Motivation-based Attacker Modelling for Cyber Risk Management: A Quantitative Content: Analysis and a Natural Experiment

Article · December 2021

DOI: 10.26735/NMMD9869

CITATIONS
0

READS
330

3 authors:

Florian Kaiser
Karlsruhe Institute of Technology
10 PUBLICATIONS **17** CITATIONS
SEE PROFILE

Marcus Wiens
Technical University Bergakademie Freiberg; Karlsruhe Institute of Technology (KIT)
81 PUBLICATIONS **290** CITATIONS
SEE PROFILE

Frank Schultmann
Karlsruhe Institute of Technology
222 PUBLICATIONS **5,161** CITATIONS
SEE PROFILE

Journal of Information Security & Cybercrimes Research 2021; Volume 4 Issue (2), 132-147

Naif Arab University for Security Sciences
Journal of Information Security & Cybercrimes Research
مجلة بحوث أمن المعلومات والجرائم السيبرانية
https://journals.nauss.edu.sa/index.php/JISCR

JISCR

Motivation-based Attacker Modelling for Cyber Risk Management: A Quantitative Content Analysis and a Natural Experiment

Florian Klaus Kaiser*, Marcus Wiens, and Frank Schultmann

Karlsruhe Institute of Technology, Karlsruhe, Germany.

Received 10 Oct. 2021; Accepted 30 Oct. 2021; Available Online 30 Dec. 2021

Abstract

Cyber-attacks have a tremendous impact on worldwide economic performance. Hence, it is vitally important to implement effective risk management for different cyber-attacks, which calls for profound attacker models. However, cyber risk modelling based on attacker models seems to be restricted to overly simplified models. This hinders the understanding of cyber risks and represents a heavy burden for efficient cyber risk management. This work aims to forward scientific research in this field by employing a multi-method approach based on a quantitative content analysis of scientific literature and a natural experiment. Our work gives evidence for the oversimplified modelling of attacker motivational patterns. The quantitative content analysis gives evidence for a broad and established misunderstanding of attackers as being illicitly malicious. The results of the natural experiment substantiate the findings of the content analysis. We thereby contribute to the improvement of attacker modelling, which can be considered a necessary prerequisite for effective cyber risk management.

I. Introduction

The digitalization of all parts of our lives has led to a huge influx of digital technologies into society and the economy. Smartphones, smartwatches, and digital applications are used by people around the globe, and they bring great comfort. Mobile work, video conferencing, industrial robots, driverless transport systems, and artificial intelligence (among others) have further changed industrial production and increased its efficiency. Also, online shopping and online marketing have changed how markets operate. However, these technological advances also come with high risks. Cyberattacks have a continually huge impact on companies worldwide.

Although there are increasing efforts to bring effective cyber risk management practices into place, many companies and individuals struggle, as the field of cyber risk management is highly dynamic. Attackers adapt strategically to countermeasures implemented by risk managers. This has led to a "dynamic cat and mouse game" [1] between attackers and risk managers.

Generating attack hypotheses is often based on highly sophisticated statistical approaches. However, even those sophisticated approaches may fail to deliver hypotheses of sufficient accuracy.

Keywords: Information security, Threat modelling, Multi-methods, Natural experiment, Quantitative literature analysis.

Production and hosting by NAUSS

* Corresponding Author: Florian-Klaus Kaiser
Email: florian-klaus.kaiser@kit.edu
doi: 10.26735/NMMD9869

One pitfall of these approaches to hunt threats is that they frequently lack an understanding of the attackers and their motivations or include immature threat models. As it is questionable whether attack history statistical approaches alone are feasible to quantify cyber risks (as the attackers may be intelligently and strategically adapting to defender strategies and the environment), it is of particular scientific importance to understand the motivational background that incentivizes an attack and determines cyber risks in order to support and adjust these statistical approaches. This involves understanding the motivation of attackers, their strategic considerations in selecting their goals, and their attack strategies. Such an understanding leads to a higher level of cyber threat detection maturity [2], [3].

