IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 5, 2020, accepted October 16, 2020, date of publication October 29, 2020, date of current version November 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3034816

Cryptocurrencies Emerging Threats and Defensive Mechanisms: A Systematic Literature Review

EMAD BADAWI AND GUY-VINCENT JOURDAN
Faculty of Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada

Corresponding author: Emad Badawi (ebada090@uottawa.ca)

This work was supported in part by the IBM Center for Advanced Studies and by the Natural Sciences and Engineering Research Council of Canada.

ABSTRACT Cryptocurrencies have been a target for cybercriminal activities because of the pseudo-anonymity and privacy they offer. Researchers have been actively working on analyzing and developing innovative defensive mechanisms to prevent these activities. A significant challenge facing researchers is collecting datasets to train defensive systems to detect and analyze these cyberattacks. Our aims in this systematic review are to explore and aggregate the state of the art threats that have emerged with cryptocurrencies and the defensive mechanisms that have been proposed. We also discuss the threats type, scale, and how efficient the defensive mechanisms are in providing early detection and prevention. We also list out the resources that have been used to collect datasets, and we identify the publicly available ones. In this study, we extracted 1,221 articles from four top scientific and engineering databases and libraries in Computer Science: IEEE Xplore, ACM Digital Library, Elsevier's Scopus, and Crarivate's Web of Science. We defined inclusion, exclusion, and quality of assessment criteria, and after a detailed review process, 66 publications were included in the final review. Our analysis revealed that the literature contains a significant amount of research to detect and analyze several attack types, such as the high yield investment programs and pump and dump. These attacks have been used to steal millions of USD, abuse millions of connected devices, and have created even more significant loss in denial of services and productivity losses. We have found that the researchers use various sources to collect training datasets. Many authors have made their dataset publicly available. We have created a list of these datasets, which we have made available along with other supplementary websites, tools, and libraries that can be used in the data collection and analysis process.

INDEX TERMS Blockchain, cryptocurrency, cryptojacking, cyberattack, fraud, HYIP, money laundering, pumb and dumb, ransomware, scam.

I. INTRODUCTION

In recent years, there has been a rise in the use of cryptocurrencies as an investment platform [1]. As of September 27th, 2020, there are 7,186 different cryptocurrencies, with a capitalization market of approximately 346 billion USD.[1] The most popular cryptocurrencies are Bitcoin and Ethereum, which have a capitalization market of approximately 199 billion USD and 40 billion USD, respectively.

Bitcoin is a decentralized cryptocurrency that has become popular in the last ten years. It is a peer-to-peer electronic currency that can be sent from one user to another without the involvement of a trusted authority such as an administrator or a central bank [2]–[4]. It first appeared in a white paper by "Satoshi Nakamoto" [2]. The actual identity of Nakamoto is still unclear. Unlike traditional currencies, bitcoin has two key features: Transparency and Pseudo-anonymity [2], [4], [5]. It is transparent because the transactions are publicly announced in a decentralized ledger called a blockchain. The Pseudo-anonymity comes from the fact that the users use pseudonyms (addresses). These pseudonyms are not related to individuals; they are computed from the user's public key [2]. Moreover, bitcoin addresses can be generated at will [4]. As a result, users can create a unique address for each transaction. This increases privacy by creating an additional layer to keep the addresses from being linked to a specific owner [2].

Cybercriminals have leveraged Bitcoin Pseudo-anonymity in their attacks. According to a report by CipherTrace,[2] the value of thefts, hacks, and scams has more than doubled

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

[1]https://coinmarketcap.com/

[2]https://ciphertrace.com/q4-2019-cryptocurrency-anti-money-laundering-report/

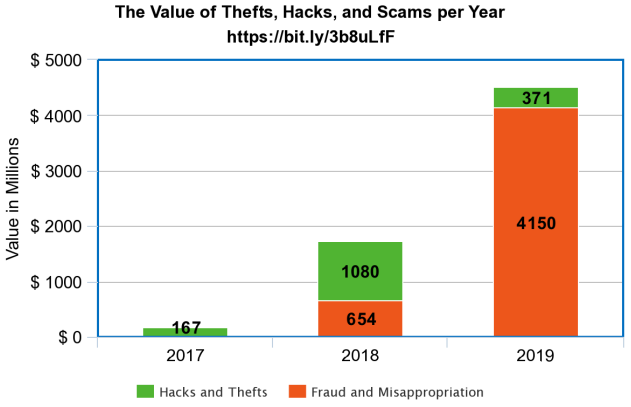

FIGURE 1. Q4 2019 cryptocurrency anti-money laundering report (reproduced from CipherTrace report[2]).

in 2019 when compared to 2018 and was more than 230 times the value of 2017; more than 4.52 billion USD was siphoned away from users and cryptocurrency exchanges in 2019 only. As shown in Figure 1, losses from cryptocurrencies exchange hacks and thefts reached 371 million USD, while the majority of the losses (4.15 billion USD) was due to fraud and misappropriation of funds. For example, the South Korean cryptocurrency exchange "Upbit" was one of the targets of the hacks and theft attacks in 2019; on November 27, Upbit CEO announced that the exchange was hacked and 342,000 Ethereum Worth 52 million USD were stolen in the attack. On the other hand, the "BitClub Network" defrauded investors of 722 million USD in a Ponzi scheme before four of it is operators got arrested.[2]

Cybercriminal attacks using Cryptocurrencies take many forms. "High yield investment programs" (HYIP) is one of the popular examples of the scams that cybercriminals carry out [3], [4], [6], [7]. HYIP is a scam in which investors are promised a high-interest rate, e.g., more than 1-2% per day [4]. Perhaps the most famous HYIP scammer was Charles Ponzi, who claimed in the early 1920s to run an arbitrage; the investors were promised a 50% profit within 45 days, or 100% profit within 90 days. Because of Charles Ponzi, HYIP is sometimes called *Ponzi scheme* [4].

Money laundering (ML) [8], [9], ransomware [10]–[12], and pump and dump (P&D) [1], [13]–[15] are other popular examples. ML describes the process of disguising the sources of illegal profits generated by criminal activity. It aims to hide the link between original criminal activities and the corresponding funds by passing the money through a complex sequence of commercial transactions or banking transfers [8], [9].

Ransomware is a denial-of-access attack in which a malicious piece of software locks and encrypts a victim's device data until a sum of money is paid [12]. Cryptocurrencies, usually Bitcoin, are often used for these payments. Recently, Riviera Beach officials voted to pay 65 bitcoins, worth 600,000 USD at the time, to a cybercriminal who seized and shut down the city's computer systems. The resulting outage forced the local fire and police departments to write down hundreds of 911 calls on paper.[3]

P&D scheme is a type of fraud in which the fraudster aims to make a profit from stock trading by artificially manipulating stock prices. In P&D, the attackers purchase stocks at a low price (pump) then spread misleading recommendations and positive statements to convince other investors to buy that stock, which increases its price. The attacker then sells (dump) their stock at a mark-up, causing a decrease in the stock price and inflicting losses to other investors [1], [14], [15]. P&D is an old fraudulent activity that started in the 1700s in London's South Sea Company. Aiming for an easy profit by selling cheap stocks at high prices, a stock owner started positive claims and statements regarding the company and its profit. This fraudulent activity becomes to be known as "the South Sea Bubble", and became an early example of a P&D scheme [1].

Another way to attack cryptocurrencies is to use a distributed denial-of-service (DDoS) attack [16]–[19]. DDoS are cyber-attacks that render a website or a service inoperable by overwhelming it with a flood of traffic. Although blockchains distributed ledgers are robust against DDoS attacks, it is still possible to attack mining nodes that use an outdated protocol [20], or to attack cryptocurrencies mining pools or exchanges [19], [21]–[23]. Although these attacks are not meant to directly steal currencies or affect the network's performance, they are affecting the value of the currency and ultimately lead to the currency's depreciation and benefit the attacker [16]–[19].

A completely different attack based on cryptocurrencies is what called "Cryptojacking" [24]. It leverages the ability of web browsers to execute code. The code in question is meant to "mine" cryptocurrencies. For example, the now-defunct website coinhive.com distributed browser-based cryptomining code that was able to mine bits of the Monero cryptocurrency. The original idea was that it was a way for a user to compensate a website provider by lending some CPU cycles of their browser when accessing the site. This was seen as an alternative to advertisement to monetize 'free access" resources. In-browser cryptomining can also be used for rate limitation as a replacement for CAPTCHAs [25]–[27]. However, this can be abused in the so-called cryptojacking attack, when this is done without the consent of the user or the site owner, or when the code is tampered with, e.g., to modify the payment address [24]–[27]. Cryptojacking attacks are easy to deploy, difficult to detect, and can be found on any Internet-connected device with CPU, such as mobiles, PCs, and the IoTs [24].

In this paper, we conduct a systematic literature review (SLR) on state-of-the-art cryptocurrency-related cybercriminal activities. This SLR aims to provide researchers with a comprehensive literature listing, which is the first step to develop more powerful defensive mechanisms against

[3]https://cbs12.com/news/local/riviera-beach-commissioners-vote-to-pay-ransom-to-hacker-who-shut-down-city-computers

these attacks. To this end, we present a summary of cybercriminal activities related to cryptocurrencies, and the scale of these crimes, as reported in the literature. We then analyze the detection methodologies proposed in the literature, the classifiers used in the process, and how effective these methodologies are. Furthermore, we summarize the sources that can be utilized to collect datasets for cryptocurrency research purposes. Finally, we list out the datasets that have been made publicly available in the literature, as well as some useful tools and resources used to collect and analyze the data. To the best of our knowledge, this SLR is the first one to discuss the different cryptocurrency cybercriminal activities, the proposed defensive mechanisms, and to provide easy access to the public datasets provided in the literature.

This paper is divided into three major sections. Section II describes the review protocol of the SLR, Section III is about data results and Section VI concludes the study.

II. RESEARCH METHOD

In this study, we adopted the standard systematic literature review (SLR) guidelines of Kitchenham and Charters [28], which is ''a means of evaluating and interpreting all available research relevant to a particular research question, topic area, or phenomenon of interest''. The review strategy consists of six steps: 1) research questions, 2) search strategy, 3) study exclusion & inclusion criteria, 4) quality assessment criteria, 5) document retrieval and data extraction, and 6) data synthesis.

Figure 2 describes the steps of the research method and review protocol.

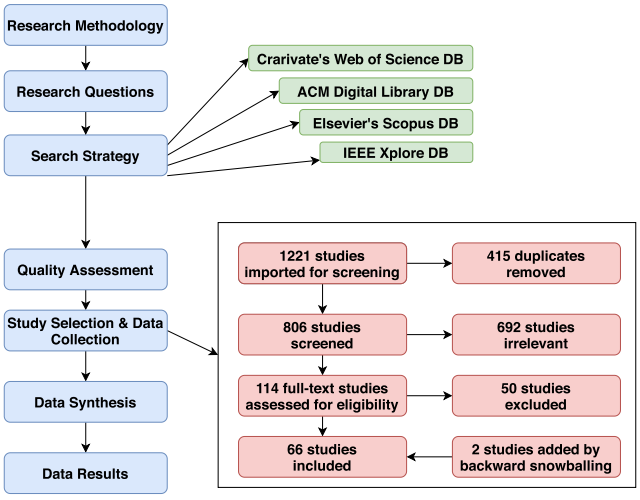

FIGURE 2. Review methodology.

