Comparing Bitcoin generators on the clear web and the dark web

Hartel, Pieter; Junger, Marianne; Staalduinen, Mark van

DOI
10.1186/s40163-025-00246-w

Publication date
2025

Document Version
Final published version

Published in
Crime Science

Citation (APA)
Hartel, P., Junger, M., & Staalduinen, M. V. (2025). Comparing Bitcoin generators on the clear web and the dark web. *Crime Science*, *14*(1), Article 4. https://doi.org/10.1186/s40163-025-00246-w

Crime Science

RESEARCH Open Access

Comparing Bitcoin generators on the clear web and the dark web

Pieter Hartel[1,3*], Marianne Junger[2†] and Mark van Staalduinen[3†]

Abstract

Objective This study examines Bitcoin generator (BG) websites on the clear and dark web. It focuses on their prevalence, revenue, and associated warnings, as these sites are suspected scams.

Method Data for the study was gathered from the Dark Web Monitor and Iknaio Cryptoasset Analytics. A four-step process was used to identify BG sites and their Bitcoin addresses from 2 million dark websites.

Results We found 832 dark web BG sites. The monetary revenue from a dark web BG site is approximately 1/3 smaller per Bitcoin address than from a clear web BG site. There is a concentration of revenue at a few BG sites. Only 24% of Bitcoin addresses on dark web BG sites have ever had money deposited on them. On the dark web, the top three clusters of crypto addresses account for 35% of the total revenue. On the clear web, the top three clusters account for 52% of the total revenue. The longer BG sites are online, the higher the revenue. There are hardly any warnings against BG sites.

Conclusion Our results fit the Rational Choice model of crime: the revenue is modest, but the effort of the offenders is also limited.

Keywords Bitcoin generator, Dark web, Cybercrime, Rational choice model, Online fraud

Introduction

Cryptocurrency scams exploit the complexity of blockchain technology to deceive victims, with the Bitcoin generator (BG) scam being a notable example. While prior research has examined such scams on the clear web, little is known about their presence on the dark web. This study fills that gap by conducting an exploratory analysis of BG scams on the dark web and comparing them to their clear web counterparts. The findings reveal that, although these scams require minimal effort to set up and pose lower risks to scammers on the dark web, their profitability is also limited. By analysing scammers' operational aspects, risks, and financial expectations, this research provides insights into the broader phenomenon of technically oriented fraud. It also raises critical questions about the evolution of these scams, the role of psychological traits in victim susceptibility, and the effectiveness of fraud prevention strategies. Ultimately, this study contributes to the growing body of research on online fraud.

The clear and dark web are contrasting environments, each with unique risk and reward structures that influence cybercriminal strategies. The clear web is generally easy to access, with authorities actively monitoring illegal activity. As a result, scammers are at greater risk of attracting the attention of law enforcement, although they also benefit from a larger audience and easier accessibility for users. Typical scams on the clear web

[†] Marianne Junger and Mark van Staalduinen have contributed equally to this work.

*Correspondence:
Pieter Hartel
pieter.hartel@tudelft.nl; pieter.hartel@cflw.com
[1] Delft University of Technology, Mekelweg 5, 2628 CD Delft, Netherlands
[2] University of Twente, Drienerlolaan 5, 7522 NB Enschede, Netherlands
[3] CFLW Cyber Strategies, Franklinstraat 1A, 2691 HB 's-Gravenzande, Netherlands

include phishing attacks and counterfeit sales (Reep-van den Bergh & Junger, 2018).

The dark web, by contrast, provides enhanced anonymity for scammers, facilitating activities such as operating illegal drug markets, trading stolen data, and distributing malware. Because the dark web exists thanks to privacy-protecting tools, law enforcement faces significant challenges in tracing activities to individuals (Winter et al., 2018).

The clear and dark web offers different advantages and challenges when running scam websites. The clear web offers broad reach and easy access but with higher levels of control. Conversely, the dark web offers privacy but at the cost of less trust from users who are likely to know that they are in a risky situation. The dark web thus offers a smaller and more "exclusive" reach to potential users. These advantages and disadvantages have been explored previously about account credentials (Villalva et al., 2018), and for counterfeit identity documents (Holt and Lee, 2022), but not yet regarding Bitcoin generators. In this article, we will focus on Bitcoin generators.

In the next section, we discuss the background and the research questions. Section "Method" explains how we collected and analysed the necessary data. Section "Results" presents the results of the data analysis. Section "Discussion" discusses the answers to the research questions. The last two sections present the limitations and the conclusions.

Background

Cryptocurrencies, such as Bitcoin, play a significant role in financial crime (Trozze et al., 2022). The current study focuses on a specific form of this: the "Bitcoin generator". Previously, Badawi et al. (2022) investigated the Bitcoin generator on the clear web. We compare our results regarding the Bitcoin generator on the dark web with those of Badawi et al. (2022) on the clear web. For a more detailed description of Bitcoin generators on the clear web, we refer the reader to Badawi et al. (2022). For a fair comparison, we focus on Bitcoin, like Badawi et al. (2022), and ignore other cryptocurrencies.

We define a Bitcoin Generator (BG) as a website that claims to earn Bitcoins using "clever technical tricks" such as:

- Exploiting flaws in the Bitcoin protocol.
- The availability of a powerful mining machine.
- The use of smart algorithms for buying and selling.

These "clever tricks" are designed to impress visitors to BG sites.

Victims assume scammers have advanced technical skills they do not understand, and victims tend to trust those they perceive as knowledgeable.

In exchange for benefitting from these "smart technologies", the visitor must pay a relatively small "mining fee". Scammers often claim that they want others to benefit from their technical expertise. They argue that spending the Bitcoins they earned would expose them to risk. The clever tricks may work well now, but at some point, they may not. Therefore, the users must quickly take advantage of the supply.

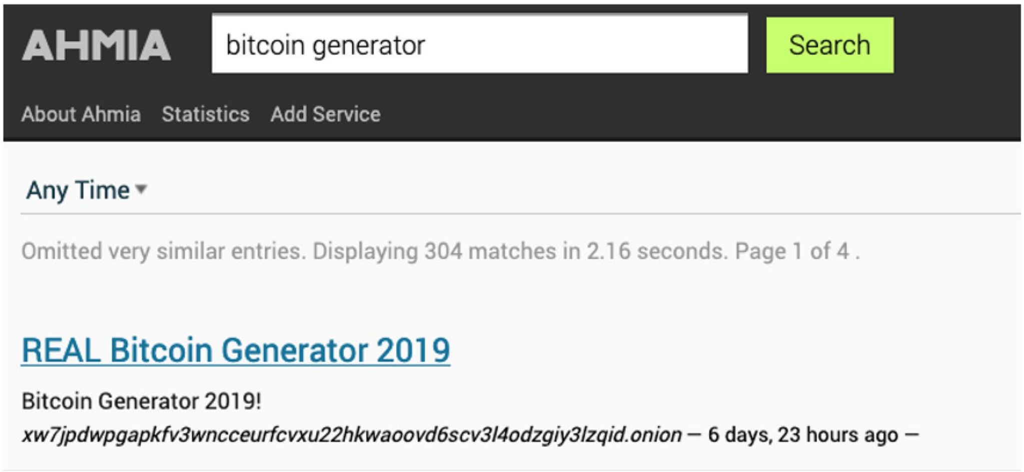
The scammers apply several marketing principles (Cialdini, 2009), such as convincing potential victims that the BG site has been built by experts (authority), encouraging people to act quickly (scarcity), and letting them benefit (liking). Perhaps the "clever trick" is not entirely legitimate, but it is all the more attractive for quick profit (Stajano and Wilson, 2011).

Running a BG site that promises to multiply users' cryptocurrency in exchange for a mining fee is generally illegal. In the United States, individuals who conduct this fraud are often prosecuted under the Wire Fraud, Securities Fraud, and Money Laundering laws. The owners of publicly available BG sites are, therefore, scammers. We are unaware of anyone ever being arrested or convicted for a BG scam. There have been lawsuits and even a conviction over the related Bitcoin Giveaway scam (Vakilinia, 2022). The Giveaway scam does not use a clever technical trick and is, therefore, outside the scope of this investigation.

The anonymity of the dark web is being abused to offer illegal products and services with a relatively small chance of being caught (Lee et al., 2019; Xia et al., 2024). Law enforcement can break the anonymity if a dark web user makes mistakes. A well-known example is Ross Ulbricht, the SilkRoad dark web marketplace owner. He inadvertently included his Gmail address in an ad recruiting engineers for a new project called 'SilkRoad" (Nurmi & Niemel, 2017).

Google and other search engines do not index the dark web. Several search and indexing platforms on the dark web enable web admins to list and promote their site's name for discovery. A common index site is ahmia.fi, see Fig. 1. Finding BG sites on the dark web is challenging without listings on index sites because dark web domain names are random characters.

We investigated the costs of setting up a BG site in a preliminary study. We rented a small Linux server from two European hosting providers for 1 Euro per month each and also bought a Raspberry PI. Then, using HTTrack Website Copier, we copied twenty existing BG sites from the dark web and installed them on the three servers. We terminated the lease of one of the rented servers after a few months and



Omitted very similar entries. Displaying 304 matches in 2.16 seconds. Page 1 of 4 .