Furthermore, the same weaknesses can be observed for cyber risk quantification. If attackers are considered to be inherently malicious, no system can be secured. However, if attackers are seen as utility-maximizing actors, efficient cyber-security can be achieved when disutility or necessary effort to circumvent the implemented countermeasures exceed the utility an attacker believes to derive from a successful attack. Consequently, cyber risk management and the engineering of secure systems calls for a clear understanding of attackers' motivation.

Criminal psychology supports the importance of modelling attackers as intelligent attackers within their incentivizing backgrounds [4]. Within psychological research related to cyber-crime, it is common knowledge that "people from various socio-economic, intellectual, and cultural backgrounds participate in a wide range of cybercrimes for many different reasons" [5]. Efficient threat modelling is hence assumed to be able to increase the efficiency of attack hypotheses generation together with risk quantification and, consequently, defensive mechanisms on cyber-attacks. To close this gap, this work relates efficient threat modelling to the motivations of attackers in the spirit of game theory, considering the strategic behavior of attackers as intelligent actors [6], [7].

Therefore, we employ a quantitative content analysis to investigate the motivations of attackers and compare this knowledge with established classifications of attackers (attacker models). In a second step, the COVID-19 pandemic is used to gain further insight into the empirical reliability of attacker motivations and models based on these. The COVID-19 pandemic has already been used for various natural experiments [8]. However, to the best of the authors' knowledge, no study has so far utilized the COVID-19 pandemic as a natural experiment for investigating the reliability of attacker models, especially the importance of specific motivational patterns and incentives. For doing so, we rely on explorative data analysis and compare the effects of COVID-19 with what would be expected from theoretical considerations. Furthermore, we describe a new framework based on psychological motivation theory to describe cyber threat agents and produce a continuous definition of different attacker attributes. This reduces the disadvantages of a combined observation of several attributes and archetypical classifications and is intended to enable the proficient use of game theoretical modelling in cyber risk quantification.

The paper is structured as follows: Section 2 presents the theoretical background, and section 3 provides a quantitative content analysis of scientific publications. The latter helps to identify recent attacker typologies and to give an overview of motivational patterns existing in scientific literature. Section 4 provides a natural experiment, which is used to investigate the fit of currently used attacker classifications and practices. Section 5 presents the conclusion.

II. Theoretical Background

A. Attacker Motivations and Models

With respect to cyber-attacks, recent academic publications show a clear consensus that there is a wide variety of different motives [9], [10]. Li [9] contributes to the literature by providing a review of different motivations for cyberattacks. The work includes a hacker ethic (ideological motivation), providing a justification based on the perceived right for a free flow of information. Furthermore, hackers are motivated by various motives, including self-expression, curiosity, practicing and showing off programming skills, financial gains, search for social

acceptance, and many more [9]. In addition, recent academic research includes works, which explicitly target the importance of the different motivational patterns. Fötinger and Ziegler [11] used the data of 599 people from the German Federal Bureau of Criminal Investigation. The study revealed economic motives to be the most significant. Woo et al. [12] based their investigations on a database of 462 defaced web pages and analyzed the motivation that could be derived from the content. They revealed ideological reasons, fun, statements about hackers' skills, and romantic statements as relevant motives. Thycotic Software Ltd [13] presented a survey based on 127 self-identified hackers. They reveal that the most important motivation for hackers is fun and thrill-seeking followed by "social consciousness or moral compass", financial gain, and notoriety. Madarie [14] analyzed hacking behaviors and motivations in a sample of 65 male hackers and potential hackers. The study identified the high importance of intellectual challenge and curiosity as motivation.

