Building Collaborative Cybersecurity for Critical Infrastructure Protection: Empirical Evidence of Collective Intelligence Information Sharing Dynamics on ThreatFox

Building Collaborative Cybersecurity for Critical Infrastructure Protection: Empirical Evidence of Collective Intelligence Information Sharing Dynamics on ThreatFox

Eric Jollès[1], Sébastien Gillard[2,3], Dimitri Percia David[1,4], Martin Strohmeier[1,5], and Alain Mermoud[1](✉)

[1] Cyber-Defence Campus, armasuisse Science and Technology, Zurich, Switzerland
mermouda@ethz.ch

[2] Information Science Institute, Geneva School of Economics and Management, University of Geneva, Geneva, Switzerland

[3] Department of Defense Economics, Military Academy at ETH Zurich, Zurich, Switzerland

[4] Institute of Entrepreneurship and Management, University of Applied Sciences HES-SO Valais-Wallis, Sierre, Switzerland

[5] Department of Computer Science, Oxford University, Oxford, UK

Abstract. This article describes three collective intelligence dynamics observed on ThreatFox, a free platform operated by abuse.ch that collects and shares indicators of compromise. These three dynamics are empirically analyzed with an exclusive dataset provided by the sharing platform. First, participants' onboarding dynamics are investigated and the importance of building collaborative cybersecurity on an established network of trust is highlighted. Thus, when a new sharing platform is created by abuse.ch, an existing trusted community with 'power users' will migrate swiftly to it, in order to enact the first sparks of collective intelligence dynamics. Second, the platform publication dynamics are analyzed and two different superlinear growths are observed. Third, the rewarding dynamics of a credit system is described - a promising incentive mechanism that could improve cooperation and information sharing in open-source intelligence communities through the gamification of the sharing activity. Overall, our study highlights future avenues of research to study the institutional rules enacting collective intelligence dynamics in cybersecurity. Thus, we show how the platform may improve the efficiency of information sharing between critical infrastructures, for example within Information Sharing and Analysis Centers using ThreatFox. Finally, a broad agenda for future empirical research in the field of cybersecurity information sharing is presented - an important activity to reduce information asymmetry between attackers and defenders.

Keywords: Information Sharing and Analysis Center · Threat Intelligence · Sharing Platform · Security Information Sharing · Collaborative Cybersecurity · Collective Intelligence · Indicator of Compromise

B. Hämmerli et al. (Eds.): CRITIS 2022, LNCS 13723, pp. 140–157, 2023.
https://doi.org/10.1007/978-3-031-35190-7_10

1 Introduction

Cybersecurity Information Sharing (CIS) is an important activity to reduce the information asymmetry between attackers and defenders [1]. This activity also allows the production of Cyber Threat Intelligence insights, which enables organizations to proactively detect cyberrisks and prevent malicious activities [2]. More than two decades ago, the first Computer Emergency Readiness Teams (CERT) [3] and Information Sharing and Analysis Centers (ISACs) [4] were established to allow critical infrastructure operators to share important information about cyberthreats [5]. Today, threat intelligence platforms help organizations aggregate, correlate, and analyze threat data from multiple sources in quasi real-time to support defensive actions [6,7]. In addition, open-source solutions, such as the MISP[1] Threat Sharing platform [8] or the AlienVault Open Threat Exchange[2] (OTX), have been proposed to counterbalance the influence of large cybercriminal networks and organizations. In March 2021, abuse.ch launched the ThreatFox[3] project, a platform used to collect and share IoCs to help IT-security professionals and threat analysts protect their customers from cyberthreats.

In this article, three collective intelligence dynamics observed on ThreatFox are empirically investigated, with the goal of better understanding the institutional rules that enact such collective intelligence dynamics. First, participants' onboarding dynamics are investigated and the importance of building collaborative cybersecurity on established networks of trust is highlighted. Second, the platform publication dynamics are analyzed and superlinear growth is observed during the first one hundred days. Third, a rewarding dynamic of a credit system is described—a promising incentive mechanism to improve information sharing in open-source intelligence communities.

The remainder of this article begins by providing a brief overview in Sect. 2 of CIS and collaborative cybersecurity. In Sect. 3, an empirical analysis of the three dynamics is conducted before presenting the obtained results in Sect. 4. Section 5 discusses this work and brings some improvement recommendations for the platform. Section 6 presents a broad research agenda on CIS and collective intelligence in cybersecurity, and Sect. 7 concludes this work.

2 Related Work

The constant evolution of cyberthreats has forced organizations and governments to develop new strategies [2] to reduce the risks of security breaches [9]. In this regard, the development of collaborative platforms as governance-strategy and knowledge-management tools has highlighted the importance of information sharing [10]. Hence, the World Economic Forum has recently recognized the fact that CIS is critical to helping improve collective security in the digital ecosystem on which society increasingly relies [11]. However, CIS faces multiple barriers.

[1] https://www.misp-project.org.

[2] https://otx.alienvault.com.

[3] https://threatfox.abuse.ch/.

First, these challenges have a social aspect; human beings tend not to optimize organizational goals [12] without selective incentive [13] and—in the case of collective action—might behave selfish in ways that do not support the overall goal of information sharing [14], leading to situations such as the prisoner's dilemma [15]. In this situation, it is in the interest of two players to cooperate on an issue; however, in the absence of communication between them, each will choose to betray the other [16]. As a result, cybersecurity professionals likely share less information than is desirable, resulting in knowledge asymmetry that benefits the attackers [1]. In particular, stakeholders strategically select their contributions to share, leading to truncated and imperfect information sharing.

In the absence of trust, commitment, and a shared vision among stakeholders, organizations are reluctant to share information for fear of disclosure, reputational risk, or loss of competitive power [17]. In this respect, information sharing can be understood as a marketplace in which transactions take place and knowledge is transferred [9].