REAL Bitcoin Generator 2019

Bitcoin Generator 2019!

xw7jpdwpgapkfv3wncceurfcvxu22hkwaoovd6scv3l4odzgiy3lzqid.onion — 6 days, 23 hours ago —

Fig. 1 Example of a BG site listed on ahmia.fi

kept the other and the Raspberry PI running for a whole year. We kept the domain names secret to prevent third parties from visiting our BG sites. The server logs showed that only we visited the BG sites.

Setting up these 3 times 20 sites took only a few hours and cost about 100 Euros.

BG scams on the clear and dark web are very similar because it is so easy to clone a BG site.

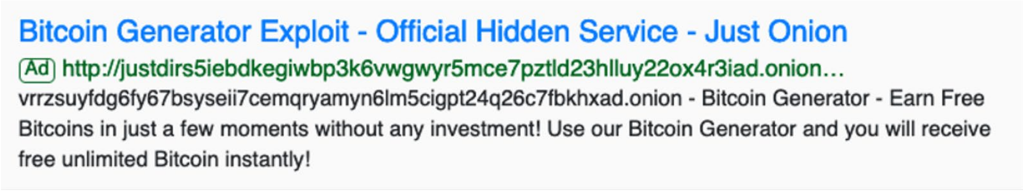
BG sites mainly trade in Bitcoins (Badawi et al., 2022), so both visitors and scammers must be able to handle Bitcoins. The victim can usually choose the mining fee.

Installing a BG site on the clear web requires two additional steps: applying for a domain name and a TLS certificate. However, the associated costs are low, sometimes even zero. Copycats can easily create BG websites on the dark and clear web by copying and pasting. The only modification required is changing the Bitcoin address. The effort necessary to set up a BG site is limited.

The risks and revenue associated with a BG site, whether operating on the clear or dark web, warrant deeper investigation. Therefore, the research question is: *How do the risks and revenue of a BG website differ on the clear and dark web?* To answer this central question, we have formulated five sub-questions. We focus on the important issues to scammers: the effort they put into launching and maintaining a website, the revenue they expect and the risks they may fear (Cornish & Clarke, 2008).

In order to answer the research question, we study five sub-questions as follows:

1. *To what extent are BG sites promoted on the dark web?* Some search engines and indexing websites, like Google Ads, allow advertisements for dark web products in exchange for a fee. Figure 2 shows an example of this. Advertising a BG site can increase the number of paying visitors. However, it also risks drawing attention from competitors, hosting providers, or law enforcement agencies (Ogbanufe et al., 2024). Visitors to a dark website must install the TOR browser, which creates a barrier and limits the number of potential visitors. The TOR network obfuscates the identity of web admins, significantly diminishing the likelihood of detection by law enforcement agencies on the dark web.
2. *How does BG site revenue differ on the clear and dark web?* According to data from TOR-metrics, there are approximately 800,000 active TOR services in use (including websites). In contrast, the number of services (including websites) added to the Internet every day is over 10 million (see Certstream Firehose). The dark web is much smaller than the clear web, so BG



Bitcoin Generator Exploit - Official Hidden Service - Just Onion

Ad http://justdirs5iebdkegiwbp3k6vwgwyr5mce7pztld23hlluy22ox4r3iad.onion...

vrrzsuyfdg6fy67bsyseii7cemqryamyn6lm5cigpt24q26c7fbkhxad.onion - Bitcoin Generator - Earn Free Bitcoins in just a few moments without any investment! Use our Bitcoin Generator and you will receive free unlimited Bitcoin instantly!

Fig. 2 Example of an advertisement for "Bitcoin generators" on the dark web from onionlandsearchengine.net

sites on the dark web likely generate less revenue than those on the clear web.

3. *To what extent are BG campaigns being conducted on the dark web?* Phishing is known to have multiple campaigns in which similar phishing emails are sent to large numbers of email addresses (Oest et al., 2020). BG sites are often, as is the case with phishing emails, near-identical copies of each other (Badawi et al., 2022). The longer a campaign runs, the greater the potential for increased visitor numbers and, consequently, higher revenue for the scammer. At the same time, a longer campaign also increases the chance that adversaries, such as law enforcement agencies, will take action against the scammer. We can only investigate campaigns on the dark web because Badawi et al. (2022) have not made the necessary clear web data available.
4. *To what extent is BG site revenue concentrated on the clear and dark web?* BG scammers sometimes collaborate or copy each other's work (Badawi et al., 2022). This type of collaboration –whether voluntary or involuntary– can be demonstrated by shared Bitcoin addresses. When two campaigns use the same Bitcoin addresses, this is a form of collaboration.
5. *How do warnings against BG scams differ on the clear and dark web?* Several organisations such as the Anti-Phishing Working Group (APWG) warn Internet users about fraud. It is up to future research to determine how effective these warnings are. As a first step, however, we want to determine whether warnings against BG scams exist. It is unclear whether organisations like APWG handle warnings about BG scams on the dark web or focus solely on clear web scams.

In the Results and Discussion sections, we will analyse and discuss these five points in the same order.

Method

The data for the study comes from the Dark Web Monitor (DWM) (Spitters et al., 2014) and Iknaio Cryptoasset Analytics (ICA) (Fröwis et al., 2020).

The Dark Web Monitor (DWM) employs a snowballing approach, a widely used web crawling technique, to collect and organise data from the dark web. This process starts with a curated list of seed addresses sourced from platforms like Ahmia.fi. The crawler iteratively retrieves and parses webpages using these initial seeds, extracting new URLs to expand its network coverage. This approach is similar to techniques used for indexing the clear web by companies such as Google.

Since its inception in 2013, DWM has built a vast dataset, capturing 16 million domains and a billion pages over 11 years. This data originates from the TOR network and other anonymised networks like I2P. DWM downloads dark web content every 18 h while sites are online to ensure complete coverage, retrying after 10 days if a site is offline. Approximately one million dark web pages are added to the dataset daily. This extensive dataset supports thematic and structural analyses, offering valuable insights to law enforcement and research into the organisation and content of the dark web (Spitters et al., 2014).

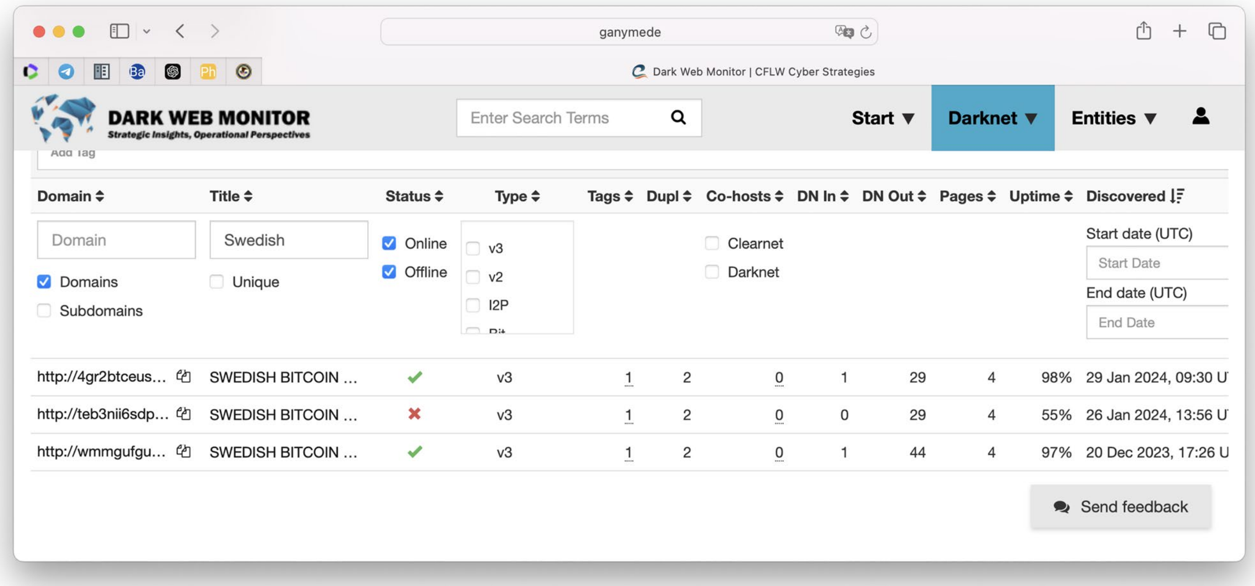
To provide an example of a DWM insight, we use our preliminary study. Figure 3 presents a screenshot depicting the status of three cloned "Swedish Bitcoin Multiplier" sites as of November 20, 2024. The first site has been running on a rented server since January 29, 2024 (as shown in the "Discovered at" column). The second site, hosted on another leased server, became unavailable mid-year when we terminated the lease. The third site has been running on a Raspberry Pi since December 20, 2023. The "uptime" column indicates the frequency at which DWM successfully accessed the servers and retrieved pages from these sites. The second site has an uptime of 55%, reflecting its availability only during the first half of the year. In contrast, the other two sites show nearly 100% uptime, having been accessed approximately 400 times since their discovery. However, neither of them achieved perfect uptime due to occasional server maintenance.

Unfortunately, there is a difference in the period of our main study: November 2019–March 2021 for the clear web versus April 2022–March 2024 for the dark web. This is because we use historical data from the DWM, which is now more complete than before.