These different motives can be further clustered. Such clustering assigns the different motivations to larger groups, e.g., attackers striving for financial benefits, the destructive aim to cause damage (maliciousness), gaining knowledge, and seeking pleasure or notoriety within a community [15]. Attackers often vary significantly regarding resources, motivational patterns, and other factors. The diversity of the attacker motivation landscape can also be seen within classifications of attackers. When mapping this diversity into the classification system, important rules of criterion purity have often been broken (e.g., attacker motivation and attacker resources were linked). This has led to the establishment of classification systems, which unify different classification criteria and do not allow a systematic classification of attackers.

Furthermore, when modelling attackers is done there is often a simplistic division between normal users and attackers [16]. Therefore, attackers are often seen as being inherently malicious. Within this chapter, we want to elaborate on the field of attacker motivations and describe the different models used for threat modelling.

Early attempts to differentiate cyber criminals were undertaken at the beginning of the nineties. Landreth [23] provided one of the first categorizations to define the hacker community, developing five categories (Novice, Student, Tourist, Crasher, and Thief) based on the involvement in hacking and their motives.

- Novices are assumed to be young and motivated by fun as they "think of hacking as play, or mischief-making, and not much more than that." [23].
- Students are defined as intelligent, curious, and bored actors searching for cognitive challenges. Their intention is to learn something that they were not aware of before. "A Student would never intentionally damage a system because there's no reason why he should, and there are many good reasons why he shouldn't." [23].
- The Tourist is searching for "adventure(s) or the challenge of solving a puzzle" [23]. Hacking a system is a mental game for Tourists, and they are searching for achievement (thrill of victory).
- Crashers are assumed to operate with the same goals as vandals or troublemakers. They are motivated by making a name for themselves.
- The Thief is a professional and criminal actor and, for the most of the hacker community does not follow the definitions of a hacker at all. It is assumed that they are seeking to benefit. In most cases, this benefit is not directly financial, but rather it is stolen data (e.g., to get a competitive advantage over a competing company).

Although the last two categories, Crackers and Thiefs, are only two kinds of attackers in cyberspace, they have a dominant (negative) impact on the perception of all hackers. This categorization, however, is not consistent with psychological theories of motivation and a lack of discriminatory power. Classifications within one of those categories may become arbitrary for many individuals (e.g., a person that searches for a cognitive challenge could be either classified as a Student or a Tourist). Thus, this classification has severe drawbacks for understanding the landscape of hackers. Although this criticism is valid for all similar classifications

and the classification does not allow for a classification of attackers from different perspectives (e.g., from the perspective of motivation or resources as they are inseparably connected by definition), similar classification systems have frequently been used to describe and classify hackers. Hollinger [24] provides differentiation in the three categories Pirates, Browsers, and Crackers. Chantler [25] furthermore relies on a classification within three categories (Elite, Nymphytes, and Losers & Lamers). Barber [26] introduced a differentiation within the groups of Script-Kiddies, Hackers, and Crackers. Warikoo [10] presents a methodology for profiling cybercriminals based on cyber threat intelligence. In doing so, they extract characteristics through statistical analysis. However, as proposed by them, their cybercriminal profiles combine all these characteristics. This leads to a broad classification into six profiles (Novice, Hacktivist, Cybercriminals, Cybercrime Syndicates, Cybers Spies, and Cyber Terrorists).

Another system of classification was provided by Smith and Rupp [27]. They differentiate between "two basic types of hackers" [27], internal and external offenders. Within these categorizations, there is a strong focus on the available resources (e.g. knowledge and know-how) to the attackers. However, it is still common knowledge today that different motivations can be observed within the cybercriminal community, including "the challenge, the excitement to succeed, (…) pure intellectual satisfaction (…) vengeance, sabotage and fraud" [27]. However, even though this classification in internal and external offenders is still valid today, it only brings limited insights to threat modelling.