2.1 Collective Intelligence Dynamics in Cybersecurity

The scientific literature confirms that sharing information security among human agents operating information systems is conducive to improving cybersecurity [1]. However, empirical analysis shows that 'sharing centers', such as ISACs do not always function optimally [18]. To improve CIS, the computer science technical literature generally focuses on getting the exchange format right, through data models, the adoption of specific technologies [19], or sharing conventions, such as the Traffic Light Protocol (TLP). This approach neglects the fact that information sharing is a human activity needing incentive mechanisms [13], which is not related to technology. Hence, CIS can be viewed as a collective intelligence process through which group intelligence emerges from repeated collaboration and collective efforts through crowdsourcing and peer-reviewing [20].

2.2 Linking Institutional Economics and Information Sharing

Institutional economics focuses on understanding the role of the evolutionary process and institutions in shaping economic behavior. With this study, a better understanding of the institutional rules enacting collective intelligence dynamics in cybersecurity is sought. By understanding and measuring these rules, an attempt is made to explain the success of abuse.ch compared with other platforms that are less successful and use different rules, such as OTX[4]. Therefore the assumption is made, that the success of sharing platforms is directly linked to the rules implemented from its creation. Hence, an institutional economics framework is used to describe the three identified dynamics. Studying these three dynamics leads to the hope that, at the same time, this study contributes to the institutional economics literature from a cybersecurity perspective, as was done in previous interdisciplinary work using a similar approach [12].

[4] The success of abuse.ch is publicly visible on Twitter, especially through the number of followers (more than 25.5K in less than a year of existence). https://twitter.com/abuse_ch.

Analyses have already been conducted on various information security sharing platforms. An analysis of the widely used open source threat sharing platform MISP [7] shows how collective action in this type of platform can increase the efficiency in the time required to fully characterize a cybersecurity threat. Their results generally informs how collective actions can be organized online at scale and in a modular fashion to address a large number of time-critical tasks.

3 Data and Methodology

3.1 abuse.ch: Community Driven Threat Intelligence

abuse.ch is a project created years ago by Roman Hüssy[5]. Initially, personally recovered malware samples were documented and shared via a blog called 'The Swiss Security Blog', which paved the way for the emergence of abuse.ch as it is known today. Subsequently, multiple platforms used to track different malware were created on the website to help participants fight cybercrime. A community and a network of trust slowly emerged behind abuse.ch, which helped feed the different datasets of the different projects. Today, abuse.ch is a web-based platform specialized in open-source threat intelligence and is composed of multiple projects used by many public and private actors to protect themselves and/or their clients against cyberthreats. Most of the threat information is generated by the community on four important platforms:

- URLHaus, launched in 2018, is a project with the goal of sharing URLs used for malware distribution.
- MalwareBazaar, launched in 2020, is a project that aims to collect and share malware samples—not easily accessible before this initiative.
- ThreatFox, launched in early 2021, is an open-source threat intelligence platform used to share, store, and collaborate on cybersecurity incidents, known as IoCs. Despite its young age, ThreatFox already has an active community.
- YARAify, launched in June, 2022, is a project from abuse.ch that allows anyone to scan suspicious files such as malware samples or process dumps against a large repository of YARA rules. With YARAhub, the platform also provides a structured way for sharing YARA rules with the community.

Overall, the goal of these platforms is to facilitate access to threat information by removing as many barriers to sharing as possible and to reduce executional costs, as described in [17]. Therefore, there is no need of a platform account to access the data.

3.2 ThreatFox Dataset Description

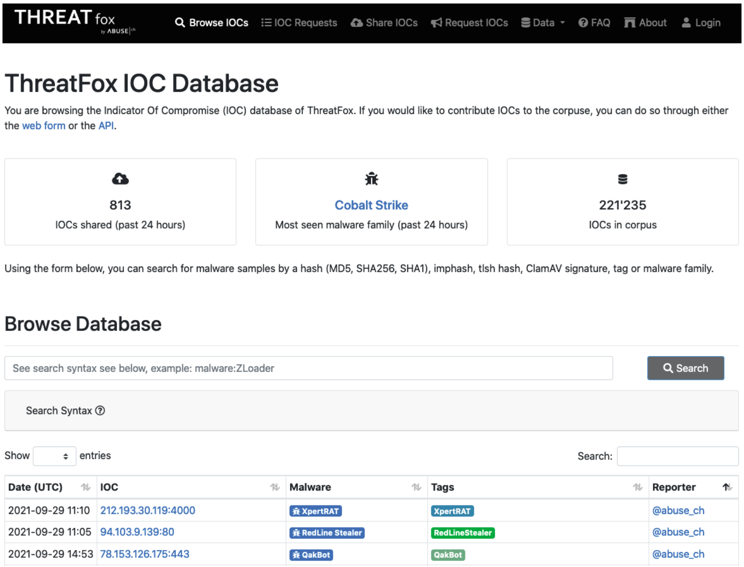
ThreatFox is an open-data threat intelligence platform, launched in March 2021 and operated by abuse.ch, on which participants can collaborate by sharing artifacts of cybersecurity incidents in the form of IoCs. These IoCs contain basic

[5] Roman Hüssy is a research associate at the Bern University of Applied Sciences (BFH) https://www.bfh.ch/fr/la-bfh/personnes/7364w3jin4k5/. The abuse.ch project has been hosted by this institution since June 1, 2021.

information, such as a URL, IP address, or a hash of a malware sample (see Table 1 for an overview of relevant fields), which can be reused by other investigators to discover the same evidence on their systems. Therefore, sharing these data with as many users as possible through sharing platforms is important. However, other platforms are either closed (only selected participants can share and receive IoCs), fee-based, or require some form of registration. From an economic perspective, these platforms can be considered as a club good (excludable and nonrivalrous). In contrast, ThreatFox is one of the first platforms to offer a public good approach (nonexcludable and nonrivalrous) with its free and open-data mindset. Moreover, ThreatFox attempts to minimize barriers and create new incentives for IoC sharing. Consequently, the user interface and API used to retrieve IoCs do not require any form of registration, and these IoCs can be downloaded in the most used formats, e.g. JSON, CSV, MISP events, and others.

