOLOGIES*

Proper understanding and analysis of the structure of the cyber-criminal and terrorists networks can lead to the path of technical insight for obstructing the illicit activities in the Dark Web. The network analysis of the terrorist and extremist groups in the Dark Web have been conducted in studies to detect and discover crime threats [28], [34], [160], [165], [170]. Network analysis with topic based model in the Dark web portals has been implemented in studies to inherent topics [94], [127], [161].

Classification of network traffic using correlation have been implemented in the field of network security monitoring and identifying attacks, this can be used as a traffic classification technique for detecting attacks on the Dark Web [70], [159], [164], [169].

Another network analysis method is locating entry and exit nodes to identify attacks [145]. Although TOR has a powerful infrastructure which is hard to break, the communication of the users and the routing behaviour can be detected by implementing some attack techniques in the TOR network [6], [19], [58], [100].

##### *d: MARKETPLACE SCRAPING*

Dark Web marketplaces are one of the main platforms for cyber criminals to spread and conduct their illicit crimes. Thus scrapping the information from these marketplaces can lead success to detect and catch the criminals [77]. According to an article by Alan Travis, with the use of anonymous marketplaces and forums to sell financial data about 5.1 million incidents of online fraud in England and wales were found [147].

Automated Dark Web marketplace scraping methodology was developed and implemented in different studies. The analysis of the scraped data could provide the basis for a subsequent investigation of suspected criminals and crimes [30], [38], [76], [117], [158], [170]. The trades in the anonymous marketplace or crypto markets have been analyzed by researchers to find the consequences and solutions towards Dark Web marketplace associated crimes [15], [16], [36], [49], [53], [139], [150].

#### e: MONITORING DARK WEB

Monitoring the Dark Web is very challenging because of the untraceable and anonymous infrastructure. Numerous methods and techniques have been established to monitor the Deep Web by the researchers. Some of the monitoring techniques of hidden parts of the Internet were discussed [39]. The monitoring of the Dark Web can be beneficial to the crime analysts as by retrieving the data for a websites on the Dark Web a prominent database can be built that can contain important information about a hidden site which can help track future illegal activities and criminals.

Monitoring extremist and hate groups to analyse the usage and content of the Dark Web data have been adopted by many researchers [1], [133], [165], [168]. Various studies were proposed to monitor malicious activities and threats in the Dark Net [14], [114], [116], [135].

#### f: HONEYPOT DEPLOYMNET

The network server is often targeted cyber-criminals for attacks as a mean of spreading malicious software across the network or breaking into the system. Monitoring the network traffic is another fruitful way to observe the activities of the criminals in the Dark Web TOR network [59]. In this context, the use of Honeypots in detecting cyber-attacks and criminals behaviour in the network traffic have been proposed in various studies [29], [89], [118].

Honeypots techniques have been applied as a detection method for Ransomware [111]. As honeypot deceives the attackers by acting a decoy computer, illicit access can be detected through this technique. Detection of comprised secure socket sell can prevent DoS attack, SSH port scanning, SSH brute-force attack, phishing attack and so on. A detection model using honeypots and machine learning was implemented to detect SSH with malicious activity [131].

#### g: TRIPWIRE IMPLEMENTATION

Credential stealing with the benefit from reusing password could result in site compromise to organizational exploit. Tripwire is another detection method used for monitoring system to notify and trigger when any compromising action could lead to exploit data. Different studies have adopted this technique to detect any comprising threats in the system that indicates attacks and hacking [4], [48].

#### h: ANOMALY DETECTION METHODS

Anomaly detection can be utilized as prevention for security breaches and attacks in the cyber space. Various user profiling models to create security profiles in anomaly detection has been discussed in [35]. Anomaly detection techniques are implemented as a cyber-incident detection method in many studies [5], [102], [132], [162].

#### i: INTRUSION DETECTION TECHNIQUES

Intrusion detection plays important role in detecting threats and attacks by analysing the system logs and network traffic. Different intrusion detection methods with the attack detecting capabilities have been analysed in the study to get more understanding about the techniques [108].

Network based, host based, signature based, anomaly based, specification based, hybrid and physical intrusion detection methods were discussed in [32], [65]. Flow based intrusion detection model was proposed in a study with the analysis of statistical and machine learning techniques [151].

### C. DISCUSSION SUMMARY OF SELECTED STUDIES

With the answers and explanation of our research questions in the above subsections A and B, this subsection summarizes the selected 65 articles contents with their contributions. The contents of the extracted 65 studies are summarized in Table 6 where each ID is referred in Appendix A, Table 11. To summarize the contents according to the contributions despite the different methods implemented, some studies have been grouped together in Table 6.

## IV. DEMOGRAPHIC INFORMATION

This section shows the distribution of the reviewed papers for the types of publications, publications by year and publication numbers based on search criteria. We have also analysed the most widely used keywords based on our studied articles. The most representative authors and the publication sources where the studied articles are most published are also discussed. We have considered the countries of publications where the most cited journals, conferences or books were published.

### A. PUBLICATION TYPE

Figure 3(a) shows the type of publications selected for our SLR. We have utilized three types of articles for our review those include journals, conferences and books. The Figure demonstrates that among 65 selected articles, the majority is journal papers with 37 numbers and 57%, followed by 25 conferences with 38% and the rest 3 books with 5%.

### B. CHRONOLOGICAL VIEW

Figure 4 shows that the reviewed papers were published between the year 2005 and 2019.The papers published before July, 2019 were covered by our SLR study at the completion of our search process for the selection of relevant articles. This Figure describes a rise in the publication numbers on Dark Web field.

### C. PUBLICATION CRITERIA

Figure 3(b) describes the selected review papers criteria wise distribution. Our main two RQs are covered in the majority portions of papers in terms of criteria. The other criteria's are relevant to our RQs as well. The research articles criteria's are Dark Web crimes, Crypto markets, Dark Web crime detection, law enforcement for Dark Web crime and threats from Dark Web.

**TABLE 6.** Overview of extracted studies.

| ID | Contribution |
|---|---|
| 4 | Identifies the needs, challenges and services of Human trafficking victims |
| 9 | Provides the conflicting impacts of Internet in social media for the understanding of mitigation |
| 54,1,11,56 | Evaluation of Silk Road marketplace in terms of criminal activities, trading, freedom speech and analysis of illicit dealings through revenue calculations |
| 55,57,59,7,53,61 | Analysis of the crypto markets focusing on the drug transactions |
| 2,3,15,29 | Overview of Deep Web, Dark Net, anonymous networks, forums, marketplaces , cybercrimes and monitoring techniques and crypto tools |
| 12 | Detect threats in Dark Net data using cyber threat intelligence tool through Dark Net ecosystem network and cyber threat breach network |
| 5,8 | Analysis of the cybercrimes related to prostitution and child prostitution in the Internet or online applications and the possible detection techniques |
| 6,17,60 | Development of Crawler for the Dark Net marketplaces and analysis of the illicit drug trading |
| 10,65 | Analysis of Terrorism and the impacts in Dark Web |
| 13,58 | Cyber threat intelligence gathering for malicious hacking with crawler and machine learning techniques |
| 14 | Development of Crawler and automatic framework to detect cyber threats for exploit markets and hacker forums in the Dark Net |
| 16 | Development of framework for top malware sellers identification in the underground forum |
| 19,18 | Exploring the cybercrime on hacking and data thieves in the stolen data markets through the analysis of the products, services ,trading procedures and revenue calculations |
| 20,21 | Overview of Bitcoin, uses, misusage, impacts and law enforcements |
| 27 | Analysis of traffic classification on anonymity of TOR I2P and JonDonym |
| 31,33,34,35 | Development of sock puppets detection method in online communities and social platforms |
| 32 | Development of alias classification and authorship detection in the Dark net marketplace |
| 38 | Development of a multilingual Dark Web forum portal for analyzing Jihadist forums |
| 36 | Analysis of law enforcement in the Dark Web with the investigation technique, legal policies and short comings |
| 40,25,41,42 | Analysis of terrorists websites in the Dark Web to determine the characteristics and terrorist activity |
| 62,63 | Development of Crawler for the analysis of services and products distributions of Dark Net marketplaces |
| 23,48,47,49,64 | Traffic network analysis for monitoring cyber attacks |
| 24,26,37 | Analysis of the extremists groups in the Dark Web forums |
| 28 | Analysis of fingerprinting techniques on TOR to get the hidden service connections and finding out the hidden services accessed |
| 39,43 | Design and implementation of crawler for TOR network and analysis of extremist and terrorist groups in Dark Web |
| 44,45,46 | Deployment of honeypot to determine attacks and attackers in the network. |
| 50 | Development of a Tor client with significant latency gains and snooping avoidance |
| 51,52 | Analysis attacks against Tor with the impacts and solutions based on experiments |
| 22 | Analysis and evaluation of different intrusion detection techniques |
| 30 | Detection method for fraud transactions by malicious attackers |

**TABLE 7.** Widely used keywords.

| Keywords | Total Papers | Sample ID |
|---|---|---|
| Dark Net, Dark Web, Deep Web | 32 | 3,26,55 |
| Anonymous, Anonymity | 35 | 1,20 |
| Crypto markets , Bit coin | 20 | 20,54 |
| Onion Routing, TOR | 30 | 3,52 |
| Hacking, Hacker, Malicious | 18 | 18,15 |
| Drugs Markets, Illicit Drugs | 20 | 54,55,57 |
| Terrorism, Cyber Threat Intelligence | 15 | 26,12 |
| Cyber Crime, Law Enforcement, Criminal, Criminal Activities, Attacks, Hidden Services, Marketplaces, Hidden Marketplace, Black Markets, Dark Net Markets | 10 | 3,11,18,44,57,60 |

### D. WIDELY USED KEYWORDS

The most widely used keywords from our studied 65 articles are represented in Table 7. These keywords are also used as search terms for searching a study of interest in this field. Keywords may vary from article to article but most popular and utilized keywords are common. However these keywords are sometimes presented in the articles with different parts of speech and tenses. Table 7 represents the most widely used keywords and the number of papers that used the keywords in their articles.

### E. REPRESENTATIVE AUTHORS

The most representative authors have been selected in terms of citation of the papers in the field of our study. We have analysed 10 top most articles with highest number of citations from the 65 selected papers for this analysis. According to our analysis Nicolas Christin has the most cited article on this field of study [ID 1, ID 11]. Although some of the authors have published in different field of studies, we considered the popularity based on the citation numbers on our field of study. Table 8 shows the most representative authors with the citation numbers and ID from Appendix A.