A. RESEARCH QUESTIONS

In this SLR, our aim was to explore the threats that emerged with cryptocurrencies and identify the proposed defensive mechanisms that were developed to prevent these new threats. Moreover, we aimed to provide easy access to the publicly available datasets in the literature. In particular, we addressed the following research questions:

RQ1: With the introduction of cryptocurrencies, what are the types and scales of cybercriminal activities reported by researchers?

RQ2: What are the proposed defensive mechanisms available to detect cybercriminal activities, and what is the reported effectiveness of these mechanisms?

RQ3: For cryptocurrency cybercrimes detection and prevention, what are the public datasets provided in the literature, and how have these datasets been collected?

B. SEARCH STRATEGY

Our search strategy was developed by identifying the two main concepts related to our research questions. The first is the concept of cryptocurrency and its related terms and synonyms. Our second concept refers to the cybercriminal activities that use cryptocurrencies and their synonyms. To increase the effectiveness of our search query, we manually searched on Google Scholar for articles that discuss cybercriminal attacks that target cryptocurrencies and extracted the synonyms of cryptocurrency and cybercriminal activities as used by other researchers. We further included the names of the most used cryptocurrencies in 2019[4],[5] which often represent the primary target for scammers.

Overall, we have collected ten different terms related to cryptocurrencies and sixteen terms related to cybercriminal activities. We then translated the different terms into Boolean logical queries that we executed on four different databases (see Section II-B1) to create our initial dataset of papers. The complete list of terms and the search query are presented in Section II-B2.

The search results on the four databases gave us the list of articles that we used to extract the different synonyms related to our two concepts. These results indicate that our query has a high possibility of returning other articles that contain any of the synonyms included in our query.

1) SOURCE DATABASES

For this systematic review, we used four different scientific and engineering databases and libraries. These databases are the top four databases suggested by our university library for conducting research in Computer Science.

These databases are:

- Elsevier's Scopus database (scopus.com).
- ACM Digital Library database (dl.acm.org).
- Crarivate's Web of Science database (apps.webofknowledge.com).
- IEEE Xplore database (ieeexplore-ieee-org).

2) ABSTRACT SEARCH QUERY

The keywords used to construct the search query are listed Table 1. The finalized search query is the following:

[4]https://www.statista.com/topics/4495/cryptocurrencies/
[5]https://leftronic.com/cryptocurrency-statistics/

TABLE 1. Search query related terms.

Concept	Synonyms
Cryptocurrency	bitcoin, ledger, blockchain, cryptocurrenc*, "crypto-currenc*", "coin mining", Ethereum, litecoin, XRP, and tether
Cybercriminal activities	scam, hyip, "yield* investment program*", ponzi, pyramid, fraud, abuse, "money laundering", ransomware, phishing,"pump & dump", pump-and-dump, *jacking, DoS, "Denial of service", and "Denial-of-service"

```
( bitcoin OR ledger OR blockchain OR
    ↪ cryptocurrenc* OR
 ''crypto-currenc*'' OR ''coin mining''
    ↪ OR Ethereum OR litecoin
 OR xrp, OR tether )
AND
( scam OR hyip OR ''yield* investment
    ↪ program*'' OR
 ponzi OR pyramid OR fraud OR abuse OR
    ↪ ''money laundering''
 OR ransomware OR phishing OR ''pump \&
    ↪ dump'' OR
 pump-and-dump OR *jacking OR DoS OR "
    ↪ Denial of service" OR "Denial-of-
    ↪ service")
```

Initially, we ran the query on the full text of the papers. However, that returned hundreds of irrelevant papers. In order to reduce the results to meaningful, manageable, and relevant results, the search was ultimately limited to the title, abstract, and keyword metadata.

C. INCLUSION CRITERIA

Although our search query is comprehensive and includes popular synonyms related to our research question, other researchers may use other synonyms that we do not know. Accordingly, our query will not detect these papers. Moreover, in our selection process, we may reject some related articles if neither the title, abstract, or keywords contained terms related to our research question. Therefore, we peruse the reference sections of the selected papers in search for additional relevant papers our search might have missed, a technique called "backward snowballing" [29].

D. EXCLUSION CRITERIA

In our search, we excluded non-peer-reviewed journals and conferences. Some of the researchers publish early results of their articles on https://arxiv.org/; we only considered the final versions published in the journals or conferences for such cases in this SLR. We limited our database search to papers written in English. We did not consider an article if the title, abstract, or keywords did not contain the keywords related to our research questions. We limited our search to the papers published after 2009, as the first successful cryptocurrency coin was introduced in 2009 [2].

Finally, We do not include articles submitted to conferences in unrelated fields, such as medical or commerce conferences. Including these conferences adds a large number of mostly unrelated papers, mainly because some of our terms such as "scam" and "fraud" are used in different contexts.

E. QUALITY ASSESSMENT CRITERIA

Identifying quality assessment criteria (QAC) improve SLRs in different ways, such as providing a more detailed inclusion/exclusion criteria and advising recommendations for further research [28]. In our SLR, we considered all the works that meet all of the following assessment criteria:

- The paper has a clear, reproducible methodology.
- The paper presents and discusses a cybercriminal attacks that target cryptocurrencies.

F. STUDY SELECTION & DATA COLLECTION

As shown in Figure 2, the search mentioned above returned 806 unique results. This was reduced to 114 unique and relevant articles after a manual selection process based on reading the papers' titles and abstracts. It was then further reduced to 64 papers once the full text of the papers was read. Finally, two papers were added thanks to the backward snowballing technique, creating a total of 66 unique and relevant papers used in our SLR. Our papers screening and selection were carried out with the aid of Covidence,[6] a web-based software platform that simplifies the production of systematic reviews. It provides an interface to import articles, screen title and abstract, screen the articles full text and export the study results in different formats.

Our data extraction approach was motivated by our research questions. The following pieces of information were manually extracted, assessed, and synthesized:

- (D1) The type of crime(s) being discussed.
- (D2) The analysis evaluation criteria.
- (D3) The cryptocurrency in the study.
- (D4) The dataset source.
- (D5) The dataset availability for public use.
- (D6) The detection algorithm.
- (D7) The detection accuracy and efficiency.
- (D8) The crime effectiveness (based on the USD value and scale).
- (D9) The evaluation of the crime effectiveness.
- (D10) Cryptocurrencies address clustering algorithm (when used).
- (D11) The features used in the classification process (if any).

Our dataset is made publicly available on our team's website and can be reused by other researchers or reproduced

[6]https://www.covidence.org/home

if necessary.[7] Our raw data includes the 11 data records mentioned above, the SLR protocol, the list of articles, the features, and the classifiers used in each article. Our full analysis is presented in section III.

III. DATA RESULTS

In recent years, researchers have been actively working on analyzing the cyberattacks that emerged with the introduction of cryptocurrencies. In 2018, 2019, and 2020 only, 55 papers studying these attacks were published. Moreover, many of these publications proposed defensive mechanisms. Overall, our analysis includes 66 papers. Figure 3 shows the number of articles published per year.

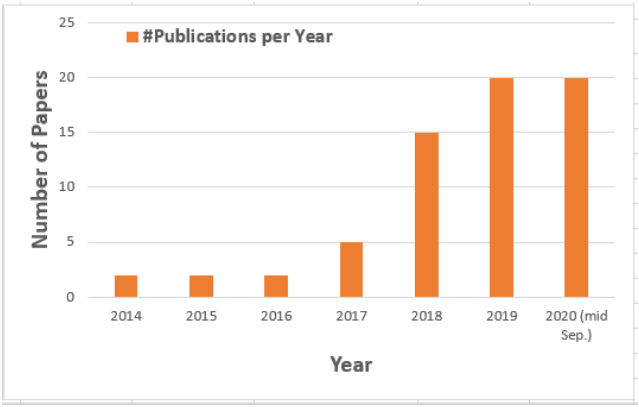

FIGURE 3. Number of papers published per year.

A summary of the papers is shown in Table 2. The table contains basic information about the papers we used in our analysis. In particular, we present the publication year, the publication location (Conference or Journal), the cybercrime type in the discussion, and the targeted cryptocurrency.

A. WITH THE INTRODUCTION OF CRYPTOCURRENCIES, WHAT ARE THE TYPES AND SCALES OF CYBERCRIMINAL ACTIVITIES REPORTED BY RESEARCHERS? [RQ.1]

1) CYBERATTACKS RELATED TO THE CRYPTOCURRENCIES

Several attacks that use cryptocurrencies as a payment medium, such as ''high yield investment programs'' (HYIP), ransomware, and money laundering (ML), have been studied in the literature. Figure 4 shows the breakdown of these attacks and the number of articles that cover each attack. Some papers cover several attacks, so the sum does not add up to the number of papers in our study. As shown Figure 5, the majority of the attacks studied in the literature target Bitcoin, Ethereum, and Monero. As of September 27th, Bitcoin and Ethereum have the highest capitalization market of approximately 199 and 40 billion USD, respectively.[1] The market capitalization of Monero is currently much smaller, just above a billion USD. It is, however, widely used in so-called ''Cryptojacking attacks'' because Monero is specifically designed to not give advantage to ASIC mining. Therefore, any computing device has a fair chance at establishing proof-of-work, and thus hijacking average computers for mining Monero can be profitable. In contrast, the same attack on, e.g. Bitcoin, has little chance of generating any revenues at all. As a result, cryptojacking can be found on any Internet-connected device with a CPU, such as mobiles, PCs, and the Internet of Things [24], [26].

[7] http://ssrg.site.uottawa.ca/slr/

2) THE SCALE OF THE CYBERATTACKS

According to Kshetri and Voas [35], the denial of services and productivity losses due to ransom attacks are in billions of USD. Furthermore, by applying their classification model on features extracted from the transactions of 100K unclassified Bitcoin addresses, Yin and Vatrapu [32] estimate that 10.95% to 29.81% of the Bitcoin addresses are involved in cybercrime activities. These addresses are involved in transactions classified into five different cybercrimes: mixing, ransomware, scam, stolen-bitcoins, and tor-market.