ThreatFox was also built on a pre-existing community from former abuse.ch platforms, such as URLHaus or MalwareBazaar, and has used all of the experiences and best practices to create a new platform that encourages sharing, such as a credit system used to reward the user for sharing an IoC.

In this article, ThreatFox data published from March 8, 2020 to July 4, 2022, is used for the analysis. During this period, 767,396 IoCs were published by anonymous users and 106 identifiable users, also called reporters. These IoCs are accessible via a web interface (see Fig. 1) or via API requests to be easily accessible. Roman Hüssy from abuse.ch kindly provided additional data (e.g., credits of the IoCs) that are not directly available on the website.



ThreatFox IOC Database

You are browsing the Indicator Of Compromise (IOC) database of ThreatFox. If you would like to contribute IOCs to the corpuse, you can do so through either the web form or the API.

813
IOCs shared (past 24 hours)

Cobalt Strike
Most seen malware family (past 24 hours)

221'235
IOCs in corpus

Using the form below, you can search for malware samples by a hash (MD5, SHA256, SHA1), imphash, tlsh hash, ClamAV signature, tag or malware family.

Browse Database

See search syntax see below, example: malware:ZLoader

Date (UTC)	IOC	Malware	Tags	Reporter
2021-09-29 11:10	212.193.30.119:4000	XpertRAT	XpertRAT	@abuse_ch
2021-09-29 11:05	94.103.9.139:80	RedLine Stealer	RedLineStealer	@abuse_ch
2021-09-29 14:53	78.153.126.175:443	QakBot	QakBot	@abuse_ch

Fig. 1. Web interface of ThreatFox

The important fields in the ThreatFox dataset are visible in Table 1.

Table 1. Important fields of the ThreatFox dataset.

Field name	Description
IoC value	The value of the IoC. This can be a hash (sha1, sha256, sha3 or md5), a url, a domain, or an ip:port pair
IoC type	The corresponding type of the IoC (hash, url, domain, or ip:port pair)
Threat type	The threat type of the IoC (payload, command & control)
Malware	The name of the corresponding malware
Timestamp	IoC post time (UTC+0)
Confidence level	Value between 25 and 100 characterizing the confidence of the contributor toward the shared IoC
Reference	Link to a web page (often tweets or MalwareBazaar pages) that gives more information and features about the IoC
Reporter	Pseudo (Twitter account) of the user who shares the IoC
Anonymous	Whether or not the IoC is shared anonymously
Credits earned	Number of credits earned by the reporter with the posted IoC

3.3 Methodology

Fitting methods are used on the dataset to highlight the dynamics of organizational integration and the cumulative dynamics of event production. In particular, the data are fitted using the most probable growth functions, starting with a visual inspection: (i) a linear growth function takes the form: $y = a \cdot x + b$, while (ii) a superlinear growth function is represented $y = x^{\beta} \Rightarrow log(y) = \beta \cdot log(x)$ for $\beta > 1$. Linear and superlinear relations are most commonly found in open collaboration platforms [21–23]. Once identified, the function is calibrated according to the data in this study by non-linear least squares.

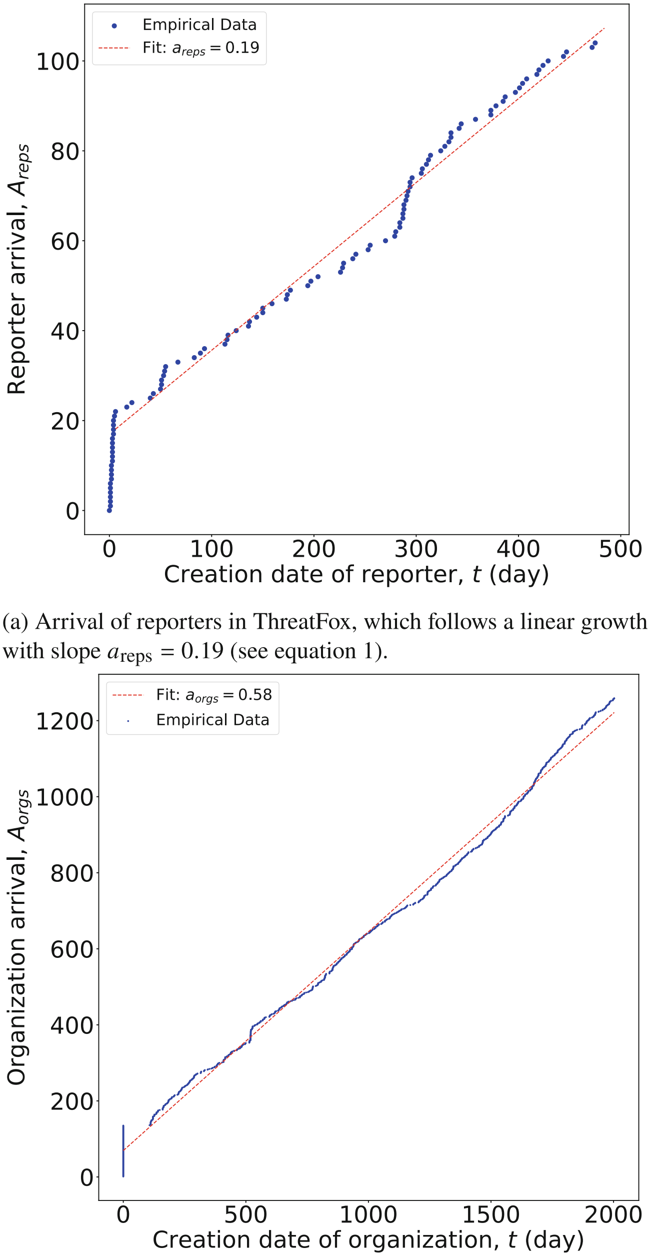
4 Results

4.1 Onboarding Dynamics of Reporters

When considering information sharing, a key aspect is the number of participants in the collective action process because each participant is expected to contribute to the collective good. It is not the case here, since 10% of users contribute to around 98% of the IoCs, which is highly skewed. We can thus conclude that, although it is admitted, free-riding [24] (e.g., leeching) occurs.