### F. POPULAR DATABASES

The statistics of extracting the articles from six databases described in subsection II.B shows that Google Scholar contains the most available resources for the study following by the IEEE Xplore. The number of articles extracted from the databases of Google Scholar, IEEE Xplore, ScieneDirect, Springer, Scopus and ACM Digital Library are 22, 18,6,9,5 and 5 respectively.

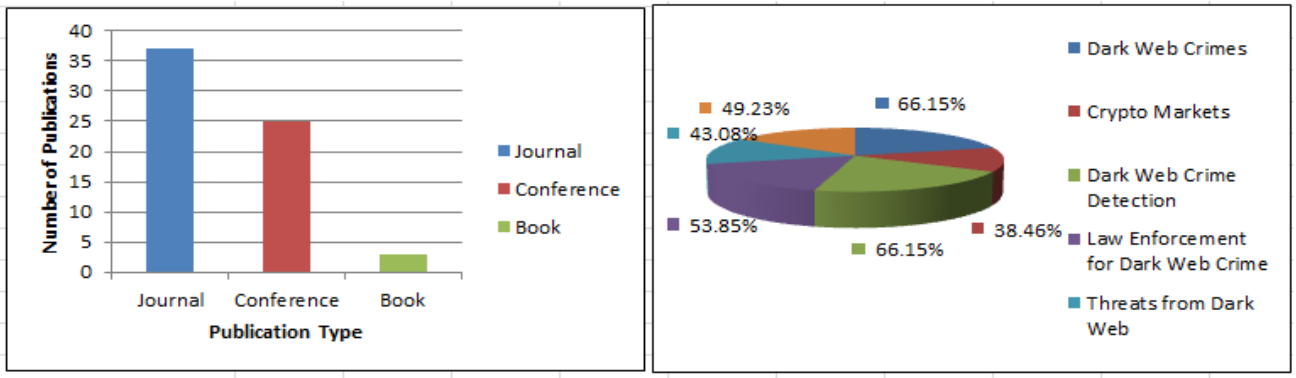


**FIGURE 3.** a) Number of Papers per Category and b) Number of Papers per Criteria.

**TABLE 8.** Most representative authors.

| ID | No of Citations | Authors |
|---|---|---|
| 1 | 672 | Nicolas Christin |
| 20 | 444 | Reuben Grinberg |
| 11 | 324 | Kyle Soska & Nicolas Christin |
| 18 | 230 | Thomas J Holt & Eric Lampke |
| 54 | 209 | Judith Aldridge & David Décary-Hétu |
| 55 | 208 | James Martin |
| 52 | 208 | Nathan S. Evans, Roger Dingledine & Christian Grothoff |
| 26 | 197 | Yilu Zhou , Edna Reid, Jialun Qin, Hsinchun Chen & Guanpi Lai |
| 3 | 163 | Daniel Moore & Thomas Rid |
| 50 | 120 | Masoud Akhoondi , Curtis Yu & Harsha V. Madhyastha |

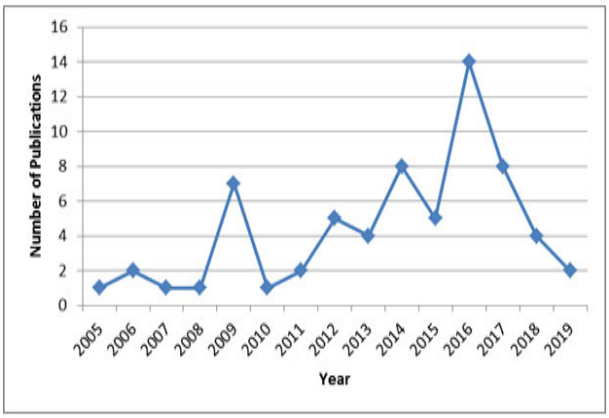


**FIGURE 4.** Number of Selected Publications per Year.

### G. MOST PUBLISHED SOURCES AND COUNTRIES

We have analysed the 65 selected articles and selected the top 10 publication sources according to the citation number. The United State of America publishes and has the headquarters of the highest number of publishing journals, conferences and books. United Kingdom has the second highest number of countries that publishes the articles on our field of study. Table 9 represents the top 10 published sources and countries.

**TABLE 9.** Most published sources and countries.

| ID | Publication Source | Country of Publication |
|---|---|---|
| 1 | International Conference on World Wide Web | Rio de Janeiro, Brazil |
| 20 | Hastings Science and Technology Law Journal | California, USA |
| 11 | USENIX Security Symposium | Washington, D.C ,USA |
| 18 | Criminal Justice Studies: A Critical Journal of Crime, Law and Society | United Kingdom |
| 54 | Social Science Research Network(SSRN) | Mableton ,Georgia ,USA |
| 55 | Palgrave Macmillan | London ,United Kingdom |
| 52 | USENIX Security Symposium | California, USA |
| 26 | IEEE Intelligent Systems | USA |
| 3 | Survival: Global Politics and Strategy | United Kingdom |
| 50 | IEEE Symposium on Security and Privacy | San Francisco, CA ,USA |

## V. ARCHITECTURAL ANALYSIS

This section reports the architectural description of the selected Dark Web article's and describes the architectural frameworks. The overall architectural view is shown in Figure 5. It is to be noted that the various elements in the architectures are not limited to what is shown in the figure.

From the reviewed articles we have found different architecture frameworks based on different crime theme analysis. All the details are discussed in Section III in details. For analysis of the overall architectures of the papers Figure 5 represents the example. Studies are focused on different themes such as Drug transactions, terrorists and ISIS, human trafficking, pornography industry, money laundering and so on. Based on these themes architectural frameworks have been created. The frame works are divided into two categories one for detection methods that includes network analysis, hash

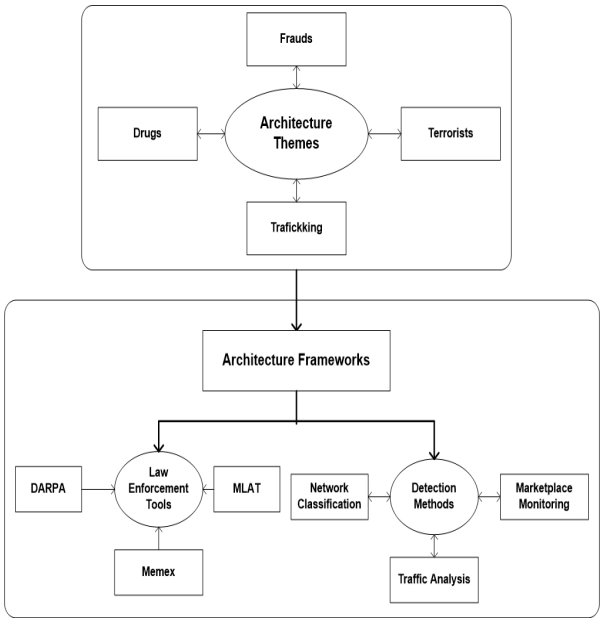


**FIGURE 5.** Overview of Architectural Analysis.

value analysis, marketplace monitoring and scraping and the other for law enforcement tools such as DARPA,MLAT and so on.

The analysis of architectural framework steps and outcomes are shown in Figure 6. To analyse the existing models obtained from the studies Figure 6 describes the generalized steps used in the models applied for different architectures. The elements of the Figures are identified from the architectures implemented in different models in the studied papers listed in Appendix A. This diagram helps to understand each key component in a model and their examples and importance. This also refers to the outcomes of each step to better understand the key criteria's. The generalized elements are described below.

*Data Collection:* This component describes the sources of data, the size of data set and the availability of the data set used for the models. The choice of collected data varies from researchers to enterprises. Many researches have used already crawled data from available online databases, many have crawled the onion sites as their research where others have used TOR traffics as their data set.

*Data Pre-processing:* This is a crucial step to process the data. Important feature selection, filtering, extraction and duplicate or noise removals are the main components of this step. In most of the studies this step is followed with their relevant requirements to feed the model.

*Data Processing:* This is the focal stage of any model which includes the algorithm implementation such as machine learning, classifications of the data such as clustering or labelling, checking the performance of the algorithms with training and testing data and applying the techniques to their respective fields.

*Results:* Depending on the implemented model the results are the final outcome targeted to create the framework. This leads to the analysis of the crimes using available software's such as Maltego, detection of crimes, locating crimes or criminals, giving system alerts of malicious activities through email or mobile applications.

The security agencies or law enforcement can take actions depending on the models outcome which is the ultimate goal of all the frameworks. It is to be noted that the techniques and examples given are not limited to the Figure 6 and the description.

We will analyse some techniques that use different models and tools from our reviewed papers based on the architectures as below.

### A. MACHINE LEARNING TECHNIQUE [ID 13]

*Introduction and Motivation:* Machine learning algorithms could play a very important role in Monitoring and detecting malicious activities. An operational system was proposed for gathering information on malicious hacking products and services from various social platforms on the Dark Net and Deep Net discussion forums and marketplaces

*Description:* They have used 10 marketplaces and two forums data for their experiment. At first, they crawled the marketplace and forum data then parsed the data for classification which could determine different topics from forums and products from marketplace which are related to malicious hacking. Machine learning algorithms of supervised and semi-supervised methods were applied for both marketplace and forums. They implemented Naive Bayes(NB), random forest (RF), support vector machine (SVM)and logistic regression (LOG-REG) algorithms as supervised and label propagation, Co-training approaches as semi supervised algorithms.

*Result and Constrains:* The proposed system collects on average 305 high-quality cyber threat warnings each week.

### B. AFFECT ANALYSIS TECHNIQUE [ID 24]

*Introduction and Motivation:* Affect analysis of the extremists groups in the Dark Web forums leads to the violence, hate and propagandas spread over there. To gather the affect related contents over extremist groups from the Dark Web forums an affect lexicon was constructed with probabilistic disambiguation technique in this study.

*Description:* Extremist forum data were collected from University of Arizona for their experiment which consisted of messages from 16 US and Middle Eastern forums. The

system consisted of two components of lexicon creation and analysis technique. At first they collected the messages and to remove affect ambiguity applied probabilistic disambiguation from the term list of collected messages to create the affect lexicon database. In the Affect Analysis phase several filtering and duplicate removal are done for affect parsing and getting intensity scores and correlations.