Several datasets and scale measurement techniques were utilized to analyze the fraud activities scale in the literature, including:

- One of the most common scale measurement techniques is estimating the value of stolen money by analyzing the blockchain transaction history of the collected cybercrime addresses; such as in the case of the crimes targeting Bitcoin and Ethereum currencies [6], [7], [11], [12], [36], [39], [43], [50], [56].
- In the case of P&D schemes, the authors inferred an estimation of the theoretical maximum possible profit based on the average P&D events per day and the currency price variation during the P&D event [1], [13]–[15].
- With the high privacy provided by Monero, and with no public available transaction history, the researchers inferred an estimation of cryptojacking attack scale by applying mathematical analysis on information extracted from the cryptojacking campaigns such as the number of visits, the visit duration, the hardware resources usage, CPU utilization, and the number of sites in each scam campaign [26], [27], [40], [44], [52]. For example, Hong *et al.* [40] used the following formula to measure the profit of each cryptojacking campaign.

$$\sum \frac{\#Visitors \times Duration \times HashSpeed}{Difficulty} \times Reward$$

  where $\#Visitors$ is the number of visitors (in millions per month), $Duration$ is the average length of time (in second) a user stays on the site, $HashSpeed$ is ''the average hashing speed of users' processors'' [40], $Difficulty$ is the current hardness of the proof of work, and $Reward$ is the block reward at the time of analysis.
- Other researchers provided an estimation based on extrapolating the results of their classification model or by applying the classifier on an unknown dataset [32], [41], [48]. For example, Yin and Vatrapu [32] reported the results of applying their classification model on 100k unclassified addresses while Chen *et al.* [41], [48] used

TABLE 2. Summary of the papers included in this SLR.

	Reference	Published Year	Journal/ Conference	Crime Type (D1)	Cryptocurrency (D3)
1	[30]	2014	Journal	ML	Bitcoin
2	[16]	2014	Journal	DDoS	Bitcoin
3	[8]	2015	Conference	ML	Bitcoin
4	[6]	2015	Conference	Services detection	Bitcoin
5	[11]	2016	Conference	Ransom	Bitcoin
6	[31]	2016	Journal	ML	Bitcoin
7	[32]	2017	Conference	Services detection	Bitcoin
8	[33]	2017	Conference	HYIP	Bitcoin
9	[34]	2017	Conference	Mining/Jacking	Bitcoin,Ethereum, etc..
10	[35]	2017	Journal	Ransom	Bitcoin
11	[17]	2017	Journal	DDoS	Bitcoin
12	[5]	2018	Conference	Services detection	Bitcoin
13	[36]	2018	Conference	Ransom	Bitcoin
14	[37]	2018	Conference	ML	Bitcoin
15	[7]	2018	Conference	HYIP	Bitcoin
16	[38]	2018	Conference	ML	Bitcoin
17	[39]	2018	Conference	Phishing	Bitcoin
18	[12]	2018	Conference	Ransom	Bitcoin
19	[1]	2018	Journal	P&D	Bitcoin,Ethereum, etc..
20	[40]	2018	Conference	Mining/Jacking	Monero
21	[41]	2018	Conference	HYIP	Ethereum
22	[42]	2018	Journal	ML	Bitcoin
23	[43]	2018	Journal	Ransom	Bitcoin
24	[44]	2018	Conference	Mining/Jacking	Monero
25	[3]	2018	Conference	HYIP	Bitcoin
26	[45]	2018	Conference	HYIP	Bitcoin
27	[46]	2019	Conference	ML	Bitcoin
28	[47]	2019	Conference	HYIP	Ethereum
29	[48]	2019	Journal	HYIP	Ethereum
30	[4]	2019	Journal	HYIP	Bitcoin
31	[25]	2019	Conference	Mining/Jacking	Monero, Litecoin, Zcash
32	[49]	2019	Journal	Mining/Jacking	Bitcoin,Ethereum, etc..
33	[50]	2019	Conference	Honeypot	Ethereum
34	[51]	2019	Conference	General	Ethereum
35	[52]	2019	Conference	Mining/Jacking	Monero and others
36	[24]	2019	Journal	Mining/Jacking	Monero
37	[14]	2019	Conference	P&D	Bitcoin
38	[26]	2019	Conference	Mining/Jacking	Monero, JSECoin
39	[27]	2019	Conference	Mining/Jacking	Monero
40	[53]	2019	Journal	ML	Bitcoin,Monero, etc..
41	[54]	2019	Conference	Mining/Jacking	Monero
42	[15]	2019	Conference	P&D	Bitcoin,Ethereum, etc..
43	[13]	2019	Conference	P&D	Bitcoin,Ethereum, etc..
44	[55]	2019	Conference	Mining/Jacking	Ethereum, Monero, and Zcash
45	[18]	2019	Conference	DDoS	Bitcoin
46	[19]	2019	Conference	DDoS	Bitcoin
47	[56]	2020	Journal	HYIP	Ethereum
48	[57]	2020	Journal	ML	General
49	[58]	2020	Journal	General	Ethereum
50	[59]	2020	Conference	ML	Bitcoin
51	[60]	2020	Conference	ML	Bitcoin
52	[61]	2020	Conference	General	EOS
53	[62]	2020	Journal	ML/ransom	Bitcoin
54	[63]	2020	Conference	Mining/Jacking	General
55	[64]	2020	Conference	Mining/Jacking	General
56	[65]	2020	Journal	Ransom	Ethereum
57	[66]	2020	Conference	honeypot	Ethereum
58	[67]	2020	Conference	Ransom	Ethereum
59	[68]	2020	Conference	ML	Bitcoin
60	[69]	2020	Conference	General	Bitcoin, Ethereum
61	[70]	2020	Journal	Ransom	Bitcoin
62	[71]	2020	Journal	Phishing	General
63	[72]	2020	Conference	ML	General
64	[73]	2020	Conference	ML	General
65	[74]	2020	Journal	General	ETH
66	[75]	2020	Journal	ML	Bitcoin

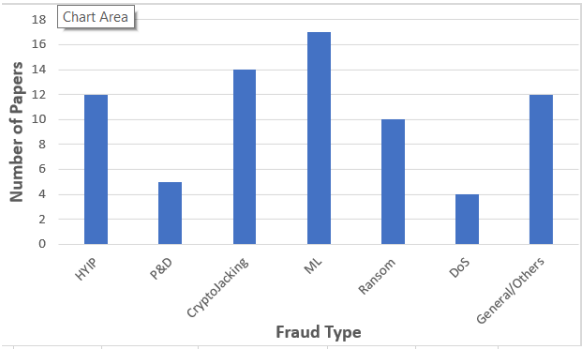

FIGURE 4. Number of published articles per fraud type.

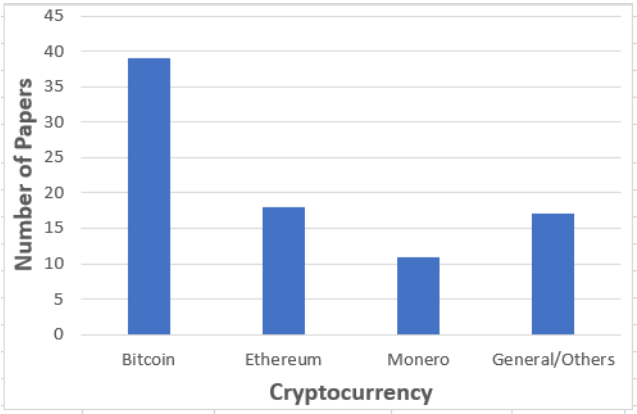

FIGURE 5. Number of published papers per currency.

the classification model precision and recall values to estimate the scale of smart Ponzi schemes on Ethereum.

In this section, we report the attacks with the highest number of victims and the ones with the highest profit for each type of crime.[8] We provide the full scale as reported in the literature in Table 7 in Appendix.

As researchers conducted different studies and analyses in the literature, the scales of the cybercrime activities were reported in many ways, even for the same cybercrime and the same cryptocurrency. A breakdown of the scale of these activities addressed in the literature is:

1) **HYIP (Bitcoin):** The Pirate@40's HYIP scheme had raised 700,000 Bitcoin from the investors before they were charged by the Security and Exchange Commission (SEC) in 2013 [45]. Bartoletti *et al.* reported an estimate of 10 million USD in [7], and Vasek and Moore [3] reported that 11,990 users have responded to 1,780 different scams on the bitcointalk forum.
2) **HYIP (Ethereum):** In the litterature, we find that 0.03% [48] to 0.15% [41] of the smart contracts are HYIP. In [56], the authors estimated the value of HYIP with Ethereum is approximately half a million USD.
3) **Phishing:** Holub and O'Connor reported that 50 million USD were stolen by the attackers in 3 years [39].
4) **Ransom:** The scale of the ransomware was reported as the payment values received by the attackers. Conti *et al.* [35], reported a ransom payment of 7,059.9 Bitcoin (∼ 2.8 million USD), Liao *et al.* [11] reported payments of 1,128.40 Bitcoin (∼ 310,000 USD) over a 5 months period, and Huang *et al.* [36] reported the highest ransom value, 16 million USD paid by 19,750 victims. However, the main monetary loss due to ransom is the denial of services and productivity losses, which are estimated in billions of USD from about 300,000 infected computers in 150 countries [35].
5) **P&D**: In the literature, it is estimated that on average, 1.6 [1] to 2 [15] P&D events are organized per day. Xu and Livshits [13] estimated that P&D events generate an aggregate, artificial trading volume of 6 million USD a month. In [14], Chen *et al.* analyzed a leaked transaction history of the Mt. Gox Bitcoin Exchange from April 2011 to November 2013, and reported that the transactions with an abnormal price involved 13.09% of the users in the dataset.
6) **Mining/Cryptojacking:** As transactions history can not be accessed with Monero, the scale of cryptojacking has been estimated using the CPU usage consumed by the mining scripts and the size of the campaign. Zimba *et al.* [24] estimated that 32% of the US users are exposed to browser-based crypto mining. Additionally, Hong *et al.* [40] estimated that 10 million web users are affected by cryptojacking monthly, at a daily cost of 59,000 USD due to 278K kWh of extra power consumption. Furthermore, in [44], the profit of each cryptojacking campaign is estimated at 14.36 USD to 31,060.80 USD per month on average, while in [27] it was estimated at 340 USD per campaign per day (about 10,200 USD per month).

[8]The full raw data is available on our public repository **http://ssrg.site.uottawa.ca/slr/**

B. FOR CRYPTOCURRENCY CYBERCRIMES DETECTION AND PREVENTION, WHAT ARE THE PUBLIC DATASETS PROVIDED IN THE LITERATURE, AND HOW HAVE THESE DATASETS BEEN COLLECTED? [RQ.3]

In this section, we present the resources used in the literature to collect datasets on which to train, detect, and analyze the attacks discussed in each paper. Some researchers collected the training data manually, e.g. by searching on online fora such as bitcointalk.org [7]. Other researchers used a semi-automated crawling process followed by manual data collection [3]–[5]. Furthermore, some datasets were collected by extracting the system resource usage data of the devices under attack [40], [49], [54]. Our analysis shows that four different resources were used to prepare the training dataset:

1) Collecting data from online fora and blogs, such as bitcointalk.org and Reddit [3]–[7], [11], [13], [16], [17], [33], [36], [43], [45], [68]. The researchers relied on crawling these fora as they are used by scammers to advertise for their schemes. For example, Vasek and

Moore [3] crawled the entire history (from June 2011 to November 2016) of the bitcointalk.org subforums that scammers use to advertise Ponzi schemes. Their crawling returned 11,424 threads, which they further refined to 2,617 threads by removing threads discussing online card games and only including threads that contain URLs or bitcoin address for the scam. In [43], the authors collected the scam addresses manually by searching online ransomware knowledge base (such as Kaspersky Lab, ESET, Symantec, and Malwarebytes), ransomware removal guides (such as MalwareTips.com, BleepingComputer.com, and 2-spyware.com), online fora where researchers and victims publish their data (such as Reddit), and available ransomware screenshots in different search engines image databases (such as Yahoo and Google).