Iknaio Cryptoasset Analytics (ICA) systematically ingests data from public blockchains of various cryptocurrencies and enhances it with attribution data from both public and non-public sources. For instance, if a cryptocurrency address appears in a dataset associated with the dark web (e.g., DWM dataset), ICA labels that address as "dark web".

ICA converts cryptocurrency transactions into their fiat currency equivalents by applying the average exchange rate from the transaction date. This time-specific conversion ensures accurate financial analysis in historical contexts.

To group related addresses, ICA employs clustering heuristics, such as the multi-input heuristic, which identifies addresses likely controlled by the same entity based on shared private keys in multi-input transactions. ICA tries to avoid false positives, such as those arising from CoinJoin transactions, where multiple users intentionally mix their inputs to obscure transaction trails. Clusters of



Domain: Domain
☑ Domains
□ Subdomains

Title: Swedish
□ Unique

☑ Online
☑ Offline

□ v3
□ v2
□ I2P

□ Clearnet
□ Darknet

Start date (UTC): Start Date
End date (UTC): End Date

Domain	Title	Status	Type	Tags	Dupl	Co-hosts	DN In	DN Out	Pages	Uptime	Discovered
http://4gr2btceus…	SWEDISH BITCOIN …	✓	v3	1	2	0	1	29	4	98%	29 Jan 2024, 09:30 U
http://teb3nii6sdp…	SWEDISH BITCOIN …	✗	v3	1	2	0	0	29	4	55%	26 Jan 2024, 13:56 U
http://wmmgufgu…	SWEDISH BITCOIN …	✓	v3	1	2	0	1	44	4	97%	20 Dec 2023, 17:26 U

Fig. 3 Screen shot of the Dark Web Monitor showing the uptime and the discovery date of the three clones of the Swedish Bitcoin Multiplier sites

addresses identified by ICA indicate control by the same individual or group, making it a valuable tool for attributing cryptocurrency activity to specific entities (Fröwis et al., 2020).

Only a small fraction of all sites in the DWM dataset are BG sites. We, therefore, use the following 4-step process to reduce the DWM dataset to a list of Bitcoin addresses where a "mining fee" can be deposited as intended by the owner of a BG site.

1. Searching for advertisements of dark web BG sites.
2. Searching for revenue of dark web BG sites.
3. Selecting Bitcoin addresses that can be considered as deposit addresses.
4. Looking up the transactions of the deposit addresses with ICA.

In the remainder of the Method section, we provide the details of the four steps.

In this study, we conduct secondary analyses on existing data from the DWM and ICA datasets. We use the Menlo Report (Bailey et al., 2012) as guidance in analysing the ethical risks and mitigations posed by our research (see Appendix A).

Searching for advertisements of dark web BG sites

There are two main ways to find advertisements for BG sites on the dark web. The first involves identifying text passages as advertisements and then analysing them to determine which website they reference. The second approach, which we have chosen, begins by identifying the names of BG sites and then searching the dark and clear web for those specific websites. Therefore, our process for locating advertisements starts with selecting recent BG sites.

One of the decisions in the study is how to search for BG sites. Badawi et al. (2022) use a list of approximately 700 search terms for BG sites on the clear web. We have updated that list by replacing specific years, such as 2019, with years 2019, 2021, ... 2024. We have also extended the list with 60 (approximately 10%) new search terms. Our list of search terms is available in the (online) Supplementary file of this article.

On the dark web, domain names usually contain random strings, offering no insight into the site's purpose. To address this, web admins typically design the landing page title to inform visitors about the site's content. Leveraging Badawi's queries and our expanded set of search terms, we focused on analysing titles.

Using the updated terms, we searched the DWM dataset by title and identified 356 domain names associated with recent dark web BG sites discovered within approximately one year (April 1, 2023, to March 10, 2024). Next, we searched for advertisements for these 356 domain names on both the clear web (manually, via Google) and the dark web (automatically, via the DWM). Since Google does not permit automated searches, we manually checked each domain name for advertisements on the clear web. To minimise the amount of manual work, we limited this search to one year.

Searching for revenue of dark web BG sites

The DWM allows automated searches, enabling us to search over a longer observation period.

We applied the updated and extended list of search terms to the title of all available pages from the 2 million sites discovered in approximately two years (April 5, 2022, to March 10, 2024). We found 2,363 sites (0.12%) where one or more pages have a title that fits one of the search terms.

Not all sites we found with the updated and extended list of search terms meet our definition of a BG site. Some sites have different goals, such as:

- The Bitcoin Giveaway, where Bitcoins are given away in exchange for an advance fee (Phillips & Wilder, 2020).
- Ponzi fraud with Bitcoin (Vasek & Moore, 2019).
- Distributing Child Sexual Abuse Material (CSAM) (van der Bruggen and Blokland, 2021).

The sites that did not meet our definition of a BG site were manually removed from the dataset. Badawi et al. (2022) used machine learning techniques to distinguish BG sites from non-BG sites. We have made this selection by hand because it is more precise and does not involve large numbers. After manual removal, 832 dark web BG sites remained, meaning that 0.04% of the dark web, according to our analysis, consists of BG sites. Table 1 lists the top five titles of the BG sites that meet our definition. For example, the first line in the table indicates that there are 493 BG sites with the title "Bitcoin Generator Exploit - Official Hidden Service". The 832 BG sites consist of 48,219 pages. Most of these pages are a snapshot of the BG site's appearance at a particular time.

Recognising Bitcoin addresses that can be considered as deposit addresses

We split all 48,219 BG pages around the Bitcoin addresses (recognised by regular expressions) into text passages, each containing a Bitcoin address.

Suppose a page contains the text "Lorem ipsum X dolor sit amet, Y consectetur adipiscing elit", where X and Y are Bitcoin addresses. The first passage of text is "Lorem ipsum X dolor sit amet," and the second is "dolor sit amet, Y consectetur adipiscing elit." Because most pages contain many Bitcoin addresses, the pages are split into 298,387 text passages. If a page does not contain a Bitcoin address, the text passage is the entire page, and the Bitcoin address is NA.

Most text passages of BG sites contain a Bitcoin address, making it easy for a human to see whether or not this is the Bitcoin address where the site owner expects visitors to pay. We call such an address a "deposit address". For example, in the snippet: "To ensure your transaction confirms consistently and reliably, pay the miners fee 0.0025 BTC for this transaction at bc1qh ...t2kfl" the Bitcoin address "bc1qh ...t2kfl" is a deposit address. (For privacy reasons, each Bitcoin address is abbreviated).

Many text passages have Bitcoin addresses, but most Bitcoin addresses come from the blockchain.

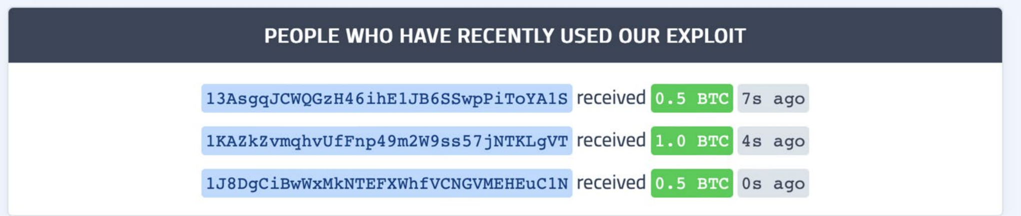
The BG site owner has used these addresses as "bait" to show how much money can be made. See Fig. 4 for an example.

We found only 117 unique text passages with a deposit address. We created a search term for each unique text passage and collected 1069 unique deposit addresses.

It is possible that the 117 manual search terms do not recognise every deposit address. To validate those search terms, we submitted a random sample of 188 different text passages with a deposit address and an equally large random sample of different text passages with a non-deposit address to chatGPT 3.5 turbo. Hence, we use chatGPT as an alternative coder of the deposit addresses (Xiao et al., 2023). We have presented chatGPT with a "prompt" for each text passage. The prompt uses few-shot learning to incorporate 10 positive and 10 negative examples from which chatGPT can learn to answer the prompt as best as possible. We did not attempt to fine-tune the chatGPT model, as we were curious about the quality of the predictions coming out of the box. We did

Table 1 Top five titles of Bitcoin generator sites on the dark web

	Title	Frequency	Percent (%)
1	Bitcoin Generator Exploit - Official Hidden Service	493	59.3
2	Bitcoin Quantum Miner	76	9.1
3	Bitcoin Generator Exploit - Make Free Bitcoins!	49	5.9
4	SWEDISH BITCOIN MULTIPLIER	16	1.9
5	Bitcoin Investment	15	1.8
	Total	832	100



PEOPLE WHO HAVE RECENTLY USED OUR EXPLOIT

13AsgqJCWQGzH46ihE1JB6SSwpPiToYA1S received 0.5 BTC 7s ago

1KAZkZvmqhvUfFnp49m2W9ss57jNTKLgVT received 1.0 BTC 4s ago

1J8DgCiBwWxMkNTEFXWhfVCNGVMEHEuC1N received 0.5 BTC 0s ago

Fig. 4 Screenshot from the "Bitcoin Generator Exploit" website showing a ticker-tape-like series of recent bait transactions designed to impress potential victims

ask chatGPT how we could best formulate the prompt. The prompt in Appendix A results from this.