*Result and Constrains:* From the US forums they found racism other than hate and violence whereas the Middle East forums are rich in hate and violence. The analysis did

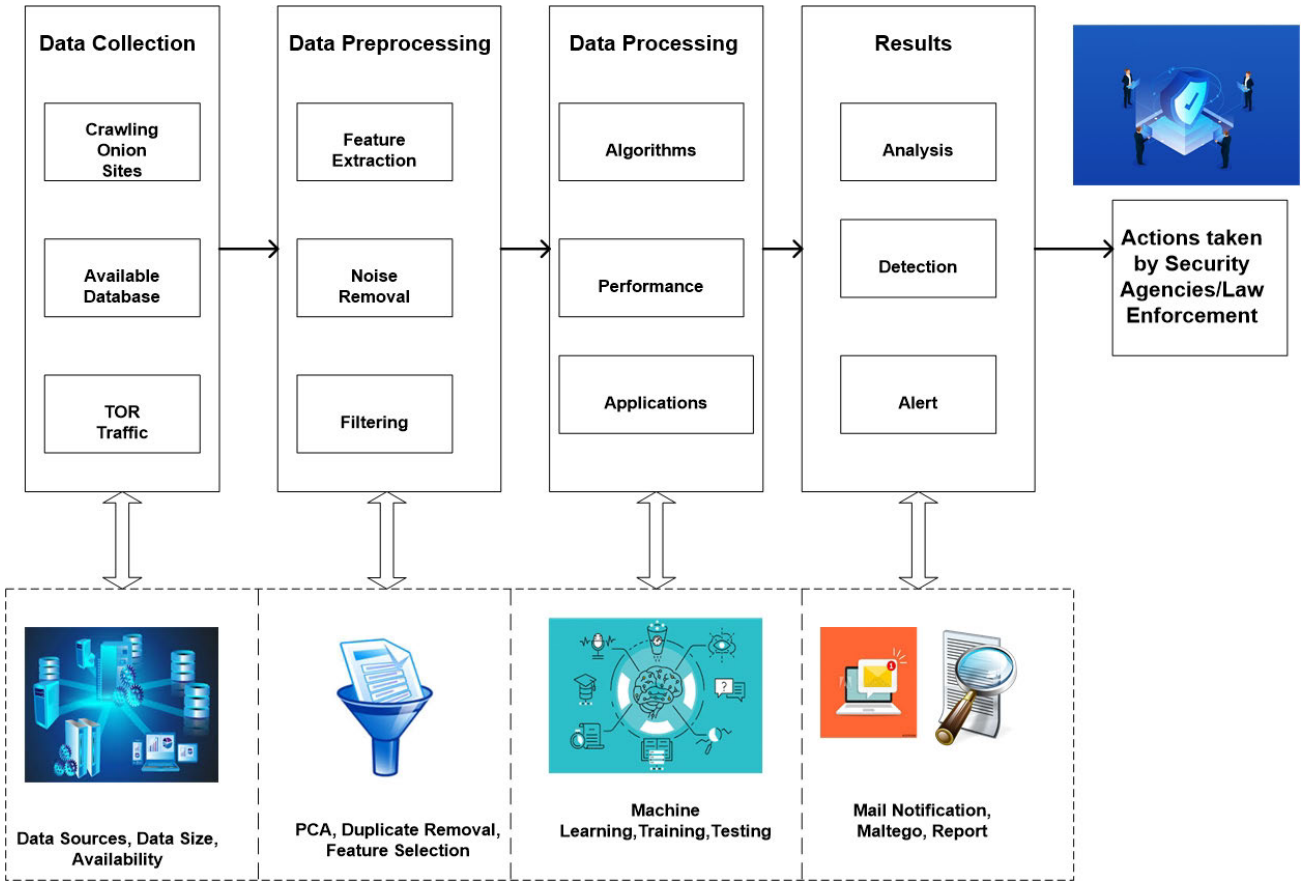


**FIGURE 6.** Overview of Architectural Framework Analysis.

not consider the forum posts which contain rich amount of features to extract data for such analysis.

### C. FINGERPRINTING TECHNIQUE [ID 28]

*Introduction and Motivation:* To gather the encrypted contents and anonymized connections website fingerprinting attacks play important role. This research aims to apply fingerprinting attacks on TOR to get the hidden service connections and finding out the hidden services accessed.

*Description:* To gather the data which is the Hidden service addresses they crawled all known hidden services search engines with a automatic tool per day from which they collected 13,243 unique addresses in four month. Alongside this, they exploited the hidden service directories in the protocols directly from the TOR network in live to gather the onion addresses. 3.6.2 TBB was utilized to access the hidden services and regular webpages via Tor. For data extraction and processing they grouped several pages together in order to avoid subpages that contain blank pages or little information. The proposed fingerprinting approach consists of two classification phases. The first phase detects the communication of clients with the hidden service and the second phase, if connection is detected it recognizes which hidden service content has been visited by the client.

*Result and Constrains:* The feasibility of a connection establishment to a hidden service is shown in their dataset to recognize the particular content which is not scalable to the existing methods applied in realistic settings.

### D. AUTHORSHIP ANALYSIS TECHNIQUE [ID 32]

*Introduction and Motivation:* To identify the suspects in the Dark Web who utilizes the anonymity service such as TOR authorship detection is necessary step in the cyber forensic. This article proposed classification setups on two major tasks of alias classification and authorship attribution for user identification in the drug trafficking area of Dark Web.

*Description:* The experimental data was collected from Black Market Reloaded (BMR) forum in the Dark Net. They collected database consists of 92,333 posts from 8,348 different users, posted in 12,923 different threads in multiple languages. English language threads and more than 250 threads posted and less than 50 threads posted were excluded while data filtering stage. For the first classification phase of alias classification they applied stylometric analysis. Topic independent features were selected for their stylometric analysis and character n-grams as features to profile users were also explored with time based features of the forum posts. For the authorship attribution a multiclass classification was

implemented. For this step they combined multiple randomly selected posts into a single instance, to represent a user's writing style and irregular or uninformative posts were ignored. They have implemented Support vector machine (SVM) as their classification algorithm.

*Result and Constrains:* In the alias classification to decide if two user accounts are aliases or not precision result of 91% at a recall of 25% were obtained for 177 users whereas for 25 users the precision was same but the recall reached 45%. The result of authorship attribution varies with the number of users. For 177 users the classification accuracy with SVM was 88% but for fewer users it could gain 94% accuracy. If any users use different style of writing in different posts this method will get biased based with the stylometric analysis thus considering more features beside character n-gram and time based is required.

### E. SOCIAL NETWORK ANALYSIS TECHNIQUE [ID 41]

*Introduction and Motivation:* social networks analysis and text mining in the terrorism activities contributes to develop methods that can fight back terrorism. This study aims to combine social network analysis and data mining techniques for identifying key members of a community for measuring social aspects which can trigger the further analysis on experts' detection, sub-groups detection, centrality measures and so on.

*Description:* The experiment was done on English Dark web forum named Ansar1. Posts were created by 376 members and extracted topics where realized over 29057 posts for this experiment. The main focus of this article is to discover the key members talking about particular subjects in the Dark web related to security threats based on a topic-based social network structure. For this they used a hybrid approach of social network analysis with latent semantic-based text mining that could lead to threatening topics and perform specific analysis on each one. For social network analysis the network configuration was done based on members' interactions through post replies. They considered two approaches for network conversion of Creator oriented Network and Last Reply oriented Network. According to each network configurations, the topic-based filtering using LDA method is used to remove all replies that are not according to the posts' topic. Key-members where determined using Hyperlink-Induced Topic Search (HITS)

*Result and Constrains:* The result obtained 47 topics using the closest 30 words on each topic and 11 topics were discarded afterwards. However, the information obtained from their network analysis could not lead to identify any specific patterns.

### F. IDENTITY DECEPTION DETECTION TECHNIQUE [ID 35]

*Introduction and Motivation:* In social media environment identity deception is increasing and detection of such identity can also play important roles in the Dark web forensics. A detection method using non-verbal behaviour for identity deception in social media is proposed in this study.

**TABLE 10. Comparative analysis of architectural techniques.**

| ID | Data Collection | Data Pre-processing | Data Processing | Result |
|---|---|---|---|---|
| 13 | Crawling | Parser for text filtering | Classification with machine learning | Collect cyber threat warning |
| 24 | Available Database | Probabilistic disambiguation for affect ambiguity | Parser for analysis | Racism, violence, and hate affects from extremist groups |
| 28 | Crawling | Grouping web pages for blank pages | Classification with fingerprinting | Detect hidden services and communications |
| 32 | Available Database | Stylometric , character level n-gram, and time based features | Classification with machine learning | User identification in TOR network |
| 41 | Available Database | Topic modeling for network filtering | Network modeling for analysis | discover potential threats based on topics |
| 35 | Available Database | Regular expression for time-independent time-dependent variable selection | Classification with machine learning | Detect deception by non-verbal behavior monitoring |
| 17 | Crawling | Apple script scrapping | Analysis with Maltego tool | Get vendor account information and vendor's origin |

*Description:* Publicly available data for Wikipedia has been used for the proposed method. Non-verbal user behaviour is monitored for deception detection using machine learning algorithms in this study. The logged data on the Wikipedia page revisions are considered as non-verbal user behaviour such as time taken between each revision. Variables of online non-verbal behaviour were divided into time-independent and time-dependent. For data retrieval from the available logs of blocked accounts, regular expressions were used to keep only sockpuppet cases with an infinite time of block issued for these accounts. For testing 7,500 cases of sockpuppets samples were considered.

*Result and Constrains:* Selecting the variables for non-verbal behaviour may vary in the result accuracy. Also new users' time window set will have significant impact on the method's effectiveness. Thus a combination of verbal detection deception with the non-verbal behaviour deception method is necessary.