2) Using dataset provided by third parties including previous studies, https://chainalysis.com/, and public blacklists [12], [14], [19], [25], [27], [32], [39], [41], [47], [50], [51], [58]–[61], [64], [66], [70], [74]. For example, Chen *et al.* [14] used a leaked transaction history of Mt. Gox Bitcoin Exchange from April 2011 to November 2013 in their analysis. Chen *et al.* [41] and Jung *et al.* [47] used the dataset from Bartoletti *et al.* [56] in their study. Ostapowicz and Zbikowski [51] used the wallets reported in *Etherscan.io* as being used in fraud activities.
3) Collecting data from free online sources, online exchanges, Telegram groups, and smart contracts with public source code. These sources require manual analysis to distinguish between scam and benign data [1], [15], [18], [26], [40], [48], [52], [56], [63], [69]. For example, Kamps and Kleinberg [1] used the CCXT python library to collected cryptocurrencies Market data from April 2018 to May 2018 from a variety of cryptocurrency exchanges, including Binance, Bittrex, Kraken, Kucoin, and Lbank. Victor and Hagemann [15] collected the price and volume of cryptocurrencies from Binance exchange, the chat histories from Telegram P&D groups (fraud ads), and general data about the currency capitalization from *coinmarketcap.com*.
4) Collecting system resources, such as system runtime parameters [40], [49], [54]. For example, Ning *et al.* [54] collected 12 system runtime parameters (such as interrupts per second, page reads/write/-fault per second, and packets received/sent per second) from 13 different devices while running 5 different application on each device.

Further breakdown is provided in Table 8 in Appendix.

In some activities such as HYIP, authors had difficulties collecting a large number of addresses. In most cases, they manually visited online fora to collect scam addresses advertised by the scammers. However, in many instances, the addresses were not included in the posts. In such cases, the authors visited the HYIP website and manually extracted the deposit address. When the websites were no longer online, the researchers tried to recover old snapshots through the Internet Archive [7]. To increase the number of collected addresses, some authors are using ''multiplier'' techniques. We have found two such techniques used in the literature:

- **Multi-input heuristic:** in this heuristic, the assumption is that the same person owns all the addresses on the input side of any transactions [4], [5], [7], [11], [33], [36], [43], [45].
- **Shadow/change address algorithm:** in this heuristic, the assumption is that if there are only two addresses in the output side of any transactions, and one address has appeared before in the blockchain while the other address has not been used before, then it can be safely assumed that the new address is a shadow/change address generated to accept the change from the transaction back to the sender, and thus is owned by the sender [4], [11], [43].

Many of the authors have disclosed their datasets, which, in turn, provides an opportunity for other researchers to use them. In fact, the dataset prepared by Bartoletti *et al.* [56] was later used in [41], [47] to implement defensive mechanisms against the cybercriminal activities that utilize cryptocurrencies as a payment medium. The full list of disclosed datasets in the literature is presented in Table 3. If a dataset is used in more than one research paper, we only show the most recent publication year in the table. Furthermore, Table 4 contains other supplementary websites and tools that can be used in the data collection and analysis process.

C. WHAT ARE THE PROPOSED DEFENSIVE MECHANISMS AVAILABLE TO DETECT CYBERCRIMINAL ACTIVITIES, AND WHAT IS THE REPORTED EFFECTIVENESS OF THESE MECHANISMS? [RQ.2]

State of the art defensive mechanisms reported in the literature are usually based on extracting distinguishing features from the training dataset and using these features to train a classifier such as random forest (RF) [4], [5], [7], [13], [33], [34], [47], [48], [51], XGBoost [14], [32], [41], [51] or support vector machine (SVM) [26], [51] to tell benign data apart from cybercrime data. The features are based on the type of cybercrime being discussed and the type of the available dataset. For example, to detect HYIP schemes in Bitcoin and Ethereum, publicly readable blockchain transaction records and smart contract code are leveraged [3], [6], [47], [48]. In Table 5, we present six different types of sources that are used in the literature to extract the features and examples of these features for each source type.[8]

The effectiveness of the proposed detection mechanisms varies from a 0-day detection model, in which the scam is detected as soon as it is posted [47], to models that require the attack to have victims as these mechanisms depend on extracting features from the scam transactions history [4]. Several measurement metrics were used in the literature to report how successful the proposed mechanisms were. The measurements most often used are:

TABLE 3. Publicly available data provided in the literature.

Dataset (D5)	Description	Reference	Published Year
https://bitbucket.org/mhuzai/mineguard/src/master/	The data and code used in the paper.	[34]	2017
https://goo.gl/sQJKdx	List of Bitcoin addresses categorized per the service they are used in.	[5]	2018
https://goo.gl/ToCho7	List of scam Bitcoin addresses.	[7]	2018
https://osf.io/827wd/	The data and code used in the paper.	[1]	2018
https://github.com/deluser8/cmtracker	The data and code used in the paper.	[40]	2018
https://www.walletexplorer.com/wallet/Btcst.com-pirateat40/addresses	Btcst.com-pirateat40 Bitcoin addresses.	[45]	2018
https://github.com/teamnsrg/outguard	The data and code used in the paper.	[26]	2019
https://github.com/hoshsadiq/adblock-nocoin-list	NoCoin adblock list. Block "browser-based crypto mining".	[25]	2019
https://goo.gl/k5PCOZ	List of scam domains and Bitcoin addresses.	[4], [33]	2019
https://github.com/pan-unit42/iocs/blob/master/6908_of_8712_coin_mining_urls_in_pandb.txt	List of URLs contains scam coin mining code.	[52]	2019
https://github.com/pan-unit42/iocs/blob/master/4457_of_4633_scam_js_urls_in_pandb.txt	List of URLs contains scam JS code.	[52]	2019
https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/25541	142 distinct DDoS attack reports on 40 Bitcoin services	[16], [19]	2019
https://github.com/blockchain-unica/ethereum-ponzi	The data and code used in the paper.	[56]	2020
https://goo.gl/CvdxBp	List of scam Bitcoin addresses.	[41], [47], [56]	2020
Hardcoded in the paper	List of Bitcoin addresses, mining script URLs, online mixers, etc...	[12], [27], [43], [50], [68]	2020
https://bit.ly/32pmC2A	Dataset and code used in the paper	[58]	2020

- **True Positive Rate (TPR)**, the number of cybercrime instances that were successfully identified as cybercrime.
- **False Positive Rate (FPR)**, the number of cybercrime instances that were wrongly identified as benign.
- **Precision**, the ratio of actual cybercrime instances to all those classified as cybercrime.
- **Recall**, the ratio of correctly classified cybercrime instances to all cybercrime instances in the training set.

In this section, we report the breakdown of the mechanisms with the highest detection rate per crime type as follows[8]:

- **HYIP (Bitcoin):** a TPR of 95% and an FPR of 4.9% was reported by Toyoda *et al.* [4], and Bartoletti *et al.* [7] proposed a detection mechanism with 96.8% TPR and a recall of 96.9%. However, the proposed mechanisms for Bitcoin HYIP detection do not provide early detection and defensive models. They depend on classifying previously reported scam campaigns to extract features from the transactions history of the addresses.
- The proposed detection mechanisms fairs better with **HYIP (Ethereum) schemes** detection. 0-day detection models that can detect HYIP schemes in smart contracts at the moment of creation were proposed in [41], [47], [48]. For example, the model proposed by Jung *et al.* [47] reports a precision of 0.99 and a recall of 0.97 for full data analysis and a precision of 0.98 and recall of 0.96 for 0-day detection.
- **P&D** detection mechanisms depend on features extracted from the market movement such as market capitalization, the price, and the volume. As a result, it was possible to develop a model that predicts the likelihood of a cryptocurrency being pumped before the actual pump event [13].
- **Cryptojacking** detection methodologies achieved a high detection rate of 99.7% [34], and TPR of more than

TABLE 4. Useful resources provided in the literature.

URL (D5)	Description	Reference
https://github.com/bitcoinponzi	A public tool provided by the author for features extractions from the transactions history	[7]
https://github.com/ccxt/ccxt	CCXT python library to collect data from variety of cryptocurrency exchanges.	[1]
https://bitcointalk.org/index.php?topic=75883.0	List of clean gampling domains.	[3]
https://go.aws/2R1Jktx	Alexa top 1M domains list.	[27]
https://coinmarketcap.com/	Market Capitalization related data.	[27]
https://chromedevtools.github.io/devtools-protocol/	Allows for tools to instrument, inspect, debug and profile Chromium, Chrome and other Blink-based browsers.	[27]
https://github.com/binance-exchange/binance-official-apidocs/blob/master/web-socket-streams.md	Binance API.	[15]
https://github.com/LonamiWebs/Telethon	Telegram's API.	[15]
https://etherscamdb.info/scams	Scam DB.	[58]
https://github.com/hashbaby-com/eos-hall-of-shame/tree/master/bots	Bot index (list of bots in ESO).	[61]
https://bit.ly/38PButR	Scam domains/addresses dataste provided by CryptoScamDB.org.	[69]
https://urlscan.io/	Online service that scans and analyze websites.	[69]
https://www.malware-traffic-analysis.net	Malware traffic analysis.	[70]
https://virusshare.com/	Malware samples.	[70]
https://github.com/ytisf/theZoo/tree/master/malwares/Binaries	Malware repository.	[70]
https://github.com/twintproject/twint	Twitter intelligence tTool to scrape tweets.	[18]

TABLE 5. Utilized sources for extracting detection features.

Sources to extract features from (D2)	Examples of extracted features (D11)	Reference
The HTML code	Used global variable, WebSocket messages, and alert text.	[40], [44], [52]
The web technology and generated traffic	The use of proxies, number of packets per minute, and the IP addresses of the cryptomining domains.	[24], [25], [55]
The system resources.	(Interrupts,page read/write)/second, parallel tasks, and L1-dcache-loads.	[26], [27], [34], [49], [54], [63], [64], [70]
The cryptocurrency addresses transaction history	The frequency of transactions, the ratio of in/out transactions, and the lifetime of the address.	[4], [5], [7], [19], [32], [33], [51], [59]–[61], [74]
The cryptocurrency market movement (price, volume, etc.)	Market capitalizations, Volumes in coin, Entropy, and stability.	[1], [13]–[15]
The Ethereum account and smart contract code	Number of in/out transactions, existence of an unconditional jump, and the frequency of all the opcodes used in the smart contracts	[41], [47], [48], [50], [58].

99% [26], [34], [54], [55]. Furthermore, the detection model proposed by Ning *et al.* [54] detects 87% of the mining scripts "instantly[9]", and detects 99% of the scripts within a window of 11 seconds.