From that, we produce the Fig. 2a that shows the cumulative number of new reporters as a function of the date on which the first organization was created once the platform was launched. ThreatFox is based on an existing community, which is why a massive arrival of new reporters in Fig. 2a in the first five days is observed. This arrival corresponds to onboarding reporters who were already present and active on previous abuse.ch platforms. The arrival of new reporters is observed to indicate a slow, linear growth after the first five days (see equation (1)), which could mean that the platform's attractiveness remains the same over time. Joining curve of reporters' arrival is:

$$A_{\text{reps}}(t) \sim a_{\text{reps}} \cdot t \quad \text{with} \quad a_{\text{reps}} = 0.19 \tag{1}$$



(a) Arrival of reporters in ThreatFox, which follows a linear growth with slope $a_{\text{reps}} = 0.19$ (see equation 1).

(b) Arrival of reporters in Computer Incident Response Center Luxembourg (CIRCL) MISP, which follows a linear growth with slope $a_{\text{orgs}} = 0.58$ (Data taken from Gillard et al. paper [7]).

Fig. 2. Onboarding dynamics of new reporters in ThreatFox and CIRCL MISP.

In Fig. 2b, a similar pattern is observed with the CIRCL MISP community, which was also built on a pre-established community.

Goldenberg and Dean [25] argued that successful information sharing depends on a combination of a common mission, a shared identity, familiarity, and trust. Trust facilitates data sharing, which in turn enhances trust itself, and is thus necessary on information sharing platforms [16]. Indeed, data on these platforms will be used, for example, in monitoring systems; therefore, they must be reliable. Lack of trust within the community can lead to collaboration issues [26] (e.g., sharing information with rivals could improve their competitive position).

Trust cannot come from anywhere; it must be built through pre-existing personal relationships [16] developed over time through formal and informal networks [25]. Thus, information sharing must rely on a core of trusted individuals, who interact formally for the information sharing process but also informally to build trust among themselves.

In this way, the suggestion is made that pre-established communities create a foundation of trust that then encourages newcomers to become more involved.

4.2 Publication Dynamics of Indicator of Compromise

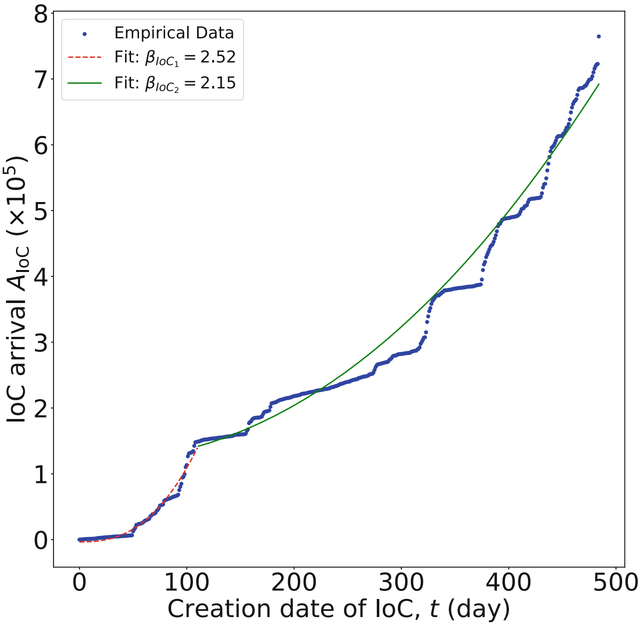


Fig. 3. Arrival of IoCs in ThreatFox, which first follows a superlinear growth function with β_{IoC_1} = 2.52 (see Eq. 2) for the first 111 d, then follows a superlinear growth function with β_{IoC_2} = 2.15 (see Eq. 3). The repeated pattern of strong/flat growth in the figure could be explained by the grouped arrival of new users from other pre-existing communities of trusted networks on abuse.ch.

Figure 3 shows the cumulative number of IoCs created over time. The publication of IoCs is composed of two distinct growth phases.

When a new sharing platform is created by abuse.ch, an existing trusted community with 'power users' will migrate swiftly to it, in order to enact the first sparks of collective intelligence dynamics.

During the first 111 d, the number of shared IoCs is observed to grow superlinearly (faster than the linear function) with the number of days since the opening of ThreatFox (and, thus, with the number of reporters) (see equation (2)).

$$A_{\mathrm{IoC}}(t) \sim t^{\beta_{\mathrm{IoC}_1}} \quad \text{with} \quad \beta_{\mathrm{IoC}_1} = 2.52 \tag{2}$$

This first step could consist of transferring the events already collected by the different reporters.

After the first 111 days, the number of IoCs published begins a second slower growing phase, also superlinear. This behavior shows a strong positive dynamics, as indicated in Eq. (3). The majority of the IoCs published in this section could correspond to new cyber events that are shared after their detection.

$$A_{\mathrm{IoC}}(t) \sim t^{\beta_{\mathrm{IoC}_2}} \quad \text{with} \quad \beta_{\mathrm{IoC}_2} = 2.15 \tag{3}$$

As reported by Müller et al. [27], a high degree of social interaction is positively associated with the quantity, quality, and frequency of information sharing. Thus, the current good dynamics of information sharing in ThreatFox is suggested as being related to the ever-growing IoC database, which was initially populated by the pre-existing community.