### G. DARK WEB SCRAPPING TECHNIQUE [ID 17]

*Introduction and Motivation:* Scrapping the Dark Web marketplaces and forums can lead to important information that

**TABLE 11.** List of Primary Reviewed Articles.

| ID | Title | Author(s) | Year |
|---|---|---|---|
| 1 | Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace | Nicolas Christin | 2013 |
| 2 | Deepweb and Cybercrime | Vincenzo Ciancaglini, Marco Balduzzi, Max Goncharov & Robert McArdle, | 2013 |
| 3 | Cryptopolitik and the Darknet | Daniel Moore & Thomas Rid | 2016 |
| 4 | Addressing the needs of victims of human trafficking: Challenges, barriers, and promising practices | Heather J. Clawson & Nicole Dutch | 2008 |
| 5 | Deviant Men, Prostitution, and the Internet: A Qualitative analysis of Men who killed Prostitutes whom they met online | Kelly Beckham & Ariane Prohaska | 2012 |
| 6 | Tor marketplaces exploratory data analysis: the drugs case | Alessandro Celestin, Gianluigi Me & Mara Mignone | 2017 |
| 7 | Buying drugs on a Darknet market: A better deal? Studying the online illicit drug market through the analysis of digital, physical and chemical data | Rhumorbarbe Damiena, Staehli Ludovica, Broséus Juliana, Rossy Quentina & Esseiva Pierre | 2016 |
| 8 | Webcam child prostitution: An exploration of current and futuristic methods of detection | Kemal Veli Açar | 2017 |
| 9 | LikeWar: The Weaponization of Social Media | Peter Warren Singer & Emerson T Brooking | 2018 |
| 10 | Going dark: Terrorism on the dark web | Gabriel Weimann | 2016 |
| 11 | Measuring the longitudinal evolution of the online anonymous marketplace ecosystem | Kyle Soska & Nicolas Christin, | 2015 |
| 12 | Dark-Net Ecosystem Cyber-Threat Intelligence (CTI) Tool | Nolan Arnold, Mohammadreza Ebrahimi, Ning Zhang, Ben Lazarine, Mark Patton & Hsinchun Chen | 2019 |
| 13 | Darknet and deepnet mining for proactive cybersecurity threat intelligence | Eric Nunes, Ahmad Diab, Andrew Gunn, Ericsson Marin , Vineet Mishra,Vivin Paliath, John Robertson, Jana Shakarian, Amanda Thart & Paulo Shakarian | 2016 |
| 14 | Darknet mining and game theory for enhanced cyber threat intelligence | John Robertson, Ahmad Diab, Ericsson Marin, Eric Nunes, Vivin Paliath, Jana Shakarian & Paulo Shakarian | 2016 |
| 15 | Exploring malicious hacker forums | Jana Shakarian ,Andrew T. Gunn & Paulo Shakarian | 2016 |
| 16 | Identifying top sellers in underground economy using deep learning-based sentiment analysis | Weifeng Li & Hsinchun Chen | 2014 |
| 17 | A Framework for More Effective Dark Web Marketplace Investigations | Darren Hayes, Francesco Cappa & James Cardon | 2018 |
| 18 | Exploring stolen data markets online: products and market forces | Thomas J Holt & Eric Lampke | 2010 |
| 19 | Exploring and estimating the revenues and profits of participants in stolen data markets | Thomas J Holt, Olga Smirnova & Yi Ting Chua | 2016 |

**TABLE 11.** *(Continued)* List of Primary Reviewed Articles.

| ID | Title | Author(s) | Year |
|---|---|---|---|
| 20 | Bitcoin: An innovative alternative digital currency | Reuben Grinberg | 2012 |
| 21 | Bitcoin, silk road, and the need for a new approach to virtual currency regulation | Jonathan Lane | 2013 |
| 22 | Intrusion detection techniques in cloud environment: A survey | Preeti Mishra, Emmanuel S. Pilli, Vijay Varadharajan & Udaya Tupakula | 2017 |
| 23 | Large-Scale Monitoring for Cyber Attacks by UsingCluster Information on Darknet Traffic Features | Hironori Nishikaze, Seiichi Ozawa, Jun Kitazono, Tao Ban, Junji Nakazato & Jumpei Shimamura | 2015 |
| 24 | Affect intensity analysis of dark web forums | Ahmed Abbasi & Hsinchun Chen | 2007 |
| 25 | Automatic detection of cyber-recruitment by violent extremists | Jacob R Scanlon & Matthew S Gerber | 2014 |
| 26 | US domestic extremist groups on the Web: link and content analysis | Yilu Zhou , Edna Reid, Jialun Qin, Hsinchun Chen & Guanpi Lai | 2005 |
| 27 | Anonymity services Tor, I2P, JonDonym: classifying in the dark | Antonio Montieri, Domenico Ciuonzo, Giuseppe Aceto & Antonio Pescapé | 2017 |
| 28 | Analysis of fingerprinting techniques for tor hidden services | Andriy Panchenko, Asya Mitseva, Martin Henze, Fabian Lanze, Klaus Wehrle & Thomas Engel | 2017 |
| 29 | Anonymous connections based on onion routing: A review and a visualization tool | Abdullah A. AlQahtani & El-Sayed M. El-Alfy | 2015 |
| 30 | Fraud and financial crime detection model using malware forensics | Ae Chan KimSeongkon KimWon Hyung Park & Dong Hoon Lee | 2014 |
| 31 | A sock puppet detection algorithm on virtual spaces | Zhan Bu ,Zhengyou Xia & Jiandong Wang | 2013 |
| 32 | Authorship analysis on dark marketplace forums | Martijn Spitters , Femke Klaver , Gijs Koot & Mark van Staalduinen | 2015 |
| 33 | Sockpuppet gang detection on social media sites | Dong Liu, Quanyuan Wu ,Weihong Han & Bin Zhou | 2016 |
| 34 | An army of me: Sockpuppets in online discussion communities | Srijan Kumar, Justin Cheng, Jure Leskovec & V.S. Subrahmanian | 2017 |
| 35 | Multiple account identity deception detection in social media using nonverbal behavior | Michail Tsikerdekis & Sherali Zeadally | 2014 |
| 36 | Searching Places Unknown: Law Enforcement Jurisdiction on the Dark Web | Ahmed Ghappour | 2017 |
| 37 | The dark side of the web: Italian right-wing extremist groups and the Internet | Manuela Caiani & Linda Parenti | 2009 |
| 38 | Dark web forums portal: searching and analyzing jihadist forums | Yulei Zhang, Shuo Zeng, Li Fan, Yan Dang, Catherine A. Larson & Hsinchun Chen | 2009 |
| 39 | Surfacing collaborated networks in dark web to find illicit and criminal content | Ahmed T. Zulkarnine, Richard Frank, Bryan Monk,Julianna Mitchell & Garth Davies | 2016 |

**TABLE 11.** *(Continued)* List of Primary Reviewed Articles.

| ID | Title | Author(s) | Year |
|---|---|---|---|
| 40 | On the topology of the dark web of terrorist groups | Jennifer Xu, Hsinchun Chen ,Yilu Zhou & Jialun Qin | 2006 |
| 41 | Topic-based social network analysis for virtual communities of interests in the dark web | Gaston L'Huillier, Hector Alvarez, Sebastián A. Ríos & Felipe Aguilera | 2011 |
| 42 | Dark web portal overlapping community detection based on topic models | Sebastián A. Ríos & Ricardo Muñoz | 2012 |
| 43 | Discovering topics from dark websites | Li Yang , Feiqiong Liu , Joseph M. Kizza & Raimund K. Ege | 2009 |
| 44 | Attacks landscape in the dark side of the web | Onur Catakoglu, Marco Balduzzi & Davide Balzarotti | 2017 |
| 45 | Honeypots deployment for the analysis and visualization of malware activity and malicious connections | Ioannis Koniaris , Georgios Papadimitriou , Petros Nicopolitidis & Mohammad Obaidat | 2014 |
| 46 | Detecting Ransomware with Honeypot techniques | Chris Moore | 2016 |
| 47 | Detecting DDoS attacks against data center with correlation analysis | Peng Xiao ,Wenyu Qu, Heng Qi & Zhiyang Li | 2015 |
| 48 | Correlation-based traffic analysis attacks on anonymity networks | Ye Zhu , Xinwen Fu ,Bryan Graham , Riccardo Bettati & Wei Zhao | 2009 |
| 49 | Locating network domain entry and exit point/path for DDoS attack traffic | Vrizlynn L. L. Thing , Morris Sloman & Naranker Dulay | 2009 |
| 50 | LASTor: A low-latency AS-aware Tor client | Masoud Akhoondi , Curtis Yu & Harsha V. Madhyastha | 2012 |
| 51 | A new cell-counting-based attack against Tor | Zhen Ling, Junzhou Luo, Wei Yu, Xinwen Fu, Dong Xuan & Weijia Jia, | 2012 |
| 52 | A Practical Congestion Attack on Tor Using Long Paths | Nathan S. Evans, Roger Dingledine & Christian Grothoff | 2009 |
| 53 | Safer scoring? Cryptomarkets, social supply and drug market violence | Monica J. Barratt, Jason A. Ferris & Adam R. Winstock | 2016 |
| 54 | Not an 'Ebay for Drugs': The Cryptomarket 'Silk Road' as a Paradigm Shifting Criminal Innovation | Judith Aldridge & David Décary-Hétu | 2014 |
| 55 | Drugs on the dark net: How cryptomarkets are transforming the global trade in illicit drugs | James Martin | 2014 |
| 56 | Constructive activism in the dark web: cryptomarkets and illicit drugs in the digital 'demimonde' | Alexia Maddox, Monica J. Barratt, Matthew Allen & Simon Lenton | 2016 |
| 57 | Hidden wholesale: The drug diffusing capacity of online drug cryptomarkets | Judith Aldridge, and David Décary-Hétu | 2016 |
| 58 | Analysis of Hacking Related Trade in the Darkweb | Othmane Cherqi , Ghita Mezzour ,Mounir Ghogho & Mohammed El Koutbi | 2018 |
| 59 | Do Police Crackdowns Disrupt Drug Cryptomarkets? A Longitudinal Analysis Of The Effects Of Operation Onymous | David Décary-Hétu & Luca Giommoni | 2017 |
| 60 | Comparing cryptomarkets for drugs. A characterisation of sellers and buyers over time | Meropi Tzanetakis | 2018 |

**TABLE 11.** *(Continued)* **List of Primary Reviewed Articles.**

| ID | Title | Author(s) | Year |
|---|---|---|---|
| 61 | Challenging the techno-politics of anonymity: the case of cryptomarket users | Angus Bancroft & Peter Scott Reid | 2016 |
| 62 | Mining the Dark Web: drugs and fake ids | Andres Baravalle, Mauro Sanchez Lopez & Sin Wee Lee | 2016 |
| 63 | Python Scrapers for Scraping Cryptomarkets on Tor | Yubao Wu, Fengpan Zhao, Xucan Chen,Pavel Skums, Eric L. Sevigny, David Maimon, Marie Ouellet, Monica Haavisto Swahn ,Sheryl M. Strasser, Mohammad Javad Feizollahi, Youfang Zhang & Gunjan Sekhon | 2019 |
| 64 | Practical Correlation Analysis between Scan and Malware Profiles against Zero-Day Attacks Based on Darknet Monitoring | Koji Nakao, Daisuke Inoue, Masashi Eto & Katsunari Yoshioka | 2009 |
| 65 | DEEP WEB, DARKWEB, INVISIBLE WEB AND THE POST ISIS WORLD | Ryan Ehney & Jack D. Shorter | 2016 |

can help law enforcement and security agencies to investigate and forensic. The article developed an automatic scraping framework and analysed with free tools to investigate the scrapped Dark Web marketplaces

*Description:* To find the lists of the onion URLs of the top marketplaces they used of Reddit and DeepDotWeb websites. They have proposed a web crawler which is written in Apple script. The script could keep track of all the accounts and add only new accounts not listed in the crawling. After scrapping the marketplace sites Maltego tool was used to identify suspected vendors on the focus drug marketplace in the Dark Web obtained from the scrapping results. They manually analysed the marketplace and over 3000 distinct vendors operating on the marketplace.