Another work relying on disjunctive criteria is the standardized Threat Agent Library (TAL) provided by Intel. These criteria (attacker attributes) were intent, access, outcome, limits, resource, skill level, objective, visibility and motivation [28], [29]. The motivational patterns included in TAL are accidental, coercion, disgruntlement, dominance, ideology, notoriety, organizational gain, personal financial gain, personal satisfaction. However, their derived classification is limited to archetypes of cyber attackers based on unique combinations of these attacker attributes. However, this classification of individuals within archetypes may lead to severe loss of information, diminishing the understanding of cyber-attacks.

Lickiewicz [4] provides a theoretical model profile of a hacker which is based on the motivations of economic incentives, curiosity, boredom, cognitive reasons, and revenge. Additionally, a differentiation between the attacker attributes (resources, intelligence, personality, social abilities, technical abilities and internet addiction) is made. The model furthermore includes environmental and biological factors such as illnesses or genetic burden, which may have an impact on intelligence and personality. Out of the combination of these factors, the model tries to derive the method of attack, the effectiveness of the attack, the methods used, and the mode of operation at the scene [4].

Rogers [30] identified eight different types of attackers, based on a differentiation of technical abilities and the motivational patterns: Novices, Cyber-Punks, Internals, Petty Thieves, Virus Writers, Old Guard hackers, Professional Criminals, and Information Warriors. However, in contrast to previously shown classification schemas of attackers, a continuous differentiation within the two disjunctive dimensions of differentiation was followed to allow a systematic differentiation between the attackers. Rogers [30] furthermore indicated that the differentiation within the eight categories might not be enough to fully understand the cyber threat landscape. However, until now attacker modeling has relied on such limited models which lack differentiation with respect to attacker types.

As shown, there have been different attempts to model attackers and their motivational patterns. However, the literature lacks a defined methodology for clustering and modeling attackers according to psychological motivation [31]. This gap may also be a reason why until now, attack models which make the rather strong assumption that attackers are malicious are the predominant models. Such models often disregard the environment in which the attackers operate (In this sense, an attacker may strive for money while causing harm as a by-product). These typical attacker models that include undefined attacker utility functions are not fully convincing when quantifying cyber risks, as

this form of modelling cyber-attacks collapses to all-powerful attackers exploiting all available flanks of vulnerability [7]. In some recent models, the economic interests and resource restrictions of the attackers have also been included in the attacker modelling. This shows a trend, according to which the importance of correct attacker modelling has been recognized and game theoretical models for quantifying cyber risks gain is in increasing importance. As a result, the models and attack hypotheses based on them are becoming more and more accurate. However, it has not yet been possible to arrive at a comprehensive conceptualization of attacker modeling that makes this differentiation possible despite many efforts.

B. Achievement, affiliation, and power motives in cyber attacks

Psychology and criminology have a long history in research on motivation. In general, psychological insights support differentiation between power motivation, achievement motivation, and motivation through the pursuit of social acceptance [17]. All types of motivations are widespread and can be found in every human being. Therefore, these motivations are also expected to be driving factors of the attackers' actions.

An aggressor can be motivated by the goal of achievement if, e.g., his attack pursues a standard of excellence ("competition with some standard of excellence"), [18], i.e., if he strives to master a task, to do something particularly well, to surpass himself or to prove himself in competition with others, as is already evident in Murray's [19] description of the need for performance (achievement). The central assumption here is that the incentive for action lies exclusively in the enjoyment of the task-related activity itself ("thrill of accomplishment", [20]) and/or in the self-evaluating emotions of success (satisfaction, pride) or failure (shame, dejection). At the core of the performance, the motive is the affective satisfaction from the self-directed coping with performance demands [21]. The flow experience is a central component in understanding the performance motive. The requirements and abilities for an attack must be in balance [20].

In all phases of their lives, people strive to make and maintain interpersonal relationships such as new acquaintances, friendships, partnerships, and family [22]. Cybercriminals are no exception. Rather, the feeling of being socially integrated is considered an important basic human need, the satisfaction of which has a positive psychological effect. A lack of social integration has negative consequences for subjective well-being and physical well-being. Accordingly, the emergence of criminal social underground networks and the emergence of hacker communities have been observed. Attackers can thereby also be motivated to take action to strive for social acceptance within their community, searching for affiliation.