4.3 Credit System Rewarding Dynamics

Easy and free sharing is one of the great incentives in sharing platforms [16]. However, without extrinsic motivations, such as money or rewards, the motivation to share decreases over time [28]. All of these motivations do not have the same impact [29]. Titmuss stated that monetary compensation can destroy the sense of civic duty and produce a net decrease in action with respect to acts of benevolence toward others, such as blood donations [30]. Gneezy and Rustichini found that increased monetary rewards lead to better performance but that small rewards are often less effective than not using a reward at all. They explained their results by stating that the addition of monetary rewards reduces intrinsic motivation. To create incentives to information sharing platforms, alternatives to monetary rewards can be explored.

Purely symbolic rewards can affect user behavior on information sharing platforms, such as Wikipedia. Jana Gallus [31] showed that such rewards can be powerful motivators, despite the fact that they do not provide material goods or benefits to the user. This reward system consists of badges given to new users if they are active on the platform. These symbols were observed to increase the number of active contributors and their number of contributions over a long period. These rewards allow users to identify as members of the Wikipedia community and gain a reputation and recognition from other community members, creating a new motivation to share.

ThreatFox introduces a credit system used to reward the sharing of IoCs in which a user earns credits when he shares an IoC. An IoC earns more credits if requested by another user. However, the credits are only symbolic and cannot be used to buy anything yet. Hence, they rather support the gamification of sharing. They are only present in the list of 'richest reporters', which is available on the ThreatFox website. Wikipedia used a similar leaderboard. Gallus et al. [31] showed that such leaderboards have a positive impact on information sharing.

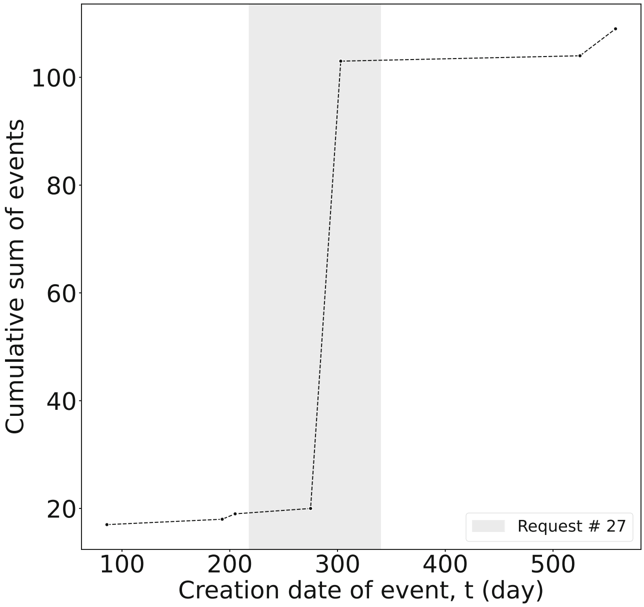


Fig. 4. Cumulative sum of the number of domains linked to Formbook botnet shared. The corresponding IoC request is represented by gray area.

Of the 105 IoCs requests made since the launch of ThreatFox, only 62 received at least one response, which results from the low number of users sharing information and the fact that IoCs are not items that the user can easily obtain or create. Instead, they are items that the user already has (e.g., because of previous incidents) or has recently received. However, the credit system seems to have—for some attacks and some users—some impact. Another explanation could be that there is an underlying cause, such as an ongoing attack, resulting in an increase of requests and share IoCs simultaneously. Indeed, Fig. 4 shows that 80% of the sharing of domains linked to Formbook botnet malware are made when a request is made. The apparition of the IoC request (IoC request #27) is linked with the apparition of a peak in the number of IoC shared. ThreatFox is still a young platform with few participants and little sharing; therefore, repeating this analysis in the future is recommended, when more data are available for analysis. However, the credit system is a good, free incentive and could be promoted more on abuse.ch website.

5 Discussion and Recommendations

Building collaborative cybersecurity through information sharing has been considered a critical path to keeping up with increasingly pervasive and innovative cyberthreats [11,18]. To formally organize such information sharing activity, a number of online and more or less open platforms have been set up [11]. ThreatFox is one of the most recent platforms to be launched and still manages to stay attractive to newcomers thanks to the pre-existing community and the positive community dynamic described in Sects. 4.1 and 4.2. This dynamic should be maintained at this degree of intensity for the platform to remain as relevant as it is now. The study results suggest that sharing platforms should be built on existing communities in which trust has already been established by numerous interactions between individuals, as shown in the case of ThreatFox, which is built on the success and established trust of URLHaus and MalwareBazaar. In contrast, sharing platforms created with a purely technical focus tend to underperform [32]. Indeed, a socio-technical approach taking into account human behavior is essential to optimize the chances of success [9].

5.1 Reduction of Executional Costs in ThreatFox

According to behavioral theory, humans are loss averse [33–36]; that is, they try harder to avoid economic losses than to realize economic benefits. An exchange relationship might involve significant transaction costs, also called 'executional costs', such as the time, material, or financial resources, that an individual must commit before an exchange can take place [37]. Therefore, if information sharing takes too long, is too laborious, or requires too much effort, individuals likely avoid the necessary resource commitments and, thus, reduce or terminate their participation [38]. Yan et al. argued, for example, that knowledge sharing is inhibited when it is time consuming [39]. The European Union's cybersecurity agency also warns that an abundance of procedures blocks information sharing activity [40]. Thus, high executional costs are likely to deter users of a sharing platform from participating [16]. ThreatFox attempts to minimize the execution cost because everyone can share without too much time, hardware, or financial resources (simple API requests, no registration required, etc.).

5.2 Anonymity in ThreatFox

One of the biggest obstacles to sharing for companies and individuals is the threat of reputational damage [41] and privacy issues. Indeed, if an incident becomes public, customers' trust in an entity can be severely affected, resulting in a loss of customers and, thus, revenue. One solution is to anonymize the shared data, an option offered by some information sharing platforms, such as ThreatFox or AlienVault OTX.