*Result and Constrains:* The Web crawler could give the account information for more than 3000 vendors from the drug marketplace and four most active vendors' origins using Maltego were identified. They did not engage in dialogue with any vendor and did not purchase any products from the marketplace which questions the authenticity of the vendors' identification. Moreover, manual entry in the Maltego software is erroneous and time consuming.

### H. COMPARATIVE ANALYSIS

To analyse the architecture of some selective techniques we have explained the methods in this section under subsection V.A to subsection V.G. To evaluate the contents of the selected papers (ID 13,17,24,28,32,35,41) the performance analysis is done in a summative way presented in Table 10. The table replicates the architectural frame work steps defined in Figure 6 with the respective article ID, the data collection procedure, the data pre-processing and processing methods and the final outcome.

## VI. CONCLUSION

In precise,this SLR incorporate a comprehensive description of the Dark Net crime threats, the technical and forensic challenges with the anonymous network structures and the detection methods, algorithms, tools and strategies applied for locating the crimes and criminals in the Dark Web. Cyber criminals are becoming more quick-witted against the enforced methods to detect them inside Dark Web. As a result challenges are elevated. For law enforcement and security agencies international boarder is one of the most hindering task. The sheer size of the hidden web necessitates more effective approaches to minimise the potential threats from the Dark Web. The black marketplace and transactions happening there must be traced to detect the criminals with advanced methods. The unindexed, fragmented and multilayer structure of Dark Web makes it more strenuous to detect the crimes. The ecosystem of the Dark Web being highly unpredictable as every day the old sites keep vanishing while new sites appear,strong digital evidences are required to collect for the forensic law agencies to ensure overcoming the obstacles in arresting and sizing the criminals.

## APPENDIX A
## PRIMARY STUDIES SELECTED FOR THIS SLR

See Table 11.

## REFERENCES

[1] A. Abbasi and H. Chen, "Affect intensity analysis of dark Web forums," presented at the IEEE Intell. Secur. Informat., May 2007.

[2] K. V. Açar, "Webcam child prostitution: An exploration of current and futuristic methods of detection," *Int. J. Cyber Criminol.*, vol. 11, no. 1, pp. 98–109, 2017.

[3] A. Afilipoaie and P. Shortis, "From dealer to doorstep-How drugs are sold on the dark net," GDPO Situation Anal., Swansea Univ., Global Drugs Policy Observatory, Swansea, U.K., Tech. Rep., 2015.

[4] I. Agrafiotis, A. Erola, M. Goldsmith, and S. Creese, "A tripwire grammar for insider threat detection," presented at the Int. Workshop Manag. Insider Secur. Threats (MIST), 2016.
[5] M. Ahmed, A. N. Mahmood, and M. R. Islam, "A survey of anomaly detection techniques in financial domain," *Future Gener. Comput. Syst.*, vol. 55, pp. 278–288, Feb. 2016.
[6] M. Akhoondi, C. Yu, and H. V. Madhyastha, "LASTor: A low-latency AS-aware tor client," presented at the IEEE Symp. Secur. Privacy, May 2012.
[7] M. W. Al Nabki, E. Fidalgo, E. Alegre, and I. de Paz, "Classifying illegal activities on tor network based on Web textual contents," presented at the 15th Conf. Eur. Chapter Assoc. Comput. Linguistics, vol. 1, 2017.
[8] J. Aldridge and D. Décary-Hétu, "Hidden wholesale: The drug diffusing capacity of online drug cryptomarkets," *Int. J. Drug Policy*, vol. 35, pp. 7–15, Sep. 2016.
[9] J. Aldridge and D. Décary-Hétu, "Not an'Ebay for Drugs': The cryptomarket'silk road'as a paradigm shifting criminal innovation," Social Sci. Res. Netw. (SSRN), The Netherlands, Tech. Rep., 2014.
[10] A. A. AlQahtani and E.-S.-M. El-Alfy, "Anonymous connections based on onion routing: A review and a visualization tool," *Procedia Comput. Sci.*, vol. 52, pp. 121–128, Jan. 2015.
[11] R. Armstrong, B. J. Hall, J. Doyle, and E. Waters, "'Scoping the scope' of a cochrane review," *J. Public Health*, vol. 33, no. 1, pp. 147–150, Mar. 2011.
[12] N. Arnold, M. Ebrahimi, N. Zhang, B. Lazarine, M. Patton, H. Chen, and S. Samtani, "Dark-net ecosystem cyber-threat intelligence (CTI) tool," presented at the IEEE Int. Conf. Intell. Secur. Informat. (ISI), Jul. 2019.
[13] B. Backman, "Follow the white rabbit: An ethnographic exploration into the drug culture concealed within the 'deep web,'" Univ. Nebraska Omaha, Omaha, NE, USA, Tech. Rep. UMI 1551711, 2013.
[14] M. Bailey, E. Cooke, F. Jahanian, A. Myrick, and S. Sinha, "Practical darknet measurement," presented at the 40th Annu. Conf. Inf. Sci. Syst., Mar. 2006.
[15] A. Bancroft and P. Scott Reid, "Challenging the techno-politics of anonymity: The case of cryptomarket users," *Inf., Commun. Soc.*, vol. 20, no. 4, pp. 497–512, Apr. 2017.
[16] A. Baravalle, M. S. Lopez, and S. W. Lee, "Mining the dark Web: Drugs and fake IDS," presented at the IEEE 16th Int. Conf. Data Mining Workshops (ICDMW), Dec. 2016.
[17] M. J. Barratt, J. A. Ferris, and A. R. Winstock, "Safer scoring? Cryptomarkets, social supply and drug market violence," *Int. J. Drug Policy*, vol. 35, pp. 24–31, Sep. 2016.
[18] R. Bates, "Tracking lone wolf terrorists," *J. Public Prof. Sociol.*, vol. 8, no. 1, p. 6, 2016.
[19] K. Bauer, D. McCoy, D. Grunwald, T. Kohno, and D. Sicker, "Low-resource routing attacks against tor," presented at the ACM Workshop Privacy Electron. Soc. (WPES), 2007.
[20] K. Beckham and A. Prohaska "Deviant men, prostitution, and the Internet: A qualitative analysis of men who killed prostitutes whom they met online," *Int. J. Criminal Justice Sci.*, vol. 7, no. 2, pp. 635–648, 2012.
[21] H. Bleau. (Apr. 24, 2019). Social Media and the Digital Transformation of Cybercrime. RSA Security. [Online]. Available: https://www.rsa.com/en-us/blog/2019-04/social-media-and-the-digital-transformation-of-cybercrime
[22] F. Breitinger, K. P. Astebol, H. Baier, and C. Busch, "MvHash-B—A new approach for similarity preserving hashing," presented at the 7th Int. Conf. IT Secur. Incident Manage. IT Forensics, Mar. 2013.
[23] F. Breitinger and H. Baier, "Similarity preserving hashing: Eligible properties and a new algorithm MRSH-v2," presented at the Int. Conf. Digit. Forensics Cyber Crime, 2012.
[24] S. Brown (Apr. 30, 2019). Cybercriminals ramping up fraud attacks on social media, says report. Cnet. [Online]. Available: https://www.cnet.com/news/cybercriminals-are-ramping-up-fraud-attacks-on-social-media-says-report/
[25] D. Bryans, "Bitcoin and money laundering: Mining for an effective solution," *Ind. LJ*, vol. 89, no. 1, p. 441, 2014.
[26] Z. Bu, Z. Xia, and J. Wang, "A sock puppet detection algorithm on virtual spaces," *Knowl.-Based Syst.*, vol. 37, pp. 366–377, Jan. 2013.
[27] J. Buxton and T. Bingham, "The rise and challenge of dark net drug markets," *Policy brief*, vol. 7, pp. 1–24, Jan. 2015.
[28] M. Caiani and L. Parenti, "The dark side of the Web: Italian right-wing extremist groups and the Internet," *South Eur. Soc. Politics*, vol. 14, no. 3, pp. 273–294, Sep. 2009.
[29] O. Catakoglu, M. Balduzzi, and D. Balzarotti, "Attacks landscape in the dark side of the Web," presented at the Symp. Appl. Comput. (SAC), 2017.
[30] A. Celestini, G. Me, and M. Mignone, "Tor marketplaces exploratory data analysis: The drugs case," presented at the Int. Conf. Global Secur., Saf., Sustainability, 2017.
[31] D. Chang, M. Ghosh, S. K. Sanadhya, M. Singh, and D. R. White, "FbHash: A new similarity hashing scheme for digital forensics," *Digit. Invest.*, vol. 29, pp. S113–S123, Jul. 2019.
[32] R. R. Chaudhari and S. P. Patil, "Intrusion detection system: Classification, techniques and datasets to implement," *Int. Res. J. Eng. Technol.* vol. 4, no. 2, 2017, Art. no. 186066.
[33] H. Chen, *Dark Web: Exploring and Data Mining the Dark Side of the Web*, vol. 30. Cham, Switzerland: Springer, 2011.
[34] H. Chen, W. Chung, J. Qin, E. Reid, M. Sageman, and G. Weimann, "Uncovering the dark Web: A case study of jihad on the Web," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 59, no. 8, pp. 1347–1359, Jun. 2008.
[35] M. Chen and A. A. Ghorbani, "A survey on user profiling model for anomaly detection in cyberspace," *J. Cyber Secur. Mobility*, vol. 8, no. 1, pp. 75–112, 2019.
[36] O. Cherqi, G. Mezzour, M. Ghogho, and M. El Koutbi, "Analysis of hacking related trade in the darkweb," presented at the IEEE Int. Conf. Intell. Secur. Informat. (ISI), Nov. 2018.
[37] M. Chertoff and T. Simon, "The impact of the dark Web on Internet governance and cyber security," Centre Int. Governance Innovation (CIGI), Waterloo, ON, Canada, Tech. Rep. 6, 2015.
[38] N. Christin, "Traveling the silk road: A measurement analysis of a large anonymous online marketplace," presented at the 22nd Int. Conf. World Wide Web, 2013.
[39] V. Ciancaglini, M. Balduzzi, M. Goncharov, and R. McArdle, "Deepweb and cybercrime," Trend Micro, Tokyo, Japan, Trend Micro Rep. 9, 2013.
[40] H. J. Clawson and N. Dutch, "Addressing the needs of victims of human trafficking: Challenges, barriers, and promising practices: Department of health and human services, office of the assistant secretary," Dept. Health Hum. Services, Washington, DC, USA, Tech. Rep., 2008.
[41] ConvenantEyes. (Sep. 7, 2011). *The Connections Between Pornography and Sex Trafficking*. [Online]. Available: https://www.covenanteyes.com/2011/09/07/the-connections-between-pornography-and-sex-trafficking/
[42] C. Cranford. (Feb. 21, 2015). *Dangerous Apps on Your Teen's Mobile Device*. [Online]. Available: https://www.cybersafetycop.com/dangerous-apps-on-your-teens-mobile-device/
[43] D. S. Cruzes and T. Dyba, "Recommended steps for thematic synthesis in software engineering," presented at the Int. Symp. Empirical Softw. Eng. Meas., Sep. 2011.
[44] M.-F. Cuellar, "The tenuous relationship between the fight against money laundering and the disruption of criminal finance," *J. Crim. L. Criminol.*, vol. 93, no. 2, p. 311, 2002.
[45] Daily Mail. (Oct. 12, 2013). *The Disturbing World of the Deep Web, Where Contract Killers and Drug Dealers Ply Their Trade on the Internet*. [Online]. Available: https://www.dailymail.co.uk/news/article-2454735/The-disturbing-world-Deep-Web-contract-killers-drug-dealers-ply-trade-internet.html
[46] I. B. Damgård, "A design principle for hash functions," presented at the Conf. Theory Appl. Cryptol., 1989.
[47] Memex(Archived), DARPA. (2019). *Defense Advanced Research Projects Agency*. [Online]. Available: https://www.darpa.mil/program/memex
[48] J. DeBlasio, S. Savage, G. M. Voelker, and A. C. Snoeren, "Tripwire: Inferring Internet site compromise," presented at the Internet Meas. Conf., Nov. 2017.
[49] D. Décary-Hétu and L. Giommoni, "Do police crackdowns disrupt drug cryptomarkets? A longitudinal analysis of the effects of operation onymous," *Crime, Law Social Change*, vol. 67, no. 1, pp. 55–75, Feb. 2017.
[50] DeepDotWeb. (2015). *Interview: 'Mr. Nice Guy' Market Admin Tells His Story*. [Online]. Available: https://gir.pub/deepdotweb/2015/06/03/interview-with-mr-niceguy-market-admin/
[51] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," Naval Res. Lab., Washington, DC, USA, Tech. Rep., 2004.
[52] DNStats. (2019). *Dark Net Stats*. [Online]. Available: https://dnstats.net/
[53] D. S. Dolliver and J. L. Kenney, "Characteristics of drug vendors on the tor network: A cryptomarket comparison," *Victims Offenders*, vol. 11, no. 4, pp. 600–620, Oct. 2016.