A breakdown of the classifiers used by the papers and the results achieved is presented in Table 6 in Appendix.

[9]The authors do not explain what "instantly" means in that context.

IV. THREATS TO VALIDITY AND LIMITATIONS

The validity of research is concerned with the alignment between reality and research conclusions [76]. This section discusses the biases, threats, and limitations that may affect this SLR and the 66 papers we reviewed.

One of our study's threats is our search query's ability to detect all the related articles. In our research query, we have included the most common terms related to cryptocurrency and cybercriminal activities, which we supplemented with different synonyms of cryptocurrency and cybercrimes used

TABLE 6. The reported detection results in the literature.

Reference	Crime Type (D1)	Detection Algorithm (D6)	Achieved Results (D7)
[32]	Service Detection	Bagging and XGBoost	80.76% accuracy (XGBoost) and 78.46% accuracy (Bagging)
[5]	Service Detection	Random forest	72% accuracy (owner-based scheme) and 70% accuracy (address-based scheme)
[51]	Service Detection	Random forest	23.67% TPR and 0.02% FPR
[33]	HYIP (Bitcoin)	Random forest	83% TPR and 4.4% FPR
[7]	HYIP (Bitcoin)	Random forest	96.8% TPR and 96.9% Recall
[4]	HYIP (Bitcoin)	Random forest	95% TPR and 4.9% FP
[41]	HYIP (Ethereum)	XGBoost	94% precision and 81% recall
[47]	HYIP (Ethereum)	Random forest	99% precision and 97% recall (full data), and 98% precision and 96% recal (0 day detection)
[48]	HYIP (Ethereum)	Random forest	95% precision and 69% recall
[34]	Mining/Jacking	Random forest	99.7% TPR and FPR less than 0.25%
[49]	Mining/Jacking	K-Nearest Neighbors	88% precision and 87% F1 score
[26]	Mining/Jacking	Support vector machine	97.9% TPR and 1.1% FPR
[54]	Mining/Jacking	Capsule Network	Detect 87% of the instance instantly and 99% of the instances within a window of 11 seconds.
[55]	Mining/Jacking	Proximity-based classification	99.7% TPR and 46.1% FPR
[15]	P&D	XGBoost	99.5% AUC, 85.5% sensitivity and 99.7% specificity
[13]	P&D	Random forest	Predicts the likelihood of a currency being pumped with an area under curve of over 90%
[58]	Service Detection	XGBoost	Accuracy of 96.3%
[59]	ML	Graph Convolutional Networks (GCN)	Accuracy of 97.4%
[60]	ML	Ensemble learning	Accuracy of 98.13%
[61]	General	RF	Accuracy of 99.55%
[70]	Ransom	Bayesian belief network (BBN)	Accuracy of 97.5%
[74]	General	XGBoost	Accuracy of more than 96%
[16]	DDoS	Word-based classifier	Accuracy of 75%
[19]	DDoS	Multilayer perceptron (MLP)	High accuracy with 12 layers and higher training epochs.

by other researchers. However, it is possible to miss some of the relevant papers that have not used those keywords. Hence, we enhanced our dataset by using backward snowballing to reduce the effect of this threat. As reported by Brings *et al.* [77], snowballing is an efficient way to complement a keywords-driven database search. In our case, only two papers were added via backward snowballing, which suggests that our keywords-driven database search was effective at finding the right literature.

Another threat is the validity of the considered studies, the extracted data, and the conducted analysis, which may be affected by the authors' biases. As a result, any imprecision and weakness in the selection or analysis stages could threaten the accuracy of the answers. To mitigate these threats, the authors have discussed the problematic papers and followed a consensus-building and analysis mechanism to validate the selected papers and answer the research questions.

Finally, utilizing the top four databases suggested by our university library for researching Computer Science, including two generic ones (Scopus and Web of Science), allows us to detect a large portion of the papers stored in scientific databases. Nevertheless, we may miss some of the relevant papers stored in other databases and not reachable with these search engines.

One of our study's main limitations is that hundreds of unrelated articles were returned when we executed our search query on the different databases. We excluded the articles submitted to non-related conferences, such as medical and commerce conferences, to reduce the results to meaningful, manageable, and relevant results, which may lead to losing some of the relevant papers. However, we believe that the most relevant articles are submitted to the scientific conferences and journals that we considered in our search,

Another limitation of our study is that we are not studying attacks targeting cryptocurrency systems themselves. We are instead looking at attacks that are somehow leveraging cryptosystems for personal gains. There is a large body of literature studying and criticizing either blockchain protocols or the protocols of the cryptocurrencies using these chains.

For example, a DDoS attack targeting Monero mining nodes running an outdated consensus protocol was discussed in [20], and a theoretical analysis of DDoS attacks targeting Bitcoin mining pools were presented in [21]–[23]. Overall, we have already identified 13 such articles during our full-text review, but discarded them because the topic did not line up with this SLR. In our future work, we would like to conduct another review of these articles and articles discussing other types of attacks, such as the double-spending attacks [78].

Finally, we limit ourselves to research published in English. Therefore, we may overlook some of the non-English literature discussing this topic.

V. RELATED WORK

Several articles have discussed the rise in cryptocurrencies' popularity and how they are used in different financial crimes. In [79], Kethineni and Cao have presented a systematic study of crimes related to virtual currencies through the lens of court documents and news reports. They have focused on understanding the role played by cryptocurrencies in criminal activities, the factors facilitating criminal activities related to cryptocurrencies, the role politics plays in regulating cryptocurrencies, and the challenges cryptocurrencies are posing for regulators and law enforcement. The authors have reported that criminals favor cryptocurrencies because they are secure, not monitored through centralized repositories, they are not regulated, and are stored virtually, which reduces the need for physical spaces. They further reported that BTC is the dominant cryptocurrency in criminal activities such as money laundering, hacking, and drug/sex trafficking. In our SLR, we conducted our analysis through the lens of technical academic papers. We have also reported the crimes' scale, the most targeted currencies by the crimes, and the proposed defensive mechanisms. Furthermore, we summarized the sources to collect datasets and listed out the disclosed ones.

In [80], [81], the authors have conducted literature reviews of scam and fraud attacks related to blockchains. Saad *et al.* [80] have categorized the attacks into three types. First, attacks targeting blockchain cryptographic constructs, such as forks and orphaned blocks. Second, attacks targeting these distributed architecture of the systems, such as DNS hijacks, eclipse attacks, or selfish mining. Third, attacks targeting blockchain application contexts, such as double-spending, wallet theft, and cryptojacking. They have also studied the relationships between the attacks and demonstrated the connections between various attack vectors. Finally, they have presented defense mechanisms proposed by researchers or already implemented by blockchain technology to mitigate these attacks' effects. Phan *et al.* [81] summarized more than 30 papers that discuss different types of attacks, such as 51% attack, selfish mining, Ponzi schemes, and Denial-of-Service attacks. They further analyzed more than 7,000 tweets to analyze people's reactions towards fraud and to identify major social media trends. They have reported that the most frequent words used in tweets that include both ''blockchain'' and ''fraud'' include ''scam'', ''bitcoin'', ''payment'' and ''crypto''. In our SLR, we have also summarized several papers discussing scam and fraud attacks related to cryptocurrencies. However, we have also analyzed and presented the sources used to extract datasets and how they were utilized to extract distinguishing features for training classifiers. We have also listed the publicly available datasets and other supplementary websites and data collection and analysis tools.

In [82], Teichmann *et al.* have conducted a qualitative study of how cryptocurrencies are used in financial crime such as money laundering, terrorist financing, and corruption. For this, the authors have interviewed international compliance experts and illegal financial services. They reported that criminals favor cryptocurrencies because of the anonymity, the absence of regulation, the reduced risk when compared to storing physical assets, and the suitability for money laundering. They concluded that an international regulatory standard should be adopted to eliminate financial crimes using cryptocurrency. They further proposed to use the Liechtenstein blockchain act as a benchmark to regulate blockchain. In our SLR, we based our analysis on the technical academic papers rather than conducting interviews. We have also focused on studying other questions: the types and scales of the cybercriminal activities, the proposed defensive mechanisms, and how to collect datasets to be used for cybercrimes detection and prevention, and whether there exist any disclosed datasets in the literature.

In [83], Chen *et al.* have surveyed three aspects of Ethereum systems security: vulnerabilities, attacks, and proposed defenses. They have reported 40 different vulnerabilities and 29 different attacks. They further grouped the vulnerabilities and attacks according to the targeted location: application layer, data layer, consensus layer, and network layer. Furthermore, the authors have reported four root causes vulnerabilities; 1) smart contract programming, 2) solidity language and toolchain, e.g. buggy compilers with insufficient toolchain support, 3) Ethereum design and implementation, e,g. not validating the input data, and 4) human factors, e.g. improper configuration of the Ethereum client. Finally, the authors have reported several consequences of the attacks on Ethereum, such as unfair income and double-spending. They presented 51 defense mechanisms to secure the Ethereum ecosystem. In our SLR, we have surveyed articles discussing different cryptocurrencies and did not limit our analysis to a specific cryptocurrency,

Other researchers have studied and analyzed a single attack or illegal activity, such as money laundering [84], [85] or cryptojacking [86]. A systematic literature review of the research articles discussing using cryptocurrencies in money laundering (cryptolaundering) from 2009 to 2018 was conducted in [84]. The study confirmed that cryptolaundering could be considered a complex socio-technical system where multiple entities interact throughout the process, including humans, technical factors, as well as organizational and social factors. In [85], Dupuis and Gleason have conducted

TABLE 7. Reported cybercrimes scale estimation in the literature.