Schultheiss [21] describes the motive of power as the ability to draw satisfaction from physical, mental, or emotional influence over others. Accordingly, human actors have an inherent need for control. This can manifest itself in the form of the power to reward and punish. Here, others are rewarded or punished for their behavior. Ideological, patriotic, or religious aspects can be cited. For the motive of power, the natural incentive is the exertion of influence, which in turn is associated with positive qualities of experience such as a feeling of strength, positive excitement, and pleasure.

C. Attacker Data and Cyber Threat Intelligence

Cyber threat intelligence (CTI) subsumes actionable information about cyber-attacks, including motives, targets, attack strategies, and attacker capabilities and their resources. CTI is based on past events and is shared through various companies. This information enables companies to understand the cyber threats they are facing [32]. To share CTI, a common structured language is necessary. Hence, the US Department of Homeland Security's Office of Cybersecurity and Communications supported MITRE to develop a structured language for threat sharing (Structured Threat Information eXpression (STIX)) [32]. Furthermore, other languages were developed including the Malware Information Sharing Platform (MISP) and a massive amount of proprietary languages. An overview of different CTI models is provided by Maveroeidis and Bromander [2].

D. Effects of COVID-19 on Cyber-attacks

IT has assumed an essential role in daily activities and penetrated almost every area of human life, including social activities, religious practices, healthcare delivery, education, and business activities. Against this background, the COVID-19 pandemic has changed the way we use IT in our daily life. In particular, personal IT devices have increasingly been used, e.g., for working from home [33]. These personal devices are often poorly protected, representing a new flank of vulnerability for many businesses during COVID-19. Furthermore, attackers leverage on the populations' need for information regarding COVID-19 and how to cope with the pandemic and the guilelessness and curiosity of the users to commit attacks [33], [34]. Attackers thereby adapt their strategies and behavior to exploit special circumstances during the COVID-19 pandemic [33]. "Prolific and opportunistic criminals are taking advantage of the COVID-19 coronavirus pandemic to launch a variety of cyber-attacks" [35]. Thus, the cybercriminal activities of attackers perceive a sharp increase [38]. This is caused by the changes in the environment. Thereby COVID-19 specific changes, e.g., increased use of home office and changed demand for information, lead to changed incentives, new vulnerabilities, and new attack possibilities (e.g., increased vulnerability of healthcare service providers and administrations). As a consequence, attackers are shifting their activities towards new targets where they can reach the highest utility under Covid-19-conditions (e.g., critical infrastructures, healthcare-related services) [36].

III. Quantitative Content Analysis

A. Methodology

The methodology of a systematic content analysis of literature ensures that the work takes up the current state of scientific knowledge and places the research in the wider scientific context. Therefore, the methodology provides a critical examination, interpretation and evaluation of the literature.

The definition of the research question was done in accordance with the paradigm of utility-maximizing attackers. An attacker is thus defined by the utility he receives through attacking, i.e., by the degree of satisfaction of needs. These needs represent the motivational drivers of the attackers. The research question of the content analysis is thus formulated as follows:

RQ1: What motivations of cyber attackers are described in the literature?

To ensure that the literature reviewed captures all relevant references for research question Q1, the selection of keywords considered the integration of synonyms and related words provided by the Merriam-Webster dictionary of the terms attacker and cracker in the sense of hacker and motivation. Test searches led to the integration of the term cyber and its respective synonyms and related words. The identified keywords were combined using Boolean logic to ensure a specific search scope. The keywords were truncated when possible to achieve a complete inclusion of all relevant sources. The literature search was done using the Scopus database. Scopus was chosen for its possibilities of systematic literature search, its high coverage of high-quality scientific journals in the relevant field, its high reputation as a trusted information source, and its highly efficient analytical tools.