The inter-rater reliability (Cohen's Kappa) of our search terms and the prediction of chatGPT was 0.73. This score does not represent the added value of using an LLM, such as chatGPT, as an alternative coder. The added value consists of analysing the differences between the prediction of chatGPT and the deposit addresses found with the manual search terms. We reviewed all 50 snippets where our search terms and chatGPT disagreed. There were nine text passages where chatGPT correctly found a deposit address but our search terms did not. We created nine new search terms to capture those cases. The new Kappa is 0.80, and all results in the paper are based on 126=117+9 search terms. The inter-rater reliability is now acceptable, meaning the search terms are reasonably complete.

Looking up the transactions of the deposit addresses with ICA

We looked up all transactions from all deposit addresses and the deposit addresses of Badawi et al. (2022) via ICA. As a result, we obtained the date and time of the transaction and the exact amount received in BTC for every Bitcoin address. The amount in USD is estimated based on the average exchange rate on the day of the transaction.

A deposit address can be used for BG scams but also for other transactions. If we were to add up all incoming transactions from a deposit address, we would overestimate the revenue from BG scams (Gomez et al., 2023). One way to reduce the risk of overestimation is by time filtering. This can be done in two ways:*course time filtering* Only transactions within the observation period are counted; *fine time filtering* Transactions within the observation period but in a month when the scam site was unavailable are excluded. For example, the revenue from deposit address A in month M is only counted if at least one BG site was available in month M with the same deposit address A.

Badawi et al. (2022) applied course time filtering but not fine time filtering. We apply both, to show the impact of overestimation.

Results

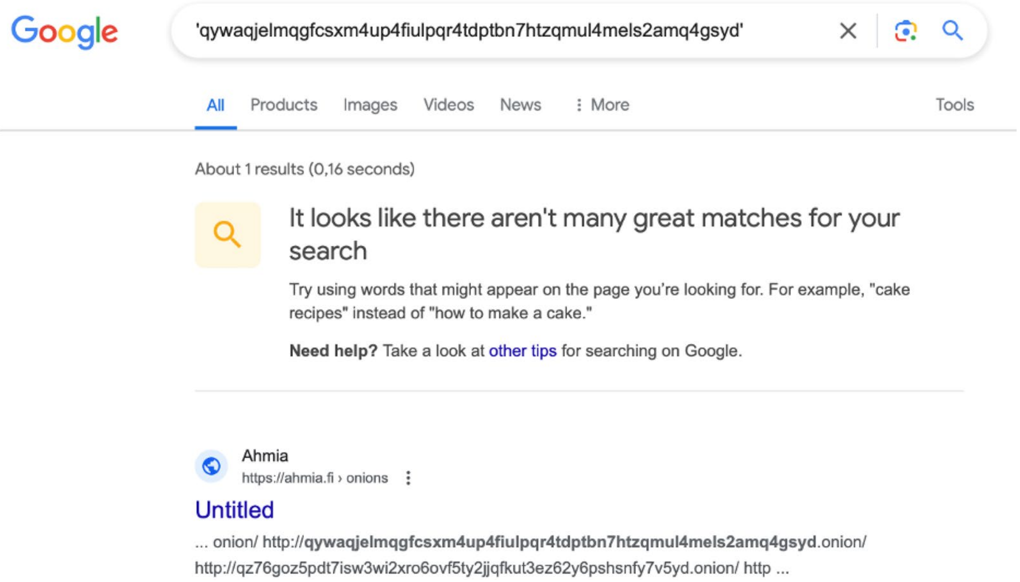
We found 832 dark web BG sites on the dark web, with 1,068 unique Bitcoin deposit addresses. We also found 1 Ethereum deposit address, which, like Badawi et al. (2022), we ignored. Below we present the results of the five sub-questions.

To what extent are BG sites promoted on the dark web?

On the clear web, according to our Google searches, we found 104 (29.2%) mentions of all 356 dark web BG sites These BG scams were mentioned 1 to 5 times. See Fig. 5 for an example. Most mentions were found on ahmia.fi. According to our DWM searches on the dark web, all 356 BG sites were mentioned, usually on an index site. The mentions include the title and the name of the BG site in all cases. We have not found any additional information, such as a review, which contains more than what is on the BG site. There are several advertisements for dark web BG sites on the clear web, as in Fig. 5. There are no ads or posts on YouTube videos or X promoting BG sites on the dark web. The easiest way to find a BG site on the dark web is to use a search and index site on the dark web.

How does BG site revenue differ on the clear and dark web?

All 1068 dark web deposit addresses were searched via ICA, and only 251 deposit addresses (23%) showed one or more transactions (reference date June 25, 2024). For comparison, Badawi et al. (2022) found more than 8000 deposit addresses on the clear web, of which transactions occurred on 3008 (38%). We also looked up all transactions from these clear web deposit addresses on the same reference date for a fair comparison. Figure 6 compares the cumulative revenue of BG sites on the clear and dark web. Unfortunately, the dark web research begins where the clear web research ends, but the figure shows that the revenue on the clear web is higher than on the dark web despite the clear web data being collected over a shorter period. The total revenue



'qywaqjelmqgfcsxm4up4fiulpqr4tdptbn7htzqmul4mels2amq4gsyd'

About 1 results (0,16 seconds)

It looks like there aren't many great matches for your search

Try using words that might appear on the page you're looking for. For example, "cake recipes" instead of "how to make a cake."

Need help? Take a look at other tips for searching on Google.

Ahmia
https://ahmia.fi › onions

Untitled

... onion/ http://**qywaqjelmqgfcsxm4up4fiulpqr4tdptbn7htzqmul4mels2amq4gsyd**.onion/ http://qz76goz5pdt7isw3wi2xro6ovf5ty2jjqfkut3ez62y6pshsnfy7v5yd.onion/ http ...

Fig. 5 Example of searching with Google for a dark web BG site, with one search result

of dark web BG site owners is 252,346 USD. On average this is USD 1005 per deposit address. The revenue from the clear web study was USD 9,477,659 in total and USD 3150 on average per deposit address. The revenue per deposit address on the clear web is three times higher than on the dark web, and the total amount is 40 times higher on the clear web for a shorter period. The clear web and dark web datasets had no common deposit addresses.

Fine time filtering

With fine time filtering, 132,403 USD of the dark web revenue is lost, leaving 119,943 USD (48%) in total revenue. The average revenue per Bitcoin address will then be 502.5 USD. Badawi et al. (2022) have not applied fine time filtering nor made the data public with which we could do fine time filtering. Based on the work of Gomez et al. (2023), we expect that almost 10 M USD clear web revenue is overestimated.

To what extent are BG campaigns being conducted on the dark web?

We assume that sites with the same title are duplicates because the title acts as a showcase for the site. A BG campaign is a series of websites with (almost) duplicate titles, possibly with different deposit addresses. We represent campaigns as a parameterised heatmap. In a heatmap, each coloured block represents a month of a campaign, with the block's colour indicating the parameter's value.

Figure 7 shows the monthly heatmap and, per title, the number of deposit addresses used by that title. For example, the title with the most deposit addresses is "Bitcoin Generator V2.0". Over 2 years, 414 different deposit addresses were used with this title, peaking at 63 in November 2023 and 61 in February 2024 (yellow blocks). Sites with this title have been online as of August 2022. Figure 7 shows that the most BG activity on the dark web was in the winter of 2023–2024.

Figures 8 and 9 sum up the revenue from deposit addresses per title and per month.

Figure 9 for the website titled "SWEDISH BITCOIN MULTIPLIER" is shown separately due to the exceptionally high revenue in March 2024 when 1 BTC (77,000 USD; the yellow block) was deposited to the address "1AUki ...swAVM". The victim could have created this transaction, but it is also possible that the scammer used this address to collect revenue from other addresses.

The transaction falls just outside the observation period.

On average, a deposit address is used by 1.12 (maximum 6) titles. At first glance, this could mean that one website title with only one deposit address probably has only one owner. However, owners may manage multiple deposit addresses. Copycats can also clone a site and change the deposit address, but we have no data to investigate this.