[54] S. Dredge. (Nov. 5, 2013). *What is Tor? A Beginner's Guide to the Privacy Tool*. The Guardian, London, U.K. [Online]. Available: https://www.theguardian.com/technology/2013/nov/05/tor-beginners-guide-nsa-browser

[55] P. Dughi. (Jun. 26, 2016). 17 times social media helped police track down thieves, murderers, and gang criminals. Medium. [Online]. Available: https://medium.com/the-mission/17-times-social-media-helped-police-track-down-thieves-murderers-and-gang-criminals-a814b6c40fb

[56] R. Ehney and J. D. Shorter, "Deep Web, dark Web, invisible Web and the post isis world," *Inf. Syst.*, vol. 17, no. 4, pp. 36–41, 2016.

[57] K. P. Erb. (Oct. 16, 2019). IRS followed bitcoin transactions, resulting in takedown of the largest child exploitation site on the Web. Forbes. [Online]. Available: https://www.forbes.com/sites/kellyphillipserb/2019/10/16/irs-followed-bitcoin-transactions-resulting-in-takedown-of-the-largest-child-exploitation-site-on-the-web/#327343231ed0

[58] N. S. Evans, R. Dingledine, and C. Grothoff, "A practical congestion attack on tor using long paths," presented at the USENIX Secur. Symp., 2009.

[59] W. Fan, Z. Du, D. Fernandez, and V. A. Villagra, "Enabling an anatomic view to investigate honeypot systems: A survey," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3906–3919, Dec. 2018.

[60] Y. Fanusie and T. Robinson, "Bitcoin laundering: An analysis of illicit flows into digital currency services," Center Sanctions Illicit Finance Memorandum, Elliptic, London, U.K., Tech. Rep., Jan. 2018.

[61] FBI. (2019). *Oversight of the Federal Bureau of Investigation*. [Online]. Available: https://www.fbi.gov/news/testimony/oversight-of-the-federal-bureau-of-investigation-072319

[62] FBI. (May 5, 2017). *Playpen' Creator Sentenced to 30 Years*. [Online]. Available: https://www.fbi.gov/news/stories/playpen-creator-sentenced-to-30-years

[63] J. Fernquist, L. Kaati, and R. Schroeder, "Political bots and the Swedish general election," presented at the IEEE Int. Conf. Intell. Secur. Informat. (ISI), Nov. 2018.

[64] K. M. Finklea, "Dark Web," in *Proc. Congressional Res. Service*, 2015, pp. 1–16.

[65] K. Gai, M. Qiu, L. Tao, and Y. Zhu, "Intrusion detection techniques for mobile cloud computing in heterogeneous 5G," *Secur. Commun. Netw.*, vol. 9, no. 16, pp. 3049–3058, Nov. 2016.

[66] A. García-Holgado and F. J. García-Peñalvo, "Mapping the systematic literature studies about software ecosystems," presented at the 6th Int. Conf. Technol. Ecosystems Enhancing Multiculturality (TEEM), 2018.

[67] Geneva. (Sep. 19, 2017). *Forced Labour, Modern Slavery and Human Trafficking*. [Online]. Available: https://www.ilo.org/global/topics/forced-labour/lang–en/index.htm

[68] A. Ghappour, "Searching places unknown: Law enforcement jurisdiction on the dark Web," *Stan. L. Rev.*, vol. 69, no. 4, p. 1075, 2017.

[69] D. Gilbert, (Mar. 19, 2018). Criminals are racing to cash out their bitcoin. Here's how they're doing it. Vice News. [Online]. Available: https://www.vice.com/en_ca/article/7xdzqa/criminals-are-racing-to-cash-out-their-bitcoin-heres-how-theyre-doing-it

[70] R. Gowsalya and S. Amali, "Naive Bayes based network traffic classification using correlation information," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 4, no. 3, 2014.

[71] H. Grant, "Cathryn Lavery of Iona College' Social Media and the New Generation of 'Computerated' Criminals," *Crim Forensic Studies*, vol. 2, no. 1, 2019, Art. no. 180022.

[72] A. Greenberg. (May 5, 2014). The FBI finally says how it 'legally' pinpointed silk road's server. WIRED. [Online]. Available: https://www.wired.com/2014/09/the-fbi-finally-says-how-it-legally-pinpointed-silk-roads-server/

[73] L. Greenemeier. (Feb. 8, 2015). Human traffickers caught on hidden Internet. Scientific American. [Online]. Available: https://www.scientificamerican.com/article/human-traffickers-caught-on-hidden-internet/

[74] R. Grinberg, "Bitcoin: An innovative alternative digital currency," *Hastings Sci. Tech. LJ*, vol. 4, no. 1, p. 159, 2012.

[75] J. Hammonds, "An inquiry into privacy concerns: Memex, the deep Web, and sex trafficking," Infosec Writers, Tech. Rep., 2015.

[76] D. Hayes, F. Cappa, and J. Cardon, "A framework for more effective dark Web marketplace investigations," *Information*, vol. 9, no. 8, p. 186, Jul. 2018.

[77] T. J. Holt and E. Lampke, "Exploring stolen data markets online: Products and market forces," *Criminal Justice Stud.*, vol. 23, no. 1, pp. 33–50, Mar. 2010.

[78] T. J. Holt, O. Smirnova, and Y. T. Chua, "Exploring and estimating the revenues and profits of participants in stolen data markets," *Deviant Behav.*, vol. 37, no. 4, pp. 353–367, Apr. 2016.

[79] Homeland Security. (2019). *Human Trafficking and the Hospitality Industry*. [Online]. Available: https://www.dhs.gov/blue-campaign/hospitalityindustry

[80] D. Ibata. (Apr. 18, 2012). Milton man arrested in international online drug bust. The Atlanta Journal-Constitution. [Online]. Available: https://www.ajc.com/news/local/milton-man-arrested-international-online-drug-bust/gYjBHvrZKens0q2kVxf3ZK/

[81] E. Jardine, *The Dark Web Dilemma: Tor, Anonymity and Online Policing* (Global Commission on Internet Governance Paper Series). Waterloo, ON, Canada: The Centre for International Governance Innovation (CIGI), 2015.

[82] I. Jayaweera, C. Sajeewa, S. Liyanage, T. Wijewardane, I. Perera, and A. Wijayasiri, "Crime analytics: Analysis of crimes through newspaper articles," presented at the Moratuwa Eng. Res. Conf. (MERCon), Apr. 2015.

[83] F. Johansson, L. Kaati, and A. Shrestha, "Detecting multiple aliases in social media," presented at the IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), 2013.

[84] D. Jovanović and P. Janićić, "Logical analysis of hash functions," presented at the Int. Workshop Frontiers of Combining Syst., 2005.

[85] G. Kaur and B. Nagpal, "Malware analysis & its application to digital forensic," *Int. J. Comput. Sci. Eng.* vol. 4, no. 4, pp. 622–626, 2012.

[86] A. C. Kim, S. Kim, W. H. Park, and D. H. Lee, "Fraud and financial crime detection model using malware forensics," *Multimedia Tools Appl.*, vol. 68, no. 2, pp. 479–496, Jan. 2014.

[87] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele Univ., Durham Univ. Joint Rep., Keele, U.K., Tech. Rep. EBSE Technical Report EBSE-2007-01, 2007.

[88] R. Koch, "Hidden in the shadow: The dark Web—A growing risk for military operations?" presented at the 11th Int. Conf. Cyber Conflict (CyCon), 2019.