The primary search results were further specified by the utilization of inclusion criteria and exclusion criteria. The final search term and the inclusion criteria are presented in Table I.

The literature search was conducted by applying the search terms and inclusion/ exclusion criteria to the title, abstract, and keywords. We extracted 260 journals.

Fig.1,2 and 3 visualize descriptive statistics regarding the selected literature. Fig. 1 plots the number of selected articles ordered by date of publication. It shows that the topic of cyber attackers and their motivation has gained importance, especially in recent years. Furthermore, Fig. 2 presents the count of citations of the selected articles. There are very few highly influential articles; most articles have a comparatively low impact. This corresponds to a classical power-law distribution of citations.

Moreover, if we assume cross-referencing in the field of attacker motivation modelling, it also shows that attacker motivation modeling is not frequently used in other fields of cyber risk research (e.g., quantification). Fig. 3 focuses on the document types. It gives an overview of the works included within this quantitative content analysis.

TABLE I
DETAILS REGARDING THE DATA SEARCH

Search details	Comments
TITLE-ABS-KEY (boost OR encouragement OR goad OR impetus OR impulse OR incentive OR incitation OR incitement OR instigation OR momentum OR provocation OR spur OR stimulant OR stimulus OR yeast OR inducement OR invitation OR antecedent OR cause OR consideration OR grounds OR motive OR occasion OR reason OR catalyst OR catalyzer OR fuel OR spark) AND TITLE-ABS-KEY (attacker OR assailant OR assaulter OR bushwhacking OR mugger OR robber OR molester OR predator OR raper OR rapist OR ravaged OR aggressor OR besieged OR invader OR raider OR counterattacked OR hacker OR cyberpunk OR cracker OR computerise OR spearhead OR geek OR propeller-head OR techie OR technocrat OR technophile OR polisher) AND TITLE-ABS-KEY (cyber OR computer OR computer AND network)	(LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (SUBJAREA , "PSYC") OR LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "ECON") OR LIMIT-TO (SUBJAREA , "DECI")) LIMIT-TO (SUBJAREA , "BUSI") AND (LIMIT-TO (LANGUAGE , "English"))

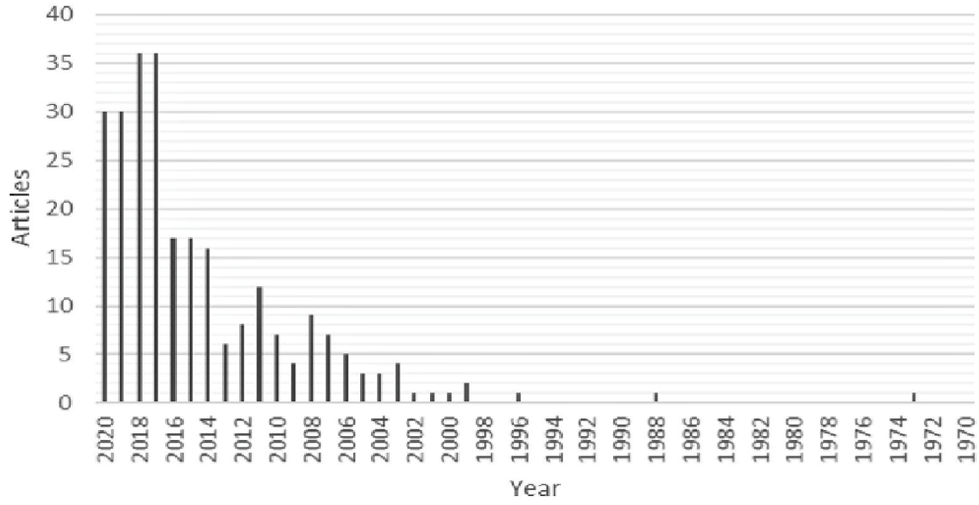

Fig. 1 Date of Publication.