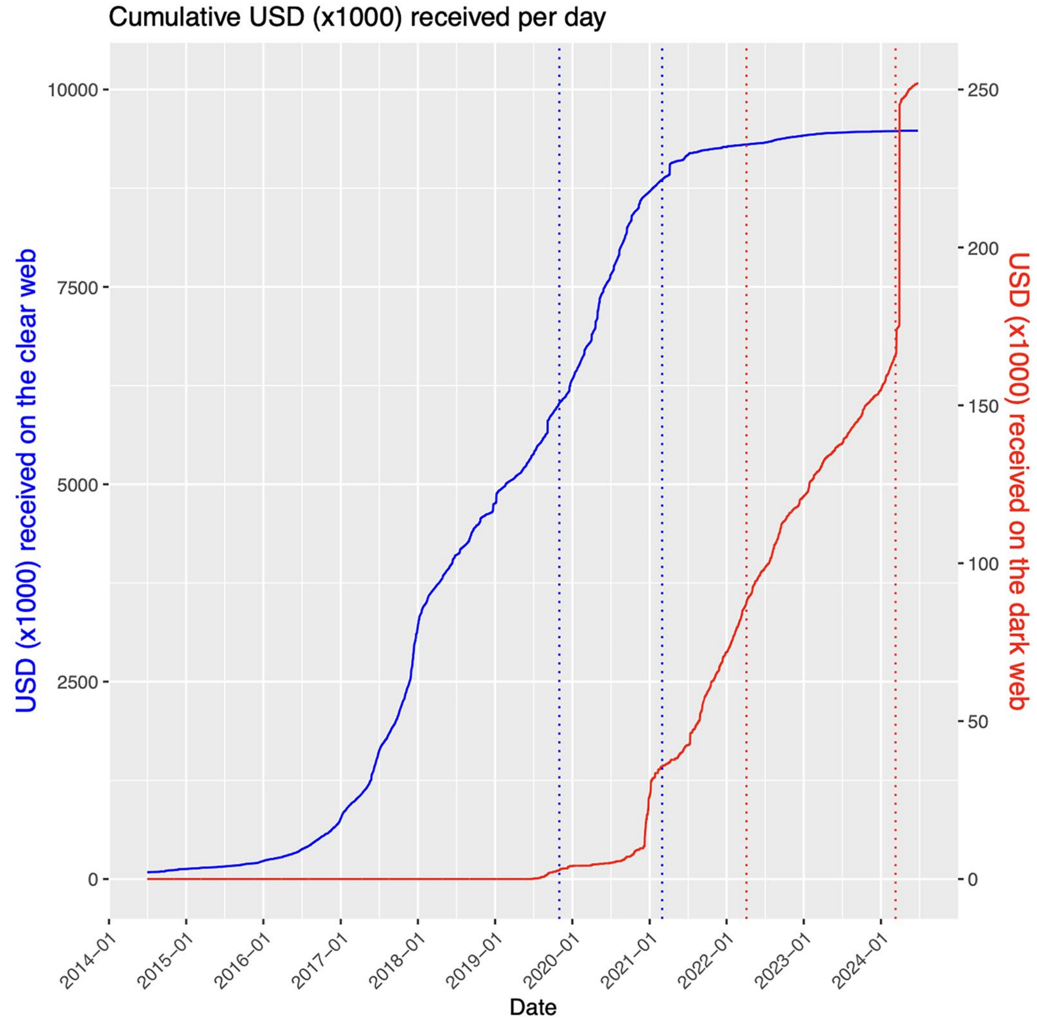


Fig. 6 The y-axes show the total cumulative revenue of BG sites in USD, left (in blue) for the clear web and right (in red) for the dark web. The dotted lines indicate the beginning and end of the study periods

To what extent is BG site revenue concentrated on the clear and dark web?

The top three clusters of clear web BG sites raised USD 2,378,092, USD 561,006 and USD 358,486, respectively, which is 52% of USD 9.5 M. On the dark web, the top three clusters raised 104K, 25K, and 21K USD, respectively, which is 35% of 288K USD. Therefore, the revenue is concentrated on the dark and clear web. The top five Bitcoin addresses of the first cluster on the clear web are shown in Table 2.

How do warnings against BG scams differ on the clear and dark web?

Several organisations warn consumers about scams like the Internet Crime Complaint Center (IC3). To investigate whether such warnings are adequate against BG sites, we sorted the deposit addresses of the clear and dark web BG sites by descending revenue. We then looked up the addresses as follows:

- We used Google to look up the top 100 deposit addresses of the clear web and the top 100 of the dark web.
- With the DWM, we looked up the top 100 deposit addresses of the clear web and the top 100 of the dark web.

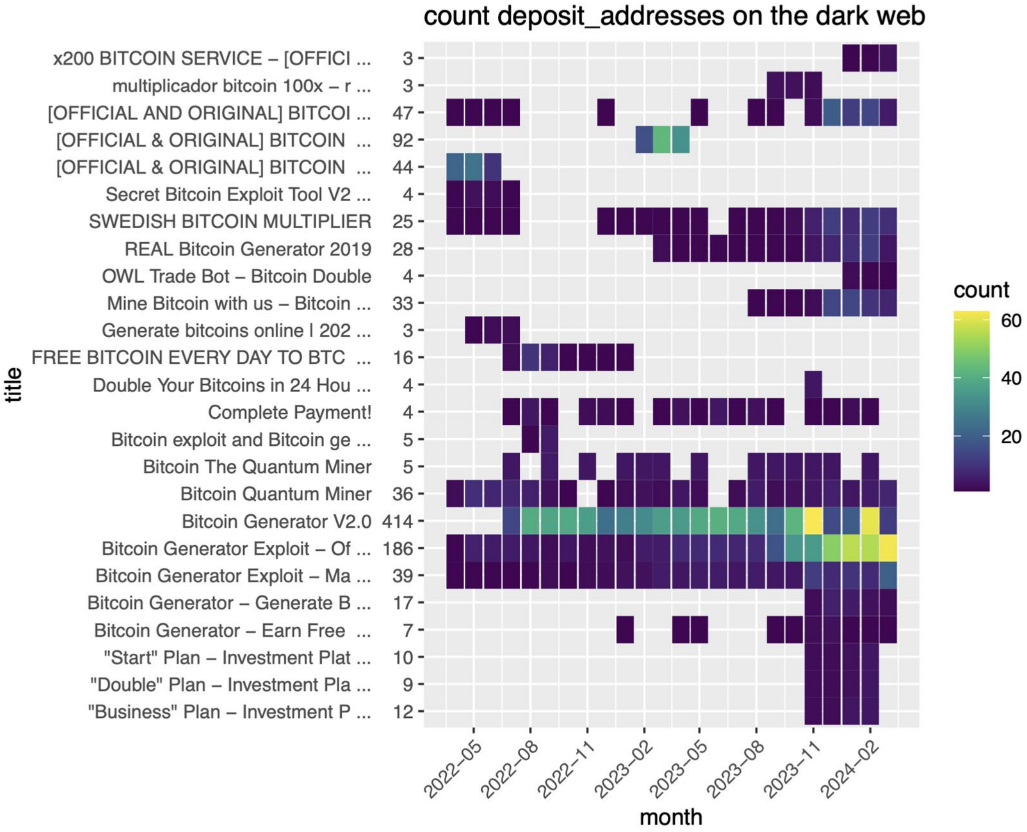


Fig. 7 Number of deposit addresses per website title over time (without time filtering). The x-axis is a timeline in months. Titles are listed in the first column along the y-axis and the total number of addresses is shown in the second column; the z-axis are different colours that indicate the number of deposit addresses per title and per month

We limited our searches to 400, divided into four sets of 100. Conducting these searches was time-consuming, often requiring reviewing several search result pages to identify relevant warnings.

Table 3 shows the numbers of hits in the columns "Total Google hits" and "Total DWM hits". We then read and coded the search results on a three-point scale: 0 = no warning, 1 = unclear warning and 2 = explicit warning. The essence of the coding is: how clear is the message that there is a warning? Would a visitor be warned if he landed on the page in question? An example of an "unclear warning" is a site with pages whose title is "anti-scam warning", but whose content consists only of a long list of Bitcoin addresses and their balances. Another example of an unclear warning is a paper by Bartoletti et al. (2018) in which Bitcoin generators are discussed, among other things.

We found an explicit warning on the clear web for eight clear web BG site deposit addresses and three dark web BG site deposit addresses. There are several clear websites where explict warnings are given such as chainabuse.com and bitcointalk.org. The screenshot in Fig. 10 is such an explicit warning. The message was posted by one of the moderators of the Bitcoin Forum (see bitcointalk.org). We did not find explicit warnings on the dark web, but we found two unclear warnings (with the already discussed title "anti-scam warning"). To summarise, there are hardly any warnings against BG sites on the dark web.

Discussion

The present study investigated 832 bitcoin generator (BG) sites on the dark web that are generally considered scams and compares the results with BG sites on the clear web, based on Badawi et al. (2022) findings.

The answers to the five sub-questions are as follows. Table 4 summarises the main results.