[89] I. Koniaris, G. Papadimitriou, P. Nicopolitidis, and M. Obaidat, "Honeypots deployment for the analysis and visualization of malware activity and malicious connections," presented at the IEEE Int. Conf. Commun. (ICC), Jun. 2014.

[90] R. Konrad and A. Trapp. (2017). *Data Science Can Help Us Fight Human Trafficking*. [Online]. Available: https://theconversation.com/data-science-can-help-us-fight-human-trafficking-81647

[91] J. Kornblum, "Identifying almost identical files using context triggered piecewise hashing," *Digit. Invest.*, vol. 3, pp. 91–97, Sep. 2006.

[92] A. Kulkarni, J. Goldman, B. Nabholz, and W. Eyre, "Detection of steganography-producing software artifacts on crime-related seized computers," *J. Digit. Forensics, Secur. Law.* vol. 4, no. 2, p. 1, 2009.

[93] S. Kumar, J. Cheng, J. Leskovec, and V. Subrahmanian, "An army of me: Sockpuppets in online discussion communities," presented at the 26th Int. Conf. World Wide Web, 2017.

[94] G. L'Huillier, H. Alvarez, S. A. Ríos, and F. Aguilera, "Topic-based social network analysis for virtual communities of interests in the dark Web," *ACM SIGKDD Explor. Newslett.*, vol. 12, no. 2, pp. 66–73, 2011.

[95] J. Lane, "Bitcoin, silk road, and the need for a new approach to virtual currency regulation," *Charleston L. Rev.*, vol. 8, no. 5, p. 511, 2013.

[96] LexisNexis. (2019). *Law Enforcement's Usage of Social Media for Investigations Infographic Risk Solutions*. [Online]. Available: https://risk.lexisnexis.com/insights-resources/infographic/law-enforcement-usage-of-social-media-for-investigations-infographic

[97] W. Li and H. Chen, "Identifying top sellers in underground economy using deep learning-based sentiment analysis," presented at the IEEE Joint Intell. Secur. Informat. Conf., Sep. 2014, pp. 64–67.

[98] S. Lightfoot and F. Pospisil, "Surveillance and privacy on the deep Web," ResearchGate, Berlin, Germany, Tech. Rep., 2017.

[99] N. Lindsey (Mar. 12, 2019). Cyber criminals have turned social media cyber crime into a $3 billion business. CPO Magazine. [Online]. Available: https://www.cpomagazine.com/cyber-security/cyber-criminals-have-turned-social-media-cyber-crime-into-a-3-billion-business/

[100] Z. Ling, J. Luo, W. Yu, X. Fu, D. Xuan, and W. Jia, "A new cell-counting-based attack against tor," *IEEE/ACM Trans. Netw.*, vol. 20, no. 4, pp. 1245–1261, Aug. 2012.
[101] D. Liu, Q. Wu, W. Han, and B. Zhou, "Sockpuppet gang detection on social media sites," *Frontiers Comput. Sci.*, vol. 10, no. 1, pp. 124–135, Feb. 2016.
[102] Q. Liu, R. Klucik, C. Chen, G. Grant, D. Gallaher, Q. Lv, and L. Shang, "Unsupervised detection of contextual anomaly in remotely sensed data," *Remote Sens. Environ.*, vol. 202, pp. 75–87, Dec. 2017.
[103] A. Maddox, M. J. Barratt, M. Allen, and S. Lenton, "Constructive activism in the dark Web: Cryptomarkets and illicit drugs in the digital 'demimonde,'" *Inf., Commun. Soc.*, vol. 19, no. 1, pp. 111–126, Jan. 2016.
[104] E. Mainas, "The analysis of criminal and terrorist organisations as social network structures: A quasi-experimental study," *Int. J. Police Sci. Manage.*, vol. 14, no. 3, pp. 264–282, 2012.
[105] S. Mancini and L. A. Tomei, "The dark Web: Defined, discovered, exploited," *Int. J. Cyber Res. Edu.*, vol. 1, no. 1, pp. 1–12, Jan. 2019.
[106] M. C. Van Hout, *Drugs on the Dark Net: How Cryptomarkets are Transforming the Global Trade in Illicit Drugs*. Cham, Switzerland: Springer, 2014.
[107] C. A. Mattmann, "Search of the deep and dark Web via darpa memex," presented at the AGU Fall Meeting Abstr., 2015.
[108] P. Mishra, E. S. Pilli, V. Varadharajan, and U. Tupakula, "Intrusion detection techniques in cloud environment: A survey," *J. Netw. Comput. Appl.*, vol. 77, pp. 18–47, Jan. 2017.
[109] D. Moher, P.-P. Group, L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement," *Systematic Rev.*, vol. 4, no. 1, p. 1, Dec. 2015.
[110] A. Montieri, D. Ciuonzo, G. Aceto, and A. Pescapé, "Anonymity services Tor, I2P, JonDonym: Classifying in the dark," presented at the 29th Int. Teletraffic Congr. (ITC), 2017.
[111] C. Moore, "Detecting ransomware with honeypot techniques," presented at the Cybersecurity Cyberforensics Conf. (CCC), Aug. 2016.
[112] D. Moore and T. Rid, "Cryptopolitik and the Darknet," *Survival*, vol. 58, no. 1, pp. 7–38, 2016.
[113] F. Mosteller and D. L. Wallace, *Applied Bayesian and Classical Inference The Case of The Federalist Papers*. Cham, Switzerland: Springer, 2012.
[114] K. Nakao, D. Inoue, M. Eto, and K. Yoshioka, "Practical correlation analysis between scan and malware profiles against zero-day attacks based on darknet monitoring," *IEICE Trans. Inf. Syst.*, vols. E92–D, no. 5, pp. 787–798, 2009.
[115] L. H. Newman. (Oct. 16, 2019). How a bitcoin trail led to a massive dark Web child-porn site takedown. Wired. [Online]. Available: https://www.wired.com/story/dark-web-welcome-to-video-takedown-bitcoin/
[116] H. Nishikaze, S. Ozawa, J. Kitazono, T. Ban, J. Nakazato, and J. Shimamura, "Large-scale monitoring for cyber attacks by using cluster information on darknet traffic features," *Procedia Comput. Sci.*, vol. 53, pp. 175–182, Jan. 2015.
[117] E. Nunes, A. Diab, A. Gunn, E. Marin, V. Mishra, V. Paliath, J. Robertson, J. Shakarian, A. Thart, and P. Shakarian, "Darknet and deepnet mining for proactive cybersecurity threat intelligence," presented at the IEEE Conf. Intell. Secur. Informat. (ISI), Sep. 2016.
[118] A. Nursetyo, D. R. Ignatius Moses Setiadi, E. H. Rachmawanto, and C. A. Sari, "Website and network security techniques against brute force attacks using honeypot," presented at the 4th Int. Conf. Informat. Comput. (ICIC), Oct. 2019.
[119] A. Panchenko, A. Mitseva, M. Henze, F. Lanze, K. Wehrle, and T. Engel, "Analysis of fingerprinting techniques for tor hidden services," presented at the Workshop Privacy Electron. Soc. (WPES), 2017.
[120] S. Pfeifer, S. Li, and W. Hamilton. (Oct. 2, 2013). *End of Silk Road for Drug Users as FBI Shuts Down Illicit Website*. [Online]. Available: https://www.latimes.com/business/la-fi-silk-road-bitcoin-20131003-story.html
[121] K. Poulsen. (Sep. 9, 2013). *FBI Admits it Controlled Tor Servers Behind Mass Malware Attack*. [Online]. Available: https://www.wired.com/2013/09/freedom-hosting-fbi/
[122] H. Prunckun, *Scientific Methods of Inquiry for Intelligence Analysis: Rowman and Littlefield*. Lanham, MD, USA: Rowman & Littlefield, 2014.
[123] E. Puffer, K. McDonald, M. Pross, and D. Hudson, "Webcam child sex tourism: An emerging global issue," Cedarville Univ., Cedarville, OH, USA, Tech. Rep., 2014.
[124] D. Rathod, "Darknet forensics," *Future*, vol. 11, no. 4, p. 12, 2017.
[125] C. Reilly. (Jul. 29, 2015). *Human Trafficking: A Crime Hard to Track Proves Harder to Fight*. [Online]. Available: https://www.pbs.org/wgbh/frontline/article/what-is-human-trafficking-and-why-is-it-so-hard-to-combat/
[126] D. Rhumorbarbe, L. Staehli, J. Broséus, Q. Rossy, and P. Esseiva, "Buying drugs on a darknet market: A better deal? Studying the online illicit drug market through the analysis of digital, physical and chemical data," *Forensic Sci. Int.*, vol. 267, pp. 173–182, Oct. 2016.
[127] S. A. Ríos and R. Muñoz, "Dark Web portal overlapping community detection based on topic models," presented at the ACM SIGKDD Workshop Intell. Secur. Informat. (ISI-KDD), 2012.
[128] J. Robertson, A. Diab, E. Marin, E. Nunes, V. Paliath, J. Shakarian, and P. Shakarian, "Darknet mining and game theory for enhanced cyber threat intelligence," *Cyber Defense Rev.*, vol. 1, no. 2, pp. 95–122, 2016.
[129] T. Roufa. (Oct. 16, 2019). How law enforcement employs social media monitoring tools. The Balance Careers. [Online]. Available: https://www.thebalancecareers.com/social-networking-and-law-enforcement-974548
[130] V. Roussev, "Data fingerprinting with similarity digests," presented at the IFIP Int. Conf. Digit. Forensics, 2010.
[131] G. K. Sadasivam, C. Hota, and B. Anand "Detection of severe SSH attacks using honeypot servers and machine learning techniques," *Softw. Netw.*, vol. 2017, no. 1, pp. 79–100, 2017.
[132] S. Sarkar, M. Almukaynizi, J. Shakarian, and P. Shakarian, "Mining user interaction patterns in the darkweb to predict enterprise cyber incidents," *Social Netw. Anal. Mining*, vol. 9, no. 1, Dec. 2019.
[133] J. R. Scanlon and M. S. Gerber, "Automatic detection of cyber-recruitment by violent extremists," *Secur. Informat.*, vol. 3, no. 1, p. 5, Dec. 2014.
[134] J. Shakarian, A. T. Gunn, and P. Shakarian, "Exploring malicious hacker forums," in *Cyber Deception*. Cham, Switzerland: Springer, 2016, pp. 259–282.
[135] A. Shimoda, T. Mori, and S. Goto, "Extended darknet: Multi-dimensional Internet threat monitoring system," *IEICE Trans. Commun.*, vol. E95.B, no. 6, pp. 1915–1923, 2012.
[136] D. Shinder, "What makes cybercrime laws so difficult to enforce," Tech. Rep., 2011.
[137] P. W. Singer and E. T. Brooking, *LikeWar: The Weaponization of Social Media*. New York, NY, USA: Eamon Dolan Books, 2018.
[138] T. Solorio, R. Hasan, and M. Mizan, "A case study of sockpuppet detection in Wikipedia," presented at the Workshop Lang. Anal. Social Media, 2013.
[139] K. Soska and N. Christin, "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem," presented at the 24th USENIX Secur. Symp., 2015.
[140] M. Spitters, F. Klaver, G. Koot, and M. van Staalduinen, "Authorship analysis on dark marketplace forums," presented at the Eur. Intell. Secur. Informat. Conf., Sep. 2015.
[141] D. Stupples, "ICITST-2013: Keynote speaker 2: Security challenge of TOR and the deep Web," presented at the 8th Int. Conf. Internet Technol. Secured Trans. (ICITST), Dec. 2013.
[142] R. Surette, "Performance crime and justice," *Current Issues Criminal Justice*, vol. 27, no. 2, pp. 195–216, 2015.
[143] D. J. B. Svantesson. (Sep. 6, 2017). It's too hard to get the data of Australian criminals when it's stored overseas. The Conversation. [Online]. Available: http://theconversation.com/its-too-hard-to-get-the-data-of-australian-criminals-when-its-stored-overseas-82828
[144] The Bitcoin News. (Feb. 9, 2017). *Anonymous Hacks Freedom Hosting II, Bringing Down Almost 20% of Active Darknet Sites*. [Online]. Available: https://thebitcoinnews.com/anonymous-hacks-freedom-hosting-ii-bringing-down-almost-20-of-active-darknet-sites/
[145] V. L. Thing, M. Sloman, and N. Dulay, "Locating network domain entry and exit point/path for DDoS attack traffic," *IEEE Trans. Netw. Service Manage.*, vol. 6, no. 3, pp. 163–174, Sep. 2009.
[146] P. V. Torres-Carrion, C. S. Gonzalez-Gonzalez, S. Aciar, and G. Rodriguez-Morales, "Methodology for systematic literature review applied to engineering and education," presented at the IEEE Global Eng. Educ. Conf. (EDUCON), Apr. 2018.
[147] A. Travis. (Oct. 16, 2015). *Crime rate in England and Wales Soars as Cybercrime is Included for First Time The Guardian*. [Online]. Available: https://www.theguardian.com/uk-news/2015/oct/15/rate-in-england-and-wales-soars-as-cybercrime-included-for-first-time