Reference	Crime Type (D1)	Currency (D3)	Scale (D8)
Analyzing the transaction history of the collected addresses			
[6]	Services detection	Bitcoin	Scam addresses received 11 million USD from 13,000 distinct victims, and returned back 4 million USD.
[11]	Ransom	Bitcoin	Scam addresses received 1,128.40 Bitcoin (310,472.38 USD) in the period from September 2013 through January 2014.
[36]	Ransom	Bitcoin	Scam addresses received 16 million USD from 19,750 victims.
[7]	HYIP	Bitcoin	scam addresses received around 10 million USD.
[39]	Phishing	Bitcoin	Scam addresses received over 50 million USD in 3 years.
[12]	Ransom	Bitcoin	Scam addresses received 169 bitcoins.
[43]	Ransom	Bitcoin	Scam addresses received 7059.9 Bitcoin (2,834,468 USD).
[50]	Honeypot	Ethereum	690 honeypot smart contracts that accumulated profit of more than 90,000 USD from 240 victims.
[56]	HYIP	Ethereum	scam addresses received almost 0.5 million USD.
[16]	DDoS	Bitcoin	7.4% of Bitcoin-related services and 60% of large mining pools have been DDoSed
[17]	DDoS	Bitcoin	Reduce the daily number of big transactions
[18]	DDoS	Bitcoin	Reduce the average trading volume during the attack
Inferred an Estimation based on the average P&D events and the currency price variation during the event			
[1]	P&D	Bitcoin, Ethereum, etc..	2,150 P&D schemes over 20 days of crawling witn an average of 1.6 P&D events per currency per day.
[14]	P&D	Bitcoin	Found 471,899 (0.04% the full dataset) abnormal price records in Mt. Gox leaked dataset. The abnormal price transactions involved 16,660 (13.09%) of the users.
[15]	P&D	Bitcoin, Ethereum, etc..	612 P&D schemes.
[13]	P&D	Bitcoin, Ethereum, etc..	Found 100 organized Telegram P&D channels that coordinates 2 P&D events per day on average. These events generates an aggregate artificial trading volume of 6 million USD a month. Furthermore, the authors reported that some online exchanges are active participants in the P&D schemes
Inferred an estimation by applying mathematical analysis on cryptojacking campaigns extracted data			
[40]	Mining/Jacking	Monero	2,770 unique cryptojacking domain, including 868 among Alexa top 100K. Estimates that the cryptojacking affects 10 million web users per month and generate over $59K daily by consuming 278K kWh extra power.
[44]	Mining/Jacking	Monero	Estimates that each cryptoJacking campaign profit from 14.36 USD to 31,060.80 USD per month on average.
[52]	Mining/Jacking	Monero and others	3,487 mining domains, including 1,295 among Alexa top 1M. Furthermore, the authors reported that many mining domains have lived more than four years and received more than tens of millions of DNS resolutions.
[26]	Mining/Jacking	Monero, JSECoin	6,302 unique cryptojacking domains, including 828 among Alexa top 1M.
[27]	Mining/Jacking	General	109 cryptojacking websites from 659 websites that employed Wasm code fro Alexa top 1M.
[63]	Mining/Jacking	Monero	Estimates that 0.2% of Alexa top 1M domains contains mining scripts and it generates up to 340 USD per day.
Estimation based on extrapolating classification model results			
[32]	Services detection	Bitcoin	Estimates that the percentage of cybercrime-related addresses is 29.81% according to Bagging classifier, and 10.95% according to Gradient Boosting classifier.
[41]	HYIP	Ethereum	Estimates that 434 (0.15%) of the contracts on Ethereum platform before May 7, 2017 are Ponzi.
[48]	HYIP	Ethereum	Estimates that 507 (0.03%) of all the contracts before May 7, 2017 are Ponzi.

TABLE 8. Sources used in the literature for dataset collection.

Reference	Crime Type (D1)	Currency D(3)	Dataset Source (D4)
[6]	Services detection	Bitcoin	Bitcointalk.org and Cryptohyips.com.
[11]	Ransom	Bitcoin	Online fora.
[32]	Services detection	Bitcoin	Dataset provided by Chainalysis.com
[33]	HYIP	Bitcoin	Bitcointalk.org and Blockchain.info/tags.
[5]	Services detection	Bitcoin	Blockchain.info/tags, WalletExplorer.com, and BitcoinTalk.org.
[36]	Ransom	Bitcoin	Executed ransom binaries and collected the addresses from the memory dump, created files, and screenshots resulted from the ransom. They further used search engines for screenshots with ransom addresses provided by previous victims.
[7]	HYIP	Bitcoin	Bitcointalk.com
[39]	Phishing	Bitcoin	isco Systems, Inc. and Ukraine Cyberpolice.
[12]	Ransom	Bitcoin	Previously reported wannacry ransom addresses.
[1]	P&D	Bitcoin, Ethereum, etc..	Cryptocurrencies market data from online exchanges (Binance, Bittrex, Kraken, Kucoin and Lbank) using CCXT python library.
[40]	Mining/Jacking	Monero	Alexa top 100k domains. regularly visits the websites to Collect traces using Hash Based Profiler (search for hashing traces) and Stack Structure Based Profiler (search for heavy workloads with repeated behavioral patterns in the stack execution).
[41], [47], [50], [66]	HYIP/Honeypot	Ethereum	Previous study.
[43]	Ransom	Bitcoin	Ransomware knowledge base, ransomware removal guides, online blogs, and available ransomware screenshots in different search engines image databases.
[3]	HYIP	Bitcoin	Bitcointalk.org subforums (Scam accusations, Games and Rounds, and Investment Games).
[45]	HYIP	Bitcoin	Pirate\@40 scheme addresses accessed through Bitcointalk.com and/or WalletExplorer.com
[48]	HYIP	Ethereum	Open source smart contracts on the Ethereum platform.
[4]	HYIP	Bitcoin	Bitcointalk.org and WalletExplorer.com
[25]	Mining/Jacking	Monero, Zcash, Litecoin	Nocoin blacklist.
[49]	Mining/Jacking	Bitcoin, Ethereum, etc..	Device side-channel magnetic field signals generated from cryptocurrencies mining algorithms.
[51]	General	Ethereum	Etherscan.io.
[52]	Mining/Jacking	Monero etc..	Monitoring Daily feed of suspicious URLs visited by Palo Alto Networks customers and Alexa top 1M domains.
[14]	P&D	Bitcoin	Mt. Gox leaked addresses transaction history.
[26]	Mining/Jacking	Monero and JSECoin	Applying existing cryptojacking detection tools to scan Alexa top 1M domains.
[27]	Mining/Jacking	Monero	previously published mining script.
[54]	Mining/Jacking	Monero	System runtime parameters of the studied computer/phone.
[15]	P&D	Bitcoin, Ethereum, etc..	Ccurrency price and volume (Binance exchange), fraud ads (Telegram P&D groups), and currency capitalization data (coinmarketcap.com).
[13]	P&D	Bitcoin, Ethereum, etc..	Telegeram channels.
[56]	HYIP	Ethereum	Open source smart contracts on the Ethereum platform.
[58]	General	ETH	Etherscamdb.
[59]	ML	Bitcoin	Elliptic dataset, a publicly available data set.
[60]	ML	Bitcoin	Elliptic dataset, a publicly available data set.
[61]	General	EOS	PeckShield and bot index.
[63]	Mining/Jacking	General	Alexa top 1M websites.
[64]	Mining/Jacking	General	Previous reports.
[68]	ML	Bitcoin	Online fora.
[69]	General	Bitcoin, ETH	CryptoScamDB and URLScan.io
[70]	Ransom	Bitcoin	Online repositories.
[74]	General	ETH	etherscan, cryptoscamdb, and GitHub.
[16]	DDoS	Bitcoin	Bitcointalk.org
[17]	DDoS	Bitcoin	Mt. Gox exchange leaked data, bitcoincharts.com, and bitcoinity.org
[18]	DDoS	Bitcoin	Bitfinex twitter feed, Bitfinex status page, and Google news search.
[19]	DDoS	Bitcoin	Previous study.

a literature review of the evasion tactics through which criminals can hide their illegal profit using money laundering. The authors have discussed six different evasion techniques, such as mixing services and using privacy coins such as Monero, Zcash, and Dash. These privacy coins were developed with anonymity in mind, implementing obfuscated public ledgers where transaction amounts, destinations, and/or sources are hidden. In [86], the authors have surveyed cryptojacking attacks that target cloud infrastructures by analyzing 11 large scale attacks. They found that most of the attacks have used Monero CPU miners and targeted the Windows platform. They further reported that techniques such as CPU-based classification and blacklisting are ineffective in detecting cryptojacking accurately as attackers use sophisticated obfuscation techniques to hide their activities.

In our study, we have surveyed articles discussing different types of cyberattacks and did not limit our analysis to a specific one. Furthermore, we have answered three different research questions that were not targeted by these studies.

VI. CONCLUSION AND FUTURE WORK

In this systematic literature review, we identified 66 research articles discussing cybercriminal activities that emerged with the introduction of the cryptocurrencies. We analyzed the papers and offered a broad perspective on the activities type, scale, and the proposed detection mechanisms. Our analysis concludes that a significant amount of research has been carried out to detect and analyze these cyberattacks. The research articles have discussed several attacks, including high yield investment programs (HYIP), ransomware, pump and dump, money laundering, and cryptojacking. The cryptocurrencies most frequently studied in the literature are Bitcoin, Ethereum, and Monero. These cyberattacks have stolen millions of USD from thousands of victims. Furthermore, millions of connected devices are abused in cryptojacking attacks. However, even greater losses are caused by ransomware denial of services and productivity losses, which are estimated in billions of USD.

In this paper, we reported the estimation of the scale of these attacks through the lens of technical academic papers only. The reality is likely to be worse, according to CipherTrace cryptocurrency anti-money laundering report.[2] In our future work, we are going to compare the scales reported in scientific papers and those reported by industrial security companies such as CipherTrace.

In the literature, we have found that four different sources have been used to collect training datasets; some are scrapping online fora, some are using data from third parties, some are using free online sources, and finally, some are using usage data of the devices under attack. Many authors have made their dataset publicly available, and we have provided a complete listing of all these datasets.

The defensive mechanisms that have been suggested in the literature relied on training classifiers such as ''random forest'' and ''support vector machine'' on distinguishing features extracted from the dataset. Our review revealed that the proposed defensive mechanisms were quite efficient with 0-day detection of HYIP in Ethereum and cryptojacking, and predicting the likelihood of a cryptocurrency being pumped before the actual pump event. However, although high accuracy late detection of HYIP in Bitcoin have been published, 0-day detection of this attack in Bitcoin is an open problem.

Finally, in our future work, we would like to extend our review to other types of cybercriminal activities such as Wallet and exchange scams [6], [87] and initial coin offering scams [88]. Furthermore, we would like to answer other research questions such as ''What are the limitations of the proposed defensive mechanisms and will they stay efficient if the attackers change their tactics?''.

APPENDIX
BREAKDOWN PER ARTICLE

In this Appendix, we provide a breakdown of the data used to conduct our analysis in Section III, as reported in the literature. We provide the reported cybercrimes scale in Table 7, the resources used to prepare the training datasets in Table 8, and a breakdown of the classifiers used in the papers and the results achieved in Table 6.

ACKNOWLEDGMENT

The authors are thankful to Professor Daniel Amyot for providing his valuable guidance throughout the development of the SLR. They are also thankful to the fellow researchers who reviewed the SLR protocol and the review paper draft.