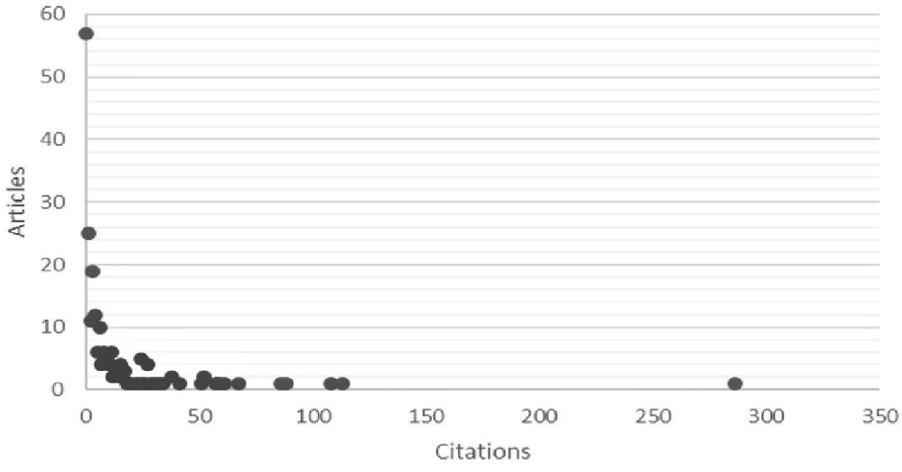

Fig. 2 Citations per Article.

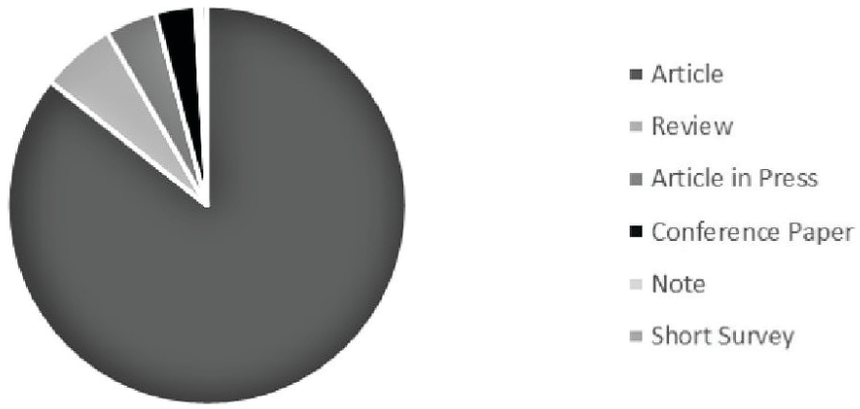

Fig. 3 Document Type of selected Articles.

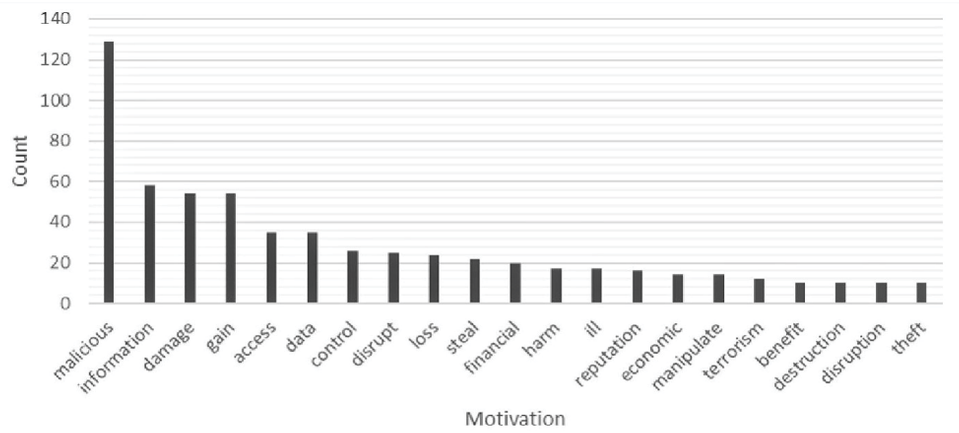

Fig. 4 Motivational terms mentioned within the reviewed literature.