1. To what extent are BG sites promoted on the dark web? Owners of BG sites on the dark web pay little attention to advertising their sites. The dark web BG sites can all be found on dark web search and index

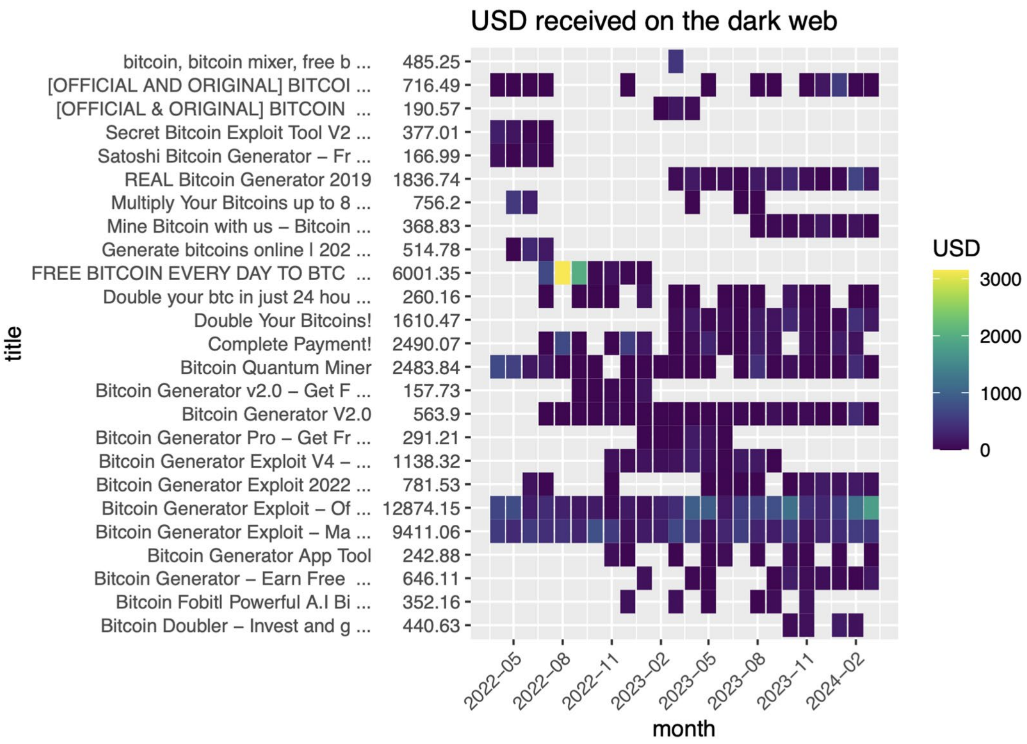


Fig. 8 Revenue from various titles per month (without time filtering). Titles are listed in the first column along the y-axis and the total revenue is shown in the second column; the z-axis are different colours indicating revenue per title and per month

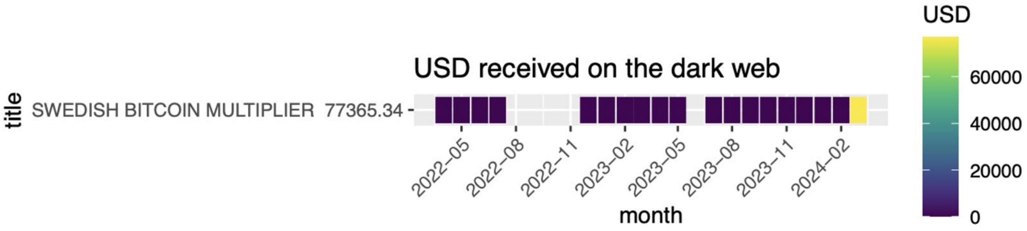


Fig. 9 Revenue from the deposit addresses for the SWEDISH BITCOIN MULTIPLIER per month (without time filtering)

sites, but there is hardly more to be found than the name and the title. So, hardly any efforts have been made to advertise BG sites. Related research shows that YouTube channels (Bouma-Sims & Reaves, 2021) and X (Cola et al., 2023) are often abused to advertise scams, but we did not find any ads for dark web BG sites on YouTube and X.

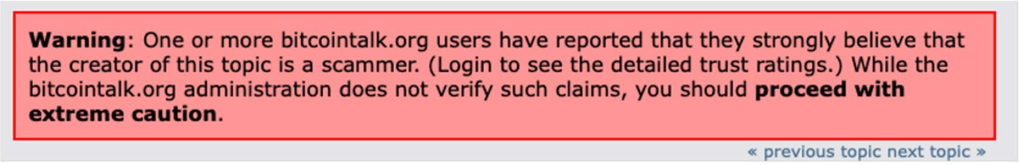
2. How does BG site revenue differ on the clear and dark web? The revenue from a dark web BG site is approximately 1/3 smaller per deposit address than the revenue from a clear web BG site. Some illegal activities are more prominent on the dark web than the clear web. For example, Villalva et al. (2018) show that stolen accounts are traded more on the

Table 2 The top cluster of deposit addresses on the clear web

Crypto address	First tx at	Last tx at	USD received
1NHrS ...njB1o	2016-08-01 05:10:07	2019-12-16 21:51:46	917,464
1FYrA ...qrYCj	2017-05-05 12:24:48	2020-11-21 11:08:53	398,854
128ks ...HScN2	2016-08-05 09:47:04	2018-07-18 14:50:01	195,647
17grN ...ocNnR	2016-08-07 03:52:25	2020-05-05 05:44:26	180,912
1BjGD ...dShdP	2016-12-23 01:07:16	2017-09-24 19:34:00	136,883
125zc ...qx2oi	2016-07-31 13:01:30	2017-10-03 06:11:44	115,055
Total			2,378,092

Table 3 The number of times a deposit address was found on the clear web and the dark web by type of warning

	Google search hits on the clear web (N=200 searched deposit addresses)				**DWM search hits on the dark web (N= 200 searched deposit addresses)**			
Found mentions of deposit address	No warning	Unclear warning	Explicit warning	Total Google hits	No warning	Unclear warning	Explicit warning	Total DWM hits
Clear web BG site deposit address	5	3	8	16	25	2	0	27
Dark web BG site deposit address	7	0	3	10	97	0	3	100
Total found deposit addresses	12	3	11	26	122	2	3	127



Warning: One or more bitcointalk.org users have reported that they strongly believe that the creator of this topic is a scammer. (Login to see the detailed trust ratings.) While the bitcointalk.org administration does not verify such claims, you should **proceed with extreme caution**.

Fig. 10 Explicit warning

Table 4 Summary of the results

Variable	**Clear web**	**Dark web**
Observation period starts	2019-11-01	2022-04-05
Observation period ends	2021-03-01	2024-03-10
Observation period length	15 months	23 months
Number of deposit addresses with transactions	3008	251
Total revenue (without time filtering)	9,477,659 USD	252,346 USD
Revenue per deposit address (without time filtering)	3,150 USD	1,005 USD
Revenue top 3 clusters (without time filtering)	35%	52%
Revenue (with time filtering)	NA	48%
Explicit warnings against the top 100 deposit addresses (assessed in May 2024)	8 out of 100	3 out of 100

dark than on the clear web. The risk of selling stolen accounts likely outweighs the risks of a BG scam. Because they must use the TOR browser, dark web users may be more tech-savvy than clear web users. Therefore, dark web users may be quicker to realise that an offer is "too good to be true".

3. To what extent are BG campaigns being conducted on the dark web? Most BG sites belong to a campaign, such as the "Swedish Bitcoin Multiplier" (See Fig. 9). Approximately half of the transactions to a deposit address are made in the months when a campaign is not online. This means that about half of the revenue cannot originate from this BG campaign. This form of overestimation of revenue can be avoided with time filtering (Gomez et al., 2023).
4. To what extent is BG site revenue concentrated on the clear and dark web? The top three clusters of deposit addresses with the most revenue are responsible for a large part of the total revenue: 35% on the clear web and 52% on the dark web. BG sites are likely connected because the Bitcoin addresses in a cluster will likely belong to the same person or group (Fröwis et al., 2020).
5. How do warnings against BG scams differ on the clear and dark web? There are hardly any warnings that a BG site could be a scam, not on the clear or dark web. We found few scam warnings, while various organisations collect scam warnings (Choi et al., 2022). Because our searches via Google also reached the main scam warning sites, no scam warnings must be posted for the Bitcoin addresses we searched. With so few warnings, we could not investigate the effect of the warning on the scammer's revenue.

Our results illustrate the rational choice principles of crime (Cornish & Clarke, 2008). The risks of setting up a BG scam are higher on the clear web, but the expected earnings are higher. The risks of setting up a BG scam on the dark web are lower, but the expected earnings are also lower. In both cases we judge that the effort is minimal.

The revenue from BG sites – on the dark web and the clear web – is relatively small compared to the total costs of cryptocurrency fraud, which run into the billions (Carpentier-Desjardins et al., 2023). Bitcoin generators do not yield much, so it will likely not become a priority for law enforcement anytime soon. However, if that changes, the top clusters are good priorities for law enforcement.

Limitations

Limitations may have arisen in our study due to how our sample composition was designed and the data were collected and analysed. There may be several reasons why BG sites or deposit addresses or adverts were not included in our research:

1. Missing sites in the dataset. If a BG site does not appear in the DWM dataset, then the site's deposit addresses are also not in the dataset. With a benchmark study, we were able to determine that the DWM dataset is more extensive than all public dark web datasets, such as ahmia.fi. It is not known whether commercial parties have more data. In addition, if we could not find any new BG sites through an extensive search, the question is how dark web users could have easily found the sites we did not find.
2. Assumption that what is stated on the BG sites is true. Based on the accompanying text of a Bitcoin address, we determine that a Bitcoin address is a deposit address. We have not been able to ask the owner of the deposit address whether this is indeed the case because the owner is anonymous.
3. Missing direct marketing data. BG site owners may advertise their sites in other ways, for example, with personal messages via email, WhatsApp, or Telegram (Chergarova et al., 2022). We have no data about personal messages at our disposal.
4. Differences between dark web and clear web research period. The investigation period of the BG sites on the clear web ends where the investigation period of the dark web begins. The differences in revenue may be due to a decline in interest in Bitcoin generators because, as the phenomenon, it is more than 10 years old (see bitcoin-generator.net on the Internet Archive).
5. Ads on the clear web may have been removed over time, meaning they may not have been visible when we searched for Google ads. To circumvent this limitation partially, we limited our search to ads on websites discovered by DWM in the last 10 months.
6. When a user searches on Google, results are selected and prioritised based on several factors, including the user's search history. This can cause a certain degree of bias in the search results. Unfortunately, the details of the search algorithms have not been made public by Google (Piasecki et al., 2018).
7. We have collected deposit addresses from BG scam websites, making it likely that incoming transactions are due to BG scams. However, we have no information whether these deposit addresses have been used for other purposes. Therefore, the revenue of BG scams on the dark web is an upper limit, just as the revenue reported by Badawi et al. (2022) is an upper limit.