REFERENCES

[1] J. Kamps and B. Kleinberg, ''To the moon: Defining and detecting cryptocurrency pump-and-dumps,'' *Crime Sci.*, vol. 7, no. 1, p. 18, Dec. 2018.
[2] S. Nakamoto and A. Bitcoin. (2008). *A peer-to-peer electronic cash system*. Bitcoin. [Online]. Available: https://bitcoin.org/bitcoin.pdf
[3] M. Vasek and T. Moore, ''Analyzing the bitcoin ponzi scheme ecosystem,'' in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Berlin, Germany: Springer, 2018, pp. 101–112.
[4] K. Toyoda, P. Takis Mathiopoulos, and T. Ohtsuki, ''A novel methodology for HYIP Operators' bitcoin addresses identification,'' *IEEE Access*, vol. 7, pp. 74835–74848, 2019.
[5] K. Toyoda, T. Ohtsuki, and P. T. Mathiopoulos, ''Multi-class bitcoin-enabled service identification based on transaction history summarization,'' in *Proc. IEEE Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber, Phys. Social Comput. (CPSCom) IEEE Smart Data (SmartData)*, Jul. 2018, pp. 1153–1160.
[6] M. Vasek and T. Moore, ''There's no free lunch, even using bitcoin: Tracking the popularity and profits of virtual currency scams,'' in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Berlin, Germany: Springer, 2015, pp. 44–61.
[7] M. Bartoletti, B. Pes, and S. Serusi, ''Data mining for detecting bitcoin ponzi schemes,'' in *Proc. Crypto Valley Conf. Blockchain Technol. (CVCBT)*, Jun. 2018, pp. 75–84.
[8] C. Brenig, R. Accorsi, and G. Müller, ''Economic analysis of cryptocurrency backed money laundering,'' ECIS, Res. Papers 20, 2015.
[9] M. Moser, R. Bohme, and D. Breuker, ''An inquiry into money laundering tools in the bitcoin ecosystem,'' in *Proc. APWG eCrime Researchers Summit*, Sep. 2013, pp. 1–14.
[10] M. Spagnuolo, F. Maggi, and S. Zanero, ''Bitiodine: Extracting intelligence from the bitcoin network,'' in *Proc. Int. Conf. Financia Cryptogr. Data Secur.* Berlin, Germany: Springer, 2014, pp. 457–468.
[11] K. Liao, Z. Zhao, A. Doupe, and G.-J. Ahn, ''Behind closed doors: Measurement and analysis of CryptoLocker ransoms in bitcoin,'' in *Proc. APWG Symp. Electron. Crime Res. (eCrime)*, Jun. 2016, pp. 1–13.
[12] S. Bistarelli, M. Parroccini, and F. Santini, ''Visualizing bitcoin flows of ransomware: Wannacry one week later,'' in *Proc. ITASEC*, 2018, pp. 1–8.

[13] J. Xu and B. Livshits, "The anatomy of a cryptocurrency pump-and-dump scheme," in *Proc. 28th USENIX Secur. Symp. USENIX Secur.*, 2019, pp. 1609–1625.
[14] W. Chen, Y. Xu, Z. Zheng, Y. Zhou, J. E. Yang, and J. Bian, "Detecting' pump & dump schemes' on cryptocurrency market using an improved apriori algorithm," in *Proc. IEEE Int. Conf. Service-Oriented Syst. Eng. (SOSE)*, Apr. 2019, pp. 293–2935.
[15] F. Victor and T. Hagemann, "Cryptocurrency pump and dump schemes: Quantification and detection," in *Proc. Int. Conf. Data Mining Workshops (ICDMW)*, Nov. 2019, pp. 244–251.
[16] M. Vasek, M. Thornton, and T. Moore, "Empirical analysis of denial-of-service attacks in the bitcoin ecosystem," in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Berlin, Germany: Springer, 2014, pp. 57–71.
[17] A. Feder, N. Gandal, J. T. Hamrick, and T. Moore, "The impact of DDoS and other security shocks on bitcoin currency exchanges: Evidence from Mt. Gox," *J. Cybersecur.*, vol. 3, no. 2, pp. 137–144, Jun. 2017.
[18] A. Abhishta, R. Joosten, S. Dragomiretskiy, and L. J. M. Nieuwenhuis, "Impact of successful DDoS attacks on a major crypto-currency exchange," in *Proc. 27th Euromicro Int. Conf. Parallel, Distrib. Netw.-Based Process. (PDP)*, Feb. 2019, pp. 379–384.
[19] U.-J. Baek, S.-H. Ji, J. T. Park, M.-S. Lee, J.-S. Park, and M.-S. Kim, "DDoS attack detection on bitcoin ecosystem using deep-learning," in *Proc. 20th Asia–Pacific Netw. Oper. Manage. Symp. (APNOMS)*, Sep. 2019, pp. 1–4.
[20] D. A. Wijaya, J. K. Liu, R. Steinfeld, and D. Liu, "Risk of asynchronous protocol update: Attacks to monero protocols," in *Proc. Australas. Conf. Inf. Secur. Privacy*. Cham, Switzerland: Springer, 2019, pp. 307–321.
[21] R. Zheng, C. Ying, J. Shao, G. Wei, H. Yan, J. Kong, Y. Ren, H. Zhang, and W. Hou, "New game-theoretic analysis of ddos attacks against bitcoin mining pools with defence cost," in *Proc. Int. Conf. Netw. Syst. Secur.* Cham, Switzerland: Springer, 2019, pp. 567–580.
[22] S. Wu, Y. Chen, M. Li, X. Luo, Z. Liu, and L. Liu, "Survive and thrive: A stochastic game for DDoS attacks in bitcoin mining pools," *IEEE/ACM Trans. Netw.*, vol. 28, no. 2, pp. 874–887, Apr. 2020.
[23] B. Johnson, A. Laszka, J. Grossklags, M. Vasek, and T. Moore, "Game-theoretic analysis of ddos attacks against bitcoin mining pools," in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Berlin, Germany: Springer, 2014, pp. 72–86.
[24] A. Zimba, Z. Wang, and M. Mulenga, "Cryptojacking injection: A paradigm shift to cryptocurrency-based Web-centric Internet attacks," *J. Organizational Comput. Electron. Commerce*, vol. 29, no. 1, pp. 40–59, Jan. 2019.
[25] M. A. Razali and S. M. Shariff, "Cmblock: In-browser detection and prevention cryptojacking tool using blacklist and behavior-based detection method," in *Proc. Int. Vis. Informat. Conf.* Cham, Switzerland: Springer, 2019, pp. 404–414.
[26] A. Kharraz, Z. Ma, P. Murley, C. Lever, J. Mason, A. Miller, N. Borisov, M. Antonakakis, and M. Bailey, "Outguard: Detecting in-browser covert cryptocurrency mining in the wild," in *Proc. World Wide Web Conf.*, 2019, pp. 840–852.
[27] M. Musch, C. Wressnegger, M. Johns, and K. Rieck, "Thieves in the browser: Web-based cryptojacking in the wild," in *Proc. 14th Int. Conf. Availability, Rel. Secur. ARES*, 2019, pp. 1–10.
[28] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," EBSE, Goyang-si, South Korea, Tech. Rep., Ver. 2.3 EBSE, 2007.
[29] S. Jalali and C. Wohlin, "Systematic literature studies: Database searches vs. Backward snowballing," in *Proc. ACM-IEEE Int. Symp. Empirical Softw. Eng. Meas. ESEM*, 2012, pp. 29–38.
[30] D. Bryans, "Bitcoin and money laundering: Mining for an effective solution," *Indiana Law J.*, vol. 89, p. 441, 2014.
[31] D. M. Sat, A. B. Kasatkin, I. A. Kornev, G. O. Krylov, and K. Evgenyevich, "Investigation of money laundering methods through cryptocurrency," *J. Theor. Appl. Inf. Technol.*, vol. 83, no. 2, pp. 244–254, 2016.
[32] H. Sun Yin and R. Vatrapu, "A first estimation of the proportion of cybercriminal entities in the bitcoin ecosystem using supervised machine learning," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2017, pp. 3690–3699.
[33] K. Toyoda, T. Ohtsuki, and P. T. Mathiopoulos, "Identification of high yielding investment programs in bitcoin via transactions pattern analysis," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.
[34] R. Tahir, M. Huzaifa, A. Das, M. Ahmad, C. Gunter, F. Zaffar, M. Caesar, and N. Borisov, "Mining on someone else's dime: Mitigating covert mining operations in clouds and enterprises," in *Proc. Int. Symp. Res. Attacks, Intrusions, Defenses*. Cham, Switzerland: Springer, 2017, pp. 287–310.
[35] N. Kshetri and J. Voas, "Do crypto-currencies fuel ransomware?" *IT Prof.*, vol. 19, no. 5, pp. 11–15, 2017.
[36] D. Y. Huang, M. M. Aliapoulios, V. G. Li, L. Invernizzi, E. Bursztein, K. McRoberts, J. Levin, K. Levchenko, A. C. Snoeren, and D. McCoy, "Tracking ransomware end-to-end," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2018, pp. 618–631.
[37] S. Mabunda, "Cryptocurrency: The new face of cyber money laundering," in *Proc. Int. Conf. Adv. Big Data, Comput. Data Commun. Syst. (icABCD)*, Aug. 2018, pp. 1–6.
[38] J. Seo, M. Park, H. Oh, and K. Lee, "Money laundering in the bitcoin network: Perspective of mixing services," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2018, pp. 1403–1405.
[39] A. Holub and J. O'Connor, "COINHOARDER: Tracking a ukrainian bitcoin phishing ring DNS style," in *Proc. APWG Symp. Electron. Crime Res. (eCrime)*, May 2018, pp. 1–5.
[40] G. Hong, Z. Yang, S. Yang, L. Zhang, Y. Nan, Z. Zhang, M. Yang, Y. Zhang, Z. Qian, and H. Duan, "How you get shot in the back: A systematical study about cryptojacking in the real world," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Jan. 2018, pp. 1701–1713.
[41] W. Chen, Z. Zheng, J. Cui, E. Ngai, P. Zheng, and Y. Zhou, "Detecting ponzi schemes on ethereum: Towards healthier blockchain technology," in *Proc. World Wide Web Conf. World Wide Web - WWW*, 2018, pp. 1409–1418.
[42] M. Campbell-Verduyn, "Bitcoin, crypto-coins, and global anti-money laundering governance," *Crime, Law Social Change*, vol. 69, no. 2, pp. 283–305, Mar. 2018.
[43] M. Conti, A. Gangwal, and S. Ruj, "On the economic significance of ransomware campaigns: A bitcoin transactions perspective," *Comput. Secur.*, vol. 79, pp. 162–189, Nov. 2018.
[44] R. K. Konoth, E. Vineti, V. Moonsamy, M. Lindorfer, C. Kruegel, H. Bos, and G. Vigna, "MineSweeper: An in-depth look into drive-by cryptocurrency mining and its defense," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Jan. 2018, pp. 1714–1730.
[45] K. Toyoda, T. Ohtsuki, and P. Mathiopoulos, "Time series analysis for bitcoin transactions: The case of pirate@ 40's hyip scheme," in *Proc. IEEE Int. Conf. Data Mining Workshops (ICDMW)*, Nov. 2018, pp. 151–155.
[46] A. A. Maksutov, M. S. Alexeev, N. O. Fedorova, and D. A. Andreev, "Detection of blockchain transactions used in blockchain mixer of coin join type," in *Proc. IEEE Conf. Russian Young Researchers Electr. Electron. Eng. (EIConRus)*, Jan. 2019, pp. 274–277.
[47] E. Jung, M. Le Tilly, A. Gehani, and Y. Ge, "Data mining-based ethereum fraud detection," in *Proc. IEEE Int. Conf. Blockchain (Blockchain)*, Jul. 2019, pp. 266–273.
[48] W. Chen, Z. Zheng, E. C.-H. Ngai, P. Zheng, and Y. Zhou, "Exploiting blockchain data to detect smart ponzi schemes on ethereum," *IEEE Access*, vol. 7, pp. 37575–37586, 2019.
[49] A. Gangwal and M. Conti, "Cryptomining cannot change its spots: Detecting covert cryptomining using magnetic side-channel," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1630–1639, 2020.
[50] C. F. Torres and M. Steichen, "The art of the scam: Demystifying honeypots in ethereum smart contracts," in *Proc. 28th USENIX Secur. Symp. USENIX Secur.*, 2019, pp. 1591–1607.
[51] M. Ostapowicz and K. Z. bikowski, "Detecting fraudulent accounts on blockchain: A supervised approach," in *Proc. Int. Conf. Web Inf. Syst. Eng.* Cham, Switzerland: Springer, 2019, pp. 18–31.
[52] O. Starov, Y. Zhou, and J. Wang, "Detecting malicious campaigns in obfuscated JavaScript with scalable behavioral analysis," in *Proc. IEEE Secur. Privacy Workshops (SPW)*, May 2019, pp. 218–223.
[53] C. Albrecht, K. M. Duffin, S. Hawkins, and V. M. Morales Rocha, "The use of cryptocurrencies in the money laundering process," *J. Money Laundering Control*, vol. 22, no. 2, pp. 210–216, May 2019.
[54] R. Ning, C. Wang, C. Xin, J. Li, L. Zhu, and H. Wu, "CapJack: Capture in-browser crypto-jacking by deep capsule network through behavioral analysis," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Apr. 2019, pp. 1873–1881.
[55] A. Zimba, M. Chishimba, C. Ngongola-Reinke, and T. F. Mbale, "Demystifying cryptocurrency mining attacks: A semi-supervised learning approach based on digital forensics and dynamic network characteristics," in *Proc. 3rd IEEE Int. Conf. IN ICTs (ICICT)*, 2019, pp. 1–6.
[56] M. Bartoletti, S. Carta, T. Cimoli, and R. Saia, "Dissecting ponzi schemes on ethereum: Identification, analysis, and impact," *Future Gener. Comput. Syst.*, vol. 102, pp. 259–277, Jan. 2020.