[148] M. Tsikerdekis and S. Zeadally, "Multiple account identity deception detection in social media using nonverbal behavior," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 8, pp. 1311–1321, Aug. 2014.
[149] P. Tucker. (Feb. 24, 2015). *How the Military Will Fight ISIS on the Dark Web*. [Online]. Available: https://www.defenseone.com/technology/2015/02/how-military-will-fight-isis-dark-web/105948/
[150] M. Tzanetakis, "Comparing cryptomarkets for drugs. A characterisation of sellers and buyers over time," *Int. J. Drug Policy*, vol. 56, pp. 176–186, Jun. 2018.
[151] M. F. Umer, M. Sher, and Y. Bi, "Flow-based intrusion detection: Techniques and challenges," *Comput. Secur.*, vol. 70, pp. 238–254, Sep. 2017.
[152] J. Van Buskirk, S. Naicker, R. Bruno, C. Breen, and A. Roxburgh, "Drugs and the Internet," Nat. Drug Alcohol Res. Centre (NDARC), UNSW, Sydney, NSW, Australia, Tech. Rep. 7, 2016.
[153] Vienna. (Jul. 30, 2018). *Cybercrime—What is the Hidden Nature of Digital Criminal Activities Nowadays?* [Online]. Available: https://www.titanium-project.eu/news/articles/cybercrime/
[154] S.-J. Wang, "Measures of retaining digital evidence to prosecute computer-based cyber-crimes," *Comput. Standards Interfaces*, vol. 29, no. 2, pp. 216–223, Feb. 2007.
[155] Z. Wang and L. Cao, "Implementation and comparison of two hash algorithms," presented at the Int. Conf. Comput. Inf. Sci., Jun. 2013.
[156] I. M. A. Warren and M. Mann. (Sep. 7, 2017). Poisoned water holes: The legal dangers of dark Web policing. News Pty Ltd. [Online]. Available: https://www.news.com.au/technology/online/poisoned-water-holes-the-legal-dangers-of-dark-web-policing/news-story/285655e36981515e35e2290360f9e646
[157] G. Weimann, "Going dark: Terrorism on the dark Web," *Stud. Conflict Terrorism*, vol. 39, no. 3, pp. 195–206, Mar. 2016.
[158] Y. Wu, F. Zhao, X. Chen, P. Skums, E. L. Sevigny, D. Maimon, and M. J. Feizollahi, "Python scrapers for scraping cryptomarkets on tor," presented at the Int. Conf. Secur., Privacy Anonymity Comput., Commun. Storage, 2019.
[159] P. Xiao, W. Qu, H. Qi, and Z. Li, "Detecting DDoS attacks against data center with correlation analysis," *Comput. Commun.*, vol. 67, pp. 66–74, Aug. 2015.
[160] J. Xu, H. Chen, Y. Zhou, and J. Qin, "On the topology of the dark Web of terrorist groups," presented at the Int. Conf. Intell. Secur. Inform., 2006.
[161] L. Yang, F. Liu, J. M. Kizza, and R. K. Ege, "Discovering topics from dark websites," presented at the IEEE Symp. Comput. Intell. Cyber Secur., Mar. 2009.
[162] H. Zenati, M. Romain, C.-S. Foo, B. Lecouat, and V. Chandrasekhar, "Adversarially learned anomaly detection," presented at the IEEE Int. Conf. Data Mining (ICDM), Nov. 2018.
[163] H. Zhang, M. A. Babar, and P. Tell, "Identifying relevant studies in software engineering," *Inf. Softw. Technol.*, vol. 53, no. 6, pp. 625–637, Jun. 2011.
[164] J. Zhang, Y. Xiang, Y. Wang, W. Zhou, Y. Xiang, and Y. Guan, "Network traffic classification using correlation information," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 104–117, Jan. 2013.
[165] Y. Zhang, S. Zeng, L. Fan, Y. Dang, C. A. Larson, and H. Chen, "Dark Web forums portal: Searching and analyzing Jihadist forums," presented at the IEEE Int. Conf. Intell. Secur. Informat., 2009.
[166] R. Zheng, J. Li, H. Chen, and Z. Huang, "A framework for authorship identification of online messages: Writing-style features and classification techniques," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 57, no. 3, pp. 378–393, Feb. 2006.
[167] X. Zheng, Y. M. Lai, K. Chow, L. C. Hui, and S. Yiu, "Detection of sockpuppets in online discussion forums," Univ. Hong Kong, Hong Kong, HKU CS Tech. Rep. TR-2011-03, 2011.
[168] Y. Zhou, E. Reid, J. Qin, H. Chen, and G. Lai, "US domestic extremist groups on the Web: Link and content analysis," *IEEE Intell. Syst.*, vol. 20, no. 5, pp. 44–51, Sep. 2005.
[169] Y. Zhu, X. Fu, B. Graham, R. Bettati, and W. Zhao, "Correlation-based traffic analysis attacks on anonymity networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 21, no. 7, pp. 954–967, Jul. 2010.
[170] A. T. Zulkarnine, R. Frank, B. Monk, J. Mitchell, and G. Davies, "Surfacing collaborated networks in dark Web to find illicit and criminal content," presented at the IEEE Conf. Intell. Secur. Informat. (ISI), Sep. 2016.

**SAIBA NAZAH** received the B.Sc. degree in computer science and engineering from the Chittagong University of Engineering and Technology, Bangladesh. She is currently pursuing the master's degree in cyber security with the School of Information Technology, Deakin University, Australia, under the supervision of Prof. Dr. Jemal. H. Abawajy. Her main research interests include cyber security, dark net, natural language processing, machine learning, the Internet of Things, and data analytics.

**SHAMSUL HUDA** received the Ph.D. degree in computer science, in 2010. He worked as an Academic with Federation University. He worked as an Assistant Professor with the Khulna University of Engineering and Technology (KUET), Bangladesh. He is currently a Lecturer with the School of Information Technology, Deakin University, Australia. He is also a member of the Cyber Security Research and Innovation Centre (CSRI), Deakin University. He is a Certified Information System Security Professional (CISSP) by The International Information System Security Certification Consortium, $(ISC)^2$. He is involved in many international cyber security projects, including cybersecurity capacity maturity for nations with the Oceania Cyber Security Centre (OCSC), Melbourne, with partnership of the Global Cyber Security Capacity Centre (GCSCC) with the University of Oxford. He has published more than 60 journal and conference papers in well reputed journals, including the IEEE TRANSACTIONS. His main research interests include communication and network security, strategies for secure operations for industrial control systems (SCADA) and critical infrastructure, intelligent counter measure for threats against mobile systems, detection of data breaches through the darknet, the IoT security, malware analysis and detection, reverse engineering for endpoint security, and malware analysis and detection for SCADA systems.

**JEMAL ABAWAJY** (Senior Member, IEEE) is currently a Full Professor with the Faculty of Science, Engineering and Built Environment, Deakin University, Australia. He has authored/coauthored over 250 refereed articles and supervised numerous Ph.D. students to completion. He has delivered over 50 keynote and seminars worldwide and has been involved in the organization of over international conferences in various capacity, including chair and general co-chair. He has also served on the editorial board of numerous international journals, including the IEEE TRANSACTIONS ON CLOUD COMPUTING.

**MOHAMMAD MEHEDI HASSAN** (Senior Member, IEEE) received the Ph.D. degree in computer engineering from Kyung Hee University, Seoul, South Korea, in February 2011. He is currently an Associate Professor with the Information Systems Department, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. He has authored or coauthored around 180+ publications, including refereed IEEE/ACM/Springer/Elsevier journals, conference papers, books, and book chapters. His research interests include edge/cloud computing, the Internet of Things, cyber security, deep learning, artificial intelligence, body sensor networks, 5G networks, and social networks. He was a recipient of a number of awards, including the Best Journal Paper Award from the IEEE SYSTEMS JOURNAL in 2018, the Best Paper Award from CloudComp conference in 2014, and the Excellence in Research Award from King Saud University (two times in a row 2015 and 2016).