Conclusion and future work

For most people, the technology behind cryptocurrencies is difficult to understand, although certain aspects are well-known. For example, many know that a Bitcoin transaction requires a mining fee. The Bitcoin Generator (BG) scam plays on this by using a convincing but fictitious technical story to mislead victims. These victims are asked to pay a mining fee, with the promise of a significant profit. The alleged profit is supposedly possible due to the technical "cleverness" of the scammers.

Much research has been into fraud that uses technical means, such as malware and phishing. BG scams have been researched on the clear web (Badawi et al., 2022) but not on the dark web. This study fills that gap by exploratory research into BG scams on the dark web while also comparing them to BG scams on the clear web of (Badawi et al., 2022). We discussed the effort scammers have to put into setting up and maintaining a website, the earnings they expect and the risks they may fear, and where possible, we compared the results for the Bitcoin Generator (BG) scam on the dark web with the same scam on the clear web.

BG scams are one of many methods to defraud gullible people with a convincing technical story. Although the money involved in BG scams is relatively small, this type of scam provides an excellent testing ground for research into similar types of scams, as well as more complex, technically oriented frauds, such as the more sophisticated AI scams that are increasingly appearing (Gressel et al. 2024). Cryptocurrency and AI technology can be complex to many people, making it easier for scammers who appear knowledgeable to be convincing. Exploratory research such as this is essential to guide future research into the societal impact of technology abuse. This study contributes to that research and raises some fundamental questions for further research:

- How do phenomena like the BG scam change over time? How have technical and social changes and law enforcement influenced the scam (Décary-Hétu & Giommoni, 2017)?
- How do individuals who engage with the dark web differ from those who primarily use the clear web regarding their performance on standardised psychological assessments, such as measures of the Big Five personality traits? (Sirola et al., 2024)?
- What influence do appropriately worded warnings have on victimisation (Howell et al., 2024)? We found a small number of warnings against BG scams. It would be interesting to investigate this further, for example, by analysing forums such as bitcointalk.org or reddit.com, in particular, whether this has a preventive effect (Siu & Hutchings, 2023).
- To what extent do crime scripts for a specific fraud differ between the dark and the clear web (Holt and Lee, 2022)? What insights does this provide for measures to hinder or prevent fraud?
- How can a community like VirusTotal be created for the dark web? VirusTotal offers internet users the possibility to verify the safety of a URL. Based on this information, users can decide whether they want to visit the URL. For the dark web, there is only one website that can check URLs for the presence of illegal content. That is too narrow a basis to get reliable answers.

Our study shows that Bitcoin generator scams on the dark web, compared to the clear web, are just as easy to set up, pose less risk to the scammer, and yield less

Appendix A ChatGPT prompt

A website with a business proposition may contain several Bitcoin addresses. The business owner may control some of these, but external entities will control others.

Instructions to pay usually contain the Bitcoin address controlled by the business owner. Here some examples of text snippets where the business owner likely controls Bitcoin addresses:

1. Referral Contest Referral Video Get for each affiliate Enter your bitcoin address,
2. Rent your ASIC Quantum CPU Premium Iron Robot Price,
3. Our professional and experienced staff gets very good and stability result in money management activity ADDRESS FOR YOUR DEPOSIT,
4. Please send BTC to,
5. Now submit the generated mining fee so that your transaction can be verified. Send BTC to,
6. takes minutes for funds to appear in your wallet after paying the miners fee of BTC Please send BTC to,
7. Lastest transactions of sending to our users with used this free bitcoin cloud mining system,
8. Send some Bitcoins and double them in just one day Bitcoin address for making deposit,
9. Sent Bitcoin Miner Network Fee for this transaction at,
10. Invoice ID Send exactly to.

A review or a list of transactions on such a website often contains Bitcoin addresses controlled by external entities. Here some examples of text snippets where others likely control Bitcoin addresses:

1. Lastest transactions of sending to our users with used this bitcoin cash cloud mining system,
2. Mining Pools Exploited Blockchain Injection Confirmed Transactions,
3. Every address that is sent too late, gets their BTC immediately sent back. Transactions for address,
4. Last Payouts DATE ADDRESS DEPOSIT PAYOUT TIME DATE,
5. Earn a referral commission for every deposit instantly,
6. To do this, you need to specify the wallet from which you made the transfer. Enter your wallet example,
7. Double my Bitcoins Latest Investments DATE ADDRESS AMOUNT TIME LEFT PAYOUT,
8. RECENT PAYOUTS TIME BITCOIN ADDRESS DEPOSITED PAID AMOUNT,
9. Reffer Your Friend Get Refferal Commission Use Link Example,
10. Date Currency Address Deposit Amount Payment Bitcoin.

Given a text, classify the Bitcoin address within it as controlled by the business owner or others. Also, return the probability of the address being controlled by the business owner.

For example, given the text:

Industry leaders when it comes to DDoS protection and data encryption. Please confirm your deposit: Send the required BTC to: 12sG ...4A75b

The output should be in JSON format with two properties:

- *Controlled* indicating whether the address is controlled by the website owner or by external entities.

- *Probability* the likelihood (in decimal format with two decimal points) that the business owner controls the address.

Example output: `{ "controlled": "owner", "probability": 0.85 }`

Here is the text to analyse:...

Appendix B Primary Dataset and the Ethics of Secondary Analysis

We describe the background and collection of the primary information, which is outside the scope of our research but relevant to the ethical aspects of the secondary analysis.

The primary datasets were collected to contribute to a safer society by analysing services on the dark web. These services are often criminal and are used as intelligence or evidence by parties working towards a safer society. By being transparent about our research, we hope to contribute to the transparency of law enforcement. DWM uses a standard snowball method to retrieve web pages from the dark web. The process starts with a series of starting addresses being downloaded. New addresses found on the downloaded pages are also retrieved, etc. Others, such as Google, use this process on the clear web. ICA ingests public blockchains of cryptocurrencies.

We now follow the Menlo Report (Bailey et al., 2012) to discuss the ethical risks (R) and mitigations (M).

B.1 Respect for Persons

R: The BG sites on the dark web contain data that may be traceable to a person, such as Bitcoin addresses. M: We have not published complete addresses and only report totals to make tracing back to a person virtually impossible.

R: In the secondary analysis, we only looked up Bitcoin addresses with evidence of being involved in fraudulent activity. This allowed us to report the extent of fraudulent activity (in dollars). M: We have not attempted to trace the owner of a BG site or Bitcoin address in any way.

B.2 Beneficence

R: Law enforcement may approach owners of BG sites more quickly than they would without using our research, which could harm those owners. M: According to the principle of proportionality, we argue that the advantage for law enforcement (and society) outweighs the disadvantage for the owners of BG sites.

B.3 Justice

R: Because the extent of the Dark web is unknown and cannot be known, it may bias our research. M: The DWM dataset is one of the largest collections that has been collected without known bias.

B.4 Respect for Law and Public Interest

R: The revenue from BG sites may be underestimated by our research, which could cause law enforcement to give this phenomenon a lower priority. M: We investigated several possible causes of under and overestimation of total revenue.

R: Is it possible that the research has overloaded the TOR network? M: The secondary data analysis generated no traffic on the TOR network. We spread our attempts to reach 94 BG sites via the TOR browser over three days.

R: Did DWM and ICA face additional risks due to the investigation? M: The secondary analysis was conducted entirely on the servers, and all data remained within the intranet hosting the DWM.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40163-025-00246-w.

Updated search terms from Badawi et al.

Acknowledgements

We thank Emad Badawi and the anonymous referees for their comments on the paper. Crypto Legal (UK), a renowned law firm specialising in blockchain and cryptocurrency forensics and legal services, has provided valuable insights into digital asset investigations. The research complies with ethical standards because all data that has been analysed is open data that has been made public to be analysed.

Author contributions

MJ was responsible for the theory and PH for the data analysis, and MvS for the data collection. All authors read and approved the final manuscript.

Funding

The work described in this paper is performed in the H2020 project STARLIGHT ("Sustainable Autonomy and Resilience for LEAs using AI against High priority Threats"). This project has received funding from the European Union's Horizon 2020 research and in- novation program under grant agreement No 101021797.

Data availability

To access the DWM-dataset, please contact https://www.cflw.com. The Bitcoin blockchain data is publicly available.

Declarations

Competing interests

The authors declare that they have no competing of interest.