As Murdoch and Leaver [18] pointed out, members of sharing communities sometimes have to hide their identities using anonymity-enabling design principles because of legal restrictions (e.g., GDPR), public relations concerns, or the sensitive nature of the information. However, anonymizing the contributor can lead to a deterioration of trust in the shared data because its origin cannot be confirmed [42]. Furthermore, anonymity is not enough to allow a cyber event to be shared. In fact, some IoCs might still contain information about a company and its users [43] and must be shared with care as defined by the TLP[6] to avoid legal issues (e.g., GDPR).

Some platforms, such as ThreatFox, only allow data to be published when marked with the TLP:WHITE flag to avoid problems; this flag means that the data cannot be abused. Others, such as MISP, create special closed communities that share these data. In both cases, these solutions do prevent the data from being abused. Therefore, events containing personally identifiable information (PII) cannot be shared freely, which creates a form of censorship. Before these events can be openly shared they must first be modified into something shareable by anonymizing the content of the shared data [43,44]. The U.S. National Institute of Standards and Technology (NIST) published a list of recommendations in its Guide to Cyber Threat Information Sharing [45] to maximize anonymity for contributors by removing sensitive information from the shared data that is not necessary for describing an incident (e.g., masking IP/MAC addresses in network packets, masking names in phishing email samples, masking user identifiers in application logs). The problem with this technique is that PII identification, extraction, and obfuscation might be incomplete, which can lead to unauthorized disclosure of intellectual property or trade secrets [45]. Disclosing this information could result in financial loss, violation of sharing agreements, legal action, or reputational damage to an organization.

In some cases, data cannot be anonymized without losing the utility of sharing (e.g., because too many fields are deleted); alternatively, anonymization via a third party might not be reliable for everyone. In these cases, alternatives to sharing information exist, such as distributed threat intelligence learning. This solution was explored by [46,47], who attempted to find a compromise in the information-sharing trade-offs between the benefits of improved threat response capability and the drawbacks of disclosing national security-related information to foreign agencies or institutions. Their solution enables secure collaboration with valuable, sensitive data that are not normally shared. Each institution

[6] The TLP was created in the early 2000s by the U.K.'s National Infrastructure Security Coordination Centre to encourage the sharing of sensitive information between individuals and organizations in a reliable and controlled manner. The data are classified into one of four classes that regulate the conditions of its disclosure https://www.cisa.gov/tlp.

retains full control of its data records, which never leave the platform's secure perimeter, whereas computations are protected by efficient and highly scalable multi-party homomorphic encryption techniques [48]. However, this solution is not flawless because the data are no longer shared and can thus only be used for specific computations. This solution also adds some overhead to the complexity of the computations, which is not addressed in this article.

5.3 Implications for ThreatFox and Generalization to Other Sharing Platforms

In a broad sense, this study has implications for the design of CIS platforms. Fundamentally, the success of abuse.ch is argued as being related to its ease of use (reduction of the main barrier 'executional cost' to sharing described in [49]), its privacy, and the trust created over the years by an existing community. Thus, when a new sharing platform is created by abuse.ch, an existing trusted community with 'power users' will migrate swiftly to it and bring the necessary critical mass to enact collective intelligence dynamics. For instance, in June, 2022, abuse.ch launched a new platform called 'YARAify'[7], which allows anyone to scan suspicious files such as malware samples or process dumps against a large repository of YARA rules. With YARAhub, the platform also provides a structured way for sharing YARA rules with the community.

These fundamental institutional rule may explain the success of abuse.ch and might be generalizable to other platforms. In a narrower way, this research also has direct implications for abuse.ch. The first recommendation is that the platform improve its statistical processing by collecting its data in the following ways.

Indeed, at the moment, the IoCs that earn the most are those requested by other users. This system seems to have some limitations because users cannot create specific IoCs on demand. The reward system could be improved by introducing new ways to earn credits, such as a system that awards credits based on the exclusivity of an IoC. The double-blind peer review process for IoCs could be improved with more reviewers to ensure quality and to improve user confidence across IoCs. This activity could also be rewarded with a credit system to encourage review.

The 'confidence level' field of shared IoCs (see Table 1), which characterizes the confidence of the contributor in the shared IoC, seems interesting at first sight. However, this field is defined by the contributor, who might have a biased view of the shared object. When possible, the level of trust should be assessed by other ThreatFox users. This field could also be combined with a new 'utility' field, which would represent the number of times a specific IoC has been used by other ThreatFox users to detect cyberattacks in their system. This new field could be very valuable in measuring the usefulness of an IP and, thus, could detect interesting patterns, such as the existence of a best performance area (inverted U-curve) of sharing communities based on the optimal number of participants in

[7] https://yaraify.abuse.ch.

the sharing community [50]. Next, in the same vein, future work could investigate if a Ringelmann effect exists on cybersecurity sharing platforms [51].

6 Research Agenda

The development of open and free sharing platforms and ThreatFox in particular opens up many research opportunities in the growing field of CIS. This study identified numerous research gaps in the literature, such as a lack of research about patterns on onboarding dynamics, a lack of research about publication dynamics, and a lack of research about reward systems to incentivize information sharing. A ‘universal’ performance indicator to measure the influence of institutional rules on CIS is a remaining gap in the literature. In the future, it would also be interesting to study the abuse.ch community structure with new “community analysis” methods adapted from the International Data-driven Research for Advanced Modeling and Analysis (iDRAMA Lab)[8]. At the technical level, building response incident rate indicators based on ThreatFox data would be interesting. At a socio-technical level, investigating the optimal size of sharing communities and the role of pre-existing communities would be interesting. In the future, it will also be necessary to better characterize the data and the users, for example by identifying associated distributions. Finally, the assumption that ’open’ sharing communities share IoCs better and faster could be benchmarked, such as against MISP data or even other sharing communities, such as COVID-19 large threat intelligence communities [14].