[57] L. Haffke, M. Fromberger, and P. Zimmermann, "Cryptocurrencies and anti-money laundering: The shortcomings of the fifth aml directive (eu) and how to address them," *J. Banking Regulation*, vol. 21, pp. 125–138, Apr. 2019.

[58] S. Farrugia, J. Ellul, and G. Azzopardi, "Detection of illicit accounts over the ethereum blockchain," *Expert Syst. Appl.*, vol. 150, Jul. 2020, Art. no. 113318.

[59] I. Alarab, S. Prakoonwit, and M. I. Nacer, "Competence of graph convolutional networks for anti-money laundering in bitcoin blockchain," in *Proc. 5th Int. Conf. Mach. Learn. Technol.*, Jun. 2020, pp. 23–27.

[60] I. Alarab, S. Prakoonwit, and M. I. Nacer, "Comparative analysis using supervised learning methods for anti-money laundering in bitcoin," in *Proc. 5th Int. Conf. Mach. Learn. Technol.*, Jun. 2020, pp. 11–17.

[61] Y. Huang, H. Wang, L. Wu, G. Tyson, X. Luo, R. Zhang, X. Liu, G. Huang, and X. Jiang, "Understanding (Mis) behavior on the EOSIO blockchain," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 4, no. 2, pp. 1–28, 2020.

[62] B. Custers, J.-J. Oerlemans, and R. Pool, "Laundering the profits of ransomware: Money laundering methods for vouchers and cryptocurrencies," *Eur. J. Crime, Criminal Law Criminal Justice*, vol. 28, no. 2, pp. 121–152, Jul. 2020.

[63] W. Bian, W. Meng, and M. Zhang, "MineThrottle: Defending against wasm in-browser cryptojacking," in *Proc. Web Conf.*, Apr. 2020, pp. 3112–3118.

[64] D. Tanana, "Behavior-based detection of cryptojacking malware," in *Proc. Ural Symp. Biomed. Eng., Radioelectronics Inf. Technol. (USBEREIT)*, May 2020, pp. 0543–0545.

[65] O. Delgado-Mohatar, J. M. Sierra-Cámara, and E. Anguiano, "Blockchain-based semi-autonomous ransomware," *Future Gener. Comput. Syst.*, vol. 112, pp. 589–603, Nov. 2020.

[66] C. F. Torres, M. Baden, and R. State, "Towards usable protection against honeypots," in *Proc. IEEE Int. Conf. Blockchain Cryptocurrency (ICBC)*, May 2020, pp. 1–2.

[67] C. Karapapas, I. Pittaras, N. Fotiou, and G. C. Polyzos, "Ransomware as a service using smart contracts and IPFS," 2020, *arXiv:2003.04426*. [Online]. Available: http://arxiv.org/abs/2003.04426

[68] J. Crawford and Y. Guan, "Knowing your bitcoin customer: Money laundering in the bitcoin economy," in *Proc. 13th Int. Conf. Systematic Approaches Digit. Forensic Eng. (SADFE)*, May 2020, pp. 38–45.

[69] R. Phillips and H. Wilder, "Tracing cryptocurrency scams: Clustering replicated advance-fee and phishing websites," 2020, *arXiv:2005.14440*. [Online]. Available: http://arxiv.org/abs/2005.14440

[70] P. S. Goyal, A. Kakkar, G. Vinod, and G. Joseph, "Crypto-ransomware detection using behavioural analysis," in *Reliability, Safety and Hazard Assessment for Risk-Based Technologies*. Singapore: Springer, 2020, pp. 239–251.

[71] K. Weber, A. E. Schütz, T. Fertig, and N. H. Müller, "Exploiting the human factor: Social engineering attacks on cryptocurrency users," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2020, pp. 650–668.

[72] F. Teichmann and M.-C. Falker, "Money laundering through cryptocurrencies," in *Proc. 13th Int. Sci. Practical Conf.-Artif. Intell. Anthropogenic Nature Vs. Social Origin*. Springer, 2020, pp. 500–511.

[73] N. Pocher, "The open legal challenges of pursuing aml/cft accountability within privacy-enhanced iom ecosystems," in *Proc. DLT@ ITASEC*, 2020, pp. 1–15.

[74] N. Kumar, A. Singh, A. Handa, and S. K. Shukla, "Detecting malicious accounts on the ethereum blockchain with supervised learning," in *Proc. Int. Symp. Cyber Secur. Cryptogr. Mach. Learn.* Cham, Switzerland: Springer, 2020, pp. 94–109.

[75] F. Teichmann and M.-C. Falker, "Blockchain: Implications of the impending token economy," in *Proc. Inst. Sci. Commun. Conf.* Cham, Switzerland: Springer, 2019, pp. 1551–1565.

[76] M. A. Pitman. (1998). *Qualitative Research Design: An Interactive Approach*. [Online]. Available: https://anthrosource.onlinelibrary.wiley.com/doi/abs/10.1525/aeq.1998.29.4.499

[77] J. Brings, M. Daun, M. Kempe, and T. Weyer, "On different search methods for systematic literature reviews and maps: Experiences from a literature search on validation and verification of emergent behavior," in *Proc. 22nd Int. Conf. Eval. Assessment Softw. Eng. - EASE*, 2018, pp. 35–45.

[78] S. Amanzholova, N. Tastan, K. Kalkamanova, and A. Yessenalina, "Valid and invalid bitcoin transactions," in *Proc. 6th Int. Conf. Eng. MIS*, Sep. 2020, pp. 1–5.

[79] S. Kethineni and Y. Cao, "The rise in popularity of cryptocurrency and associated criminal activity," *Int. Criminal Justice Rev.*, vol. 30, no. 3, pp. 325–344, Sep. 2020.

[80] M. Saad, J. Spaulding, L. Njilla, C. Kamhoua, S. Shetty, D. Nyang, and A. Mohaisen, "Exploring the attack surface of blockchain: A systematic overview," 2019, *arXiv:1904.03487*. [Online]. Available: http://arxiv.org/abs/1904.03487

[81] L. Phan, S. Li, and K. Mentzer, "Blockchain technology and the current discussion on fraud," *Comput. Inf. Syst. J. Articles*, 2019, Paper 28. [Online]. Available: https://digitalcommons.bryant.edu/cisjou/28/

[82] F. M. J. Teichmann and M.-C. Falker, "Cryptocurrencies and financial crime: Solutions from liechtenstein," *J. Money Laundering Control*, Jun. 2020, doi: 10.1108/JMLC-05-2020-0060.

[83] H. Chen, M. Pendleton, L. Njilla, and S. Xu, "A survey on ethereum systems security: Vulnerabilities, attacks, and defenses," *ACM Comput. Surv.*, vol. 53, no. 3, pp. 1–43, 2020.

[84] D. B. Desmond, D. Lacey, and P. Salmon, "Evaluating cryptocurrency laundering as a complex socio-technical system: A systematic literature review," *J. Money Laundering Control*, vol. 22, no. 3, pp. 480–497, Jul. 2019.

[85] D. Dupuis and K. Gleason, "Money laundering with cryptocurrency: Open doors and the regulatory dialectic," *J. Financial Crime*, Aug. 2020.

[86] K. Jayasinghe and G. Poravi, "A survey of attack instances of cryptojacking targeting cloud infrastructure," in *Proc. 2nd Asia Pacific Inf. Technol. Conf.*, Jan. 2020, pp. 100–107.

[87] P. Xia, H. Wang, B. Zhang, R. Ji, B. Gao, L. Wu, X. Luo, and G. Xu, "Characterizing cryptocurrency exchange scams," *Comput. Secur.*, vol. 98, Nov. 2020, Art. no. 101993.

[88] D. A. Zetzsche, R. P. Buckley, D. W. Arner, and L. Fohr, "The ico gold rush: It's a scam, it's a bubble, it's a super challenge for regulators," *Harv. Int. LJ*, vol. 60, p. 267, Jul. 2019.

EMAD BADAWI received the bachelor's degree in computer systems engineering from the Arab American University (AAUP), and the master's degree in computer engineering from the American University of Sharjah (AUS), UAE, in 2017. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Ottawa.

His research thesis was about parallel implementations for finite state machines mutants' elimination algorithms, based on OpenMP, MPI, and CUDA. His current research is about software security, in particular, cybercrime detection and prevention.

GUY-VINCENT JOURDAN received the Ph.D. degree in the area of distributed systems analysis from l'université de Rennes/INRIA, France, in 1995. He joined the School of Electrical Engineering and Computer Science, University of Ottawa, as an Associate Professor in June 2004, after seven years of experience in the private sector as the C.T.O. and then the C.E.O. of Ottawa-based Decision Academic Graphics. He is currently a Full Professor with the Faculty of Engineering, University of Ottawa. His research interests include distributed systems modeling and analysis, software security, as well as cybercrime detection and prevention.