Received: 4 October 2024 Accepted: 20 February 2025
Published online: 13 April 2025

References

Badawi, E., Jourdan, G.-V., & Onut, I.-V. (2022). The bitcoin generator scam. *Blockchain: Research and Applications, 3*(3), 100084. https://doi.org/10.1016/j.bcra.2022.100084

Bailey, M., Dittrich, D., Kenneally, E., & Maughan, D. (2012). The Menlo report. *IEEE Security and Privacy, 10*(2), 71–75. https://doi.org/10.1109/MSP.2012.52

Bartoletti, M., Pes, B., & Serusi, S. (2018). Data mining for detecting Bitcoin Ponzi schemes. In: *Crypto Valley Conference on Blockchain Technology (CVCBT)*. IEEE. https://doi.org/10.1109/CVCBT.2018.00014

Bouma-Sims, E., & Reaves, B. (2021). A First Look at Scams on YouTube. In *Workshop on Measurements, Attacks, and Defenses for the Web (MADWeb) Workshop on measurements, attacks, and defenses for the web (madweb) (PP 1-7)*. Internet Society. https://doi.org/10.14722/madweb.2021.23001

Carpentier-Desjardins, C. , Paquet-Clouston, M. , Kitzler, S. , & Haslhofer, B. (2023). *Mapping the DeFi crime landscape: An evidence-based picture (ArXiv) Université de Montréal*. https://doi.org/10.48550/arXiv.2310.04356

Chergarova, V., Arcanjo, V., Tomeo, M., Bezerra, J., Vera, L. M., & Uloa, A. (2022). Cryptocurrency fraud: A study on the characteristics of criminals who are using fake profiles on a social media platform to persuade individuals to invest into cryptocurrency. *Issues in Information Systems, 23*(3), 242–252. https://doi.org/10.48009/3_iis_2022_120

Choi, J., Kim, J., Song, M., Kim, H., Park, N., Seo, M., & Shin, S. (2022). A large-scale Bitcoin abuse measurement and clustering analysis utilizing public reports. *IEICE Transactions on Information and Systems, 105*(D–7), 1296–1307. https://doi.org/10.1587/transinf.2021EDP7182

Cialdini, R. B. (2009). *Influence: The psychology of persuasion*. Harper Collins.

Cola, G. , Mazza, M. , & Tesconi, M. (2023). From tweet to theft: Tracing the flow of stolen cryptocurrency. In *The Italian Conference on CyberSecurity (ItaSec) (P Article 11)*. CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3488/paper11.pdf

Cornish, D. B., & Clarke, R. V. (2008). The rational choice perspective. In R. Wortley & L. Mazerolle (Eds.), *Environmental criminology and crime analysis (ECCA)* (pp. 21–47). Willan Publishing. https://doi.org/10.4324/9780203118214-10

Décary-Hétu, D., & Giommoni, L. (2017). Do police crackdowns disrupt drug cryptomarkets? A longitudinal analysis of the effects of operation onymous. *Crime, Law & Social Change, 67*(1), 55–75. https://doi.org/10.1007/s10611-016-9644-4

Fröwis, M., Gottschalk, T., Haslhofer, B., Rückert, C., & Pesch, P. (2020). Safeguarding the evidential value of forensic cryptocurrency investigations. *Forensic Science International: Digital Investigation, 33*, 200902. https://doi.org/10.1016/j.fsidi.2019.200902

Gomez, G. , van Liebergen, K. , & Caballero, J. (2023). Cybercrime Bitcoin revenue estimations: Quantifying the impact of methodology and coverage. In *ACM SIGSAC conference on computer and communications security (ccs)* (pp. 3183–3197). ACM. https://doi.org/10.1145/3576915.3623094

Gressel, G. , Pankajakshan, R. , & Mirsky, Y. (2024) Discussion paper: Exploiting llms for scam automation: A looming threat. In *3rd ACM workshop on the security implications of deepfakes and cheapfakes (wdc)* (pp. 20–24). ACM. https://doi.org/10.1145/3660354.3660356

Holt, T. J., & Lee, J. .R. (2022). A crime script analysis of counterfeit identity document procurement online. *Deviant Behavior, 43*(3), 285–302. https://doi.org/10.1080/01639625.2020.1825915

Howell, C. J., Maimon, D., Perkins, R. C., Burruss, G. W., Ouellet, M., & Wu, Y. (2024). Risk avoidance behavior on darknet marketplaces. *Crime & Delinquency, 70*(2), 519–538. https://doi.org/10.1177/00111287221092713

Lee, S. , Yoon, C. , Kang, H. , Kim, Y. , Kim, Y. , Han, D., & Shin, S. (2019). Cybercriminal minds: An investigative study of cryptocurrency abuses in the dark web. In *Network and distributed systems security (NDSS)* (pp. 1–15). Internet Society. https://doi.org/10.14722/ndss.2019.23055

Nurmi, J., & Niemel, M.S. (2017). Tor de-anonymisation techniques. In *11th international conference network and system security (NSS)* (vol 10394, pp. 657–671). Springer. https://doi.org/10.1007/978-3-319-64701-2

Oest, A. , Zhang, P. , Wardman, B. , Nunes, E. , Burgis, J. , Zand, A., & Ahn, G- J. (2020). Sunrise to sunset: Analyzing the end-to-end life cycle and effectiveness of phishing attacks at scale. In *29th USENIX security symposium* (pp. 361–377). USENIX. https://www.usenix.org/conference/usenixsecurity20/presentation/oest-sunrise

Ogbanufe, O., Wolfe, J., & Baucum, F. (2024). Towards a conceptual typology of darknet risks. *Journal of Computer Information Systems, 64*(4), 565–576. https://doi.org/10.1080/08874417.2023.2234323

Phillips, R. & Wilder, H. (2020). Tracing cryptocurrency scams: Clustering replicated advance-fee and phishing websites. In *International conference on blockchain and cryptocurrency (ICBC)* (pp. 1–8). IEEE. https://doi.org/10.1109/ICBC48266.2020.9169433

Piasecki, J., Waligora, M., & Dranseika, V. (2018). Google search as an additional source in systematic reviews. *Science and Engineering Ethics, 24*, 809–810. https://doi.org/10.1007/s11948-017-0010-4

Reep-van den Bergh, C.M.M., & Junger, M. (2018). Victims of cybercrime in Europe: a review of victim surveys. *Crime Science Journal, 7*, 5. https://doi.org/10.1186/s40163-018-0079-3

Sirola, A., Savolainen, I., & Oksanen, A. (2024). Who uses the dark web? Cross-national and longitudinal evidence on. *Personality and Individual Differences, 227*, 112709. https://doi.org/10.1016/j.paid.2024.112709

Siu, G.A. & Hutchings, A. (2023). "Get a higher return on your savings!": Comparing adverts for cryptocurrency investment scams across platforms . In *European symposium on security and privacy workshops (euros &pw)* (pp. 1–12). IEEE. https://doi.org/10.1109/EuroSPW59978.2023.00023

Spitters, M. , Verbruggen, S. , & Staalduinen, M.V. (2014). Towards a comprehensive insight into the thematic organization of the Tor hidden services. In *Joint Intelligence and Security Informatics Conference*. (pp. 220–223). IEEE. https://doi.org/10.1109/JISIC.2014.40

Stajano, F., & Wilson, P. (2011). Understanding scam victims: seven principles for systems security. *Communications of the ACM, 54*(3), 70–75. https://doi.org/10.1145/1897852.1897872

Trozze, A., Kamps, J., Akartuna, E. A., Hetzel, F. J., Kleinberg, B., Davies, T., & Johnson, S. D. (2022). Cryptocurrencies and future financial crime. *Crime Science Journal, 11*, 1. https://doi.org/10.1186/s40163-021-00163-8

Vakilinia, I. (2022). Cryptocurrency giveaway scam with YouTube live stream. In *13th annual ubiquitous computing, electronics & mobile communication conference (uemcon)* (pp. 195–200). IEEE. https://doi.org/10.1109/UEMCON54665.2022.9965686

van der Bruggen, M., & Blokland, A. (2021). A crime script analysis of child sexual exploitation material fora on the darkweb. *Sexual Abuse, 33*(8), 950–974. https://doi.org/10.1177/1079063220981063

Vasek, M., & Moore, T. (2019). Analyzing the Bitcoin Ponzi scheme ecosystem. In *Financial cryptography and data security* (VOL. 10958, pp. 101–112). Springer. https://doi.org/10.1007/978-3-662-58820-8_8

Villalva, D. B., Onaolapo, J., Stringhini, G., & Musolesi, M. (2018). Under and over the surface: a comparison of the use of leaked account credentials in the dark and surface web. *Crime Science Journal, 7*, 17. https://doi.org/10.1186/s40163-018-0092-6

Winter, P. , Edmundson, A. , Roberts, L.M. , Dutkowska-Zuk, A. , Chetty, M. , & Feamster, N. (2018). How do Tor users interact with onion services? In *27th USENIX security symposium* (pp. 411–428). USENIX https://www.usenix.org/conference/usenixsecurity18/presentation/winter

Xia, P. , Yu, Z. , Wang, K. , Ma, K. , Chen, S. , Luo, X., & Wang, H. (2024). The devil behind the mirror: Tracking the campaigns of cryptocurrency abuses on the dark web arXiv. Beijing University of Posts and Telecommunications. https://doi.org/10.48550/arXiv.2401.04662

Xiao, Z. , Yuan, X. , Liao, Q.V. , Abdelghani, R. , & Oudeyer, P- Y. (2023) Supporting qualitative analysis with large language models: Combining codebook with GPT-3 for deductive coding. In 28th Int. Conf. on Intelligent User Interfaces (IUI) Companion (pp. 75–78). ACM. https://doi.org/10.1145/3581754.3584136

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.