7 Conclusion

This study used an exclusive ThreatFox dataset to describe how pre-existing communities and pre-established networks of trust have an important impact on the success of newly created information sharing platforms. An initial study of ThreatFox’s promising credit system is also provided and some IoC requests were shown to be effective in creating new IoC sharing. However, this incentive, although rarely found on such information sharing platforms, is not mature enough and could be improved to be more impactful. Finally, this study sheds light on the onboarding dynamics of both reporters and a publication of IoCs. Overall, the results of this study provide a better understanding of the institutional rules enacting collective intelligence dynamics in cybersecurity. Finally, a broad agenda is presented for future empirical research in the field of CIS, which is an important activity to reduce information asymmetry between attackers and defenders.

Acknowledgments. We would like to thank abuse.ch and especially its founder, Roman Hüssy, for the technical support and data availability, as well as Enago.com for the English language review. Comments received from Vincent Lenders, Thomas Maillart, Marcus Keupp, Philipp Fischer, William Lacube, Michael Tsesmelis and three anonymous reviewers all helped to improve previous versions of this article.

[8] https://idrama.science.

References

1. Laube, S., Böhme, R.: Strategic aspects of cyber risk information sharing. In: ACM Computing Surveys 50.5 (2017). https://doi.org/10.1145/3124398
2. Meier, R., et al.: FeedRank: a tamper- resistant method for the ranking of cyber threat intelligence feeds. In: 2018 10th International Conference on Cyber Conflict (CyCon) (2018). https://doi.org/10.23919/CYCON.2018.8405024
3. Sridhar, K., et al.: Cybersecurity information sharing: analysing an email corpus of coordinated vulnerability disclosure. In: The 20th Annual Workshop on the Economics of Information Security (2021)
4. Gal-Or, E., Ghose, A.: The economic incentives for sharing security information. Inf. Syst. Res. **16**(2) (2005). https://doi.org/10.1287/isre.1050.0053
5. EricWeiss, N.: Legislation to facilitate cybersecurity information sharing: economic analysis. In: Econ. Anal. (2015)
6. He, M., Devine, L., Zhuang, J.: Perspectives on cybersecurity information sharing among multiple stakeholders using a decision-theoretic approach: cybersecurity information sharing. Risk Anal. **38**(2) (2018). https://doi.org/10.1111/risa.12878
7. Gillard, S., et al.: Efficient collective action for tackling time-critical cybersecurity threats. Tech. rep. arXiv:2206.15055. [physics] type: article. arXiv, (2022)
8. Wagner, C., et al.: MISP: the design and implementation of a collaborative threat intelligence sharing platform. In: Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security. (2016). https://doi.org/10.1145/2994539.2994542
9. Percia David, D., Matthias Keupp, M., Mermoud, A.: Knowledge absorption for cyber-security: the role of human beliefs. Comput. Human Behav. **106** (2020). https://doi.org/10.1016/j.chb.2020.106255
10. Sohrabi Safa, N., Von Solms, R.: An information security knowledge sharing model in organizations. Comput. Hum. Behav. **57** (2016). https://doi.org/10.1016/j.chb.2015.12.037
11. World Economic Forum. Cyber Information Sharing: Building Collective Security. Tech. Rep. (2020)
12. Mermoud, A., MatthiasKeupp, M., Percia David, D.: Governance models preferences for security information sharing: an institutional economics perspective for critical infrastructure protection. Critical Inf. Infrastruct. Secur. (2019). https://doi.org/10.1007/978-3-030-05849-4_14
13. Oliver, P.: Rewards and punishments as selective incentives for collective action: theoretical investigations. Am. J. Sociol. **85**(6) (1980). Publisher: The University of Chicago Press. https://doi.org/10.1086/227168
14. Bouwman, X., et al.: Helping hands: measuring the impact of a large threat intelligence sharing community. In: Proceedings of the 31st USENIX Security Symposium (2022)
15. Poundstone, W.: Prisoner's Dilemma/John Von Neumann, game theory and the puzzle of the bomb. Anchor (1993)
16. Mermoud, A.: Three articles on the behavioral economics of security information sharing: a theoretical framework, an empirical test, and policy recommendations". PhD Thesis. Université de Lausanne, Faculté des hautes études commerciales, (2019)
17. Mermoud, A., et al.: Incentives for human agents to share security information: a model and an empirical test. In: 17th Workshop on the Economics of Information Security (WEIS) (2018)

18. Murdoch, S., Leaver, N.: Anonymity vs trust in cyber-security collaboration. In: Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security (2015). https://doi.org/10.1145/2808128.2808134
19. Burger, E., et al.: Taxonomy model for cyber threat intelligence information exchange technologies. In: Proceedings of the 2014 ACM Workshop on Information Sharing & Collaborative Security - WISCS '14 (2014). https://doi.org/10.1145/2663876.2663883
20. Miorandi, D., Maggi, L.: Programming Social Collective Intelligence. In: IEEE Technology and Society Magazine, Conference Name: IEEE Technology and Society Magazine vol. 33. no. 3 (2014). https://doi.org/10.1109/MTS.2014.2345206
21. Sornette, D., Maillart, T., Ghezzi, G.: HowMuch is the whole really more than the sum of its Parts? 1 + 1 = 2.5: Superlinear Productivity in Collective Group Actions. In: PLoS ONE 9.8 (2014). https://doi.org/10.1371/journal.pone.0103023
22. Scholtes, I., Mavrodiev, P., Schweitzer, F.: From aristotle to ringelmann: a large-scale analysis of team productivity and coordination in open source software projects. Tech. Rep. Gesellschaft für Informatik e.V. (2016)
23. Muri, G., et al.: Collaboration drives individual productivity. In: Proceedings of the ACM on Human-Computer Interaction 3.CSCW (2019). https://doi.org/10.1145/3359176
24. Anesi, V.: Moral hazard and free riding in collective action. Soc. Choice and Welfare **32**(2) (2008). https://doi.org/10.1007/s00355-008-0318-8
25. Goldenberg, I., Dean, W.: Enablers and barriers to information sharing in military and security operations: lessons learned. In: Enablers and Barriers to Information Sharing in Military and Security Operations: Lessons Learned (2017). https://doi.org/10.1007/978-3-319-42819-2_16
26. Koepke, P.: Cybersecurity information sharing incentives and barriers. In: Working Paper CISL (2017)
27. Müller, J.M., Veile, J.W., Voigt, K.-I.: Prerequisites and incentives for digital information sharing in industry 4.0 an international comparison across data types". In: Computers & Industrial Engineering 148 (2020). https://doi.org/10.1016/j.cie.2020.106733
28. Zibak, A., Simpson, A.: Cyber threat information sharing: perceived benefits and barriers. In: Proceedings of the 14th International Conference on Availability, Reliability and Security (2019). https://doi.org/10.1145/3339252.3340528
29. Wagner, T.D., et al.: Cyber threat intelligence sharing: survey and research directions. Comput. Secur. **87** (2019). https://doi.org/10.1016/j.cose.2019.101589
30. Mellström, C., Johannesson, M.: Crowding out in blood donation: was titmuss right? J. Eur. Econ. Assoc **6**(4) (2008). https://doi.org/10.1162/JEEA.2008.6.4.845
31. Gallus, J.: Fostering Public good contributions with symbolic awards: a large-scale natural field experiment at wikipedia. Manage. Sci. **63**(12)(2017). https://doi.org/10.1287/mnsc.2016.2540
32. Stojkovski, B., et al.: What is in a cyber threat intelligence sharing platform? A mixed-methods user experience investigation of MISP. Ann. Comput. Secur. Appl. Conf. (2021). https://doi.org/10.1145/3485832.3488030
33. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. In: World Scientific Handbook in Financial Economics Series, vol. 4 (2013). https://doi.org/10.1142/9789814417358_0006
34. Tversky, A., Kahneman, D.: Loss aversion in riskless choice: a reference-dependent model. Q. J. Econ. **106**(4) (1991). https://doi.org/10.2307/2937956

35. Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertain. **5**(4) (1992). https://doi.org/10.1007/BF00122574
36. Tom, S.M., et al.: The neural basis of loss aversion in decision-making under risk. Science **315**(5811) (2007). https://doi.org/10.1126/science.1134239
37. Williamson, O.E.: The economics of organization: the transaction cost approach. Am. J. Sociol. **87**(3) (1981). https://doi.org/10.1086/227496
38. Luiijf, E., Klaver, M.: On the sharing of cyber security information. In: Critical Infrastructure Protection IX, vol. 466. Series Title: IFIP Advances in Information and Communication Technology (2015). https://doi.org/10.1007/978-3-319-26567-4_3
39. Yan, Z., et al.: Knowledge sharing in online health communities: a social exchange theory perspective. Inf. Manage. **53**(5) (2016). https://doi.org/10.1016/j.im.2016.02.001
40. Alkalabi, W., Simpson, L., Morarji, H.: Barriers and incentives to cybersecurity threat information sharing in developing countries: a case study of Saudi Arabia. In: 2021 Australasian Computer Science Week Multiconference (2021). https://doi.org/10.1145/3437378.3437391
41. Mavroeidis, V., Bromander, S.: Cyber threat intelligence model: an evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. In: 2017 European Intelligence and Security Informatics Conference (EISIC) (2017). https://doi.org/10.1109/EISIC.2017.20
42. Danezis, G., et al.: Privacy and data protection by design - from policy to engineering. In: arXiv:1501.03726 [cs] (2014). https://doi.org/10.2824/38623
43. Pang, R., et al.: The devil and packet trace anonymization. ACM SIGCOMM Comput. Commun. Rev. **36**(1) (2006). https://doi.org/10.1145/1111322.1111330
44. Fathi, Z., Rafsanjani, A.J., Habibi, F.: Anon-ISAC: anonymity-preserving cyber threat information sharing platform based on permissioned blockchain. In: 2020 28th Iranian Conference on Electrical Engineering (ICEE) (2020). https://doi.org/10.1109/ICEE50131.2020.9261029
45. Johnson, C.S., et al.: Guide to cyber threat information sharing. Tech. rep. NIST SP 800–150. National Institute of Standards and Technology (2016). https://doi.org/10.6028/NIST.SP.800-150
46. Froelicher, D., et al.: Truly privacy-preserving federated analytics for precision medicine with multiparty homomorphic encryption. Nat. Commun. **12**(1) (2021). Publisher: Nature Publishing Group
47. Trocoso-Pastoriza, J., et al.: Orchestrating collaborative cybersecurity: a secure framework for distributed privacy-preserving threat intelligence sharing. In: ACM Digital Threats: Research and Practice, Special Issue on Information Sharing (2023)
48. Froelicher, D., et al.: DRYNX: decentralized, secure, verifiable system for statistical queries and machine learning on distributed datasets (2020). arXiv:1902.03785
49. Mermoud, A., et al.: To share or not to share: a behavioral perspective on human participation in security information sharing. J. Cybersecurity 5.tyz006 (2019)
50. Dejean, S., Pénard, T., Suire, R.: Olson's Paradox Revisited: an empirical analysis of incentives to contribute in p2p file-sharing communities. SSRN Scholarly Paper (2010). https://doi.org/10.2139/ssrn.1299190
51. Maillart, T., Sornette, D.: Aristotle vs. Ringelmann: on superlinear production in open source software. Physica A: Stat. Mech. Appl. **523** (2019). https://doi.org/10.1016/j.physa.2019.04.130