To analyze the text, we designed a coding scheme, which was used to analyze each sentence of an article distinct from the other sentences of an article. We encode each word describing a motivational pattern to the same code as words with the same root. The defined codes were used to identify sentences that include motivational patterns and then to identify the attackers' incentives within these patterns (although both could not be differentiated all the time strictly).

B. Results from the quantitative content analysis and their discussion

We identified the most common terms, including information about attacker motivation. The most common terms (≥10 entries) are presented in Fig. 4. We extracted 377 terms that describe the motivation and the attackers' incentives. Those terms provide a picture of attackers where maliciousness, the will to destroy something or pathology or abnormality, is in the foreground and therefore predominant in the scientific discourse (see Fig. 4). Other motivational patterns, e.g., the aim to acquire knowledge (information, data) and others, could be found. However, the predominance of malicious motives remains the most striking result of the systematic content analysis. This can be put into perspective, as the malicious attacker is often used as a synonym for all hackers. This distinction between normal users on the one side and abnormal or malicious attackers on the other is too rough and undifferentiated, creating an artificial dichotomy. Within this simplified paradigm, the understanding of the attackers and the possibility to model them adequately is greatly weakened. This oversimplification of the attackers' motivations often leads to the fact that the attackers' utility functions are inconsistently or just poorly considered. However, in contrast to natural events,

attackers and defenders should also be regarded as intelligent players. These adapt to the conditions of a system and to the respective actions of the counterpart and, in addition, constantly develop new attack techniques. Such dynamic development of attacking techniques requires up-to-date data records. Hence, plausible attack frequencies and probabilities can be derived from models, which take the attackers' motivation explicitly into account. This is even more important as long as there are no sufficiently large and up-to-date data sets available for statistical analyses. However, the possibility of creating such data sets is further reduced by the fact that the willingness to share information about cyber-attacks is very low since, among other things, critical company information could be released to competitors and other interested parties.

To summarize, the content analysis gives insights into how security researchers around the world perceive hackers. It reveals a highly negative connotation of hackers with respect to their presumed motivation (e.g., bad, malicious, and/or illicit; respectively damage, disrupt, and harm for the incentives). The attackers' backgrounds are significantly different [16], [37], and there is no obvious differentiation from the rest of the population [11], [31]. The high dominance of malicious motives is not supported in scientific studies trying to quantify hacker motives. Therefore, the dominance of maliciousness stands in contrast to the results of these studies, which were presented in more detail within the content analysis (e.g. [11]–[14]).

This raises the question whether our understanding of cyber attackers is biased in the direction of an overrepresentation of specific motivations like maliciousness (although used as a collective term), or whether currently used attacker profiles are not actionable for further research purposes (e.g. usage for threat modelling and quantification of cyber risks). The strong negative connotation of hackers as being illicitly malicious may limit our understanding of attackers. It could thus be a heavy burden for the quantification and mitigation of cyber risks and hence cybersecurity. Jordan and Taylor [38] oppose their perception of hackers as being pathologically bad. This is in line with the theory of rational choice stating that criminality is nothing illicit nor unnatural but a straightforward outcome of rational choice [39]. Furthermore, the differentiated view on attackers may increase our ability to differentiate between abnormal malicious and abnormal benign users (e.g., by analyzing whether an action may result in a reward that is reasonable for the effort). In doing so, an attack hypothesis may be enhanced through attacker modelling and attack hypothesis validation based on attacker motives and observed behavior.

C. Limitations of the content analysis

Through the quantitative evaluation and coding of the assessed documents to infer the frequency of different types of motivation, we needed to extract from the context (at least from the wider context, as we performed a sentence-based analytical procedure). Hence, there might be a bias towards frequently used motivational patterns (e.g., malicious) used as a differentiator between users with legitimate interests and those without them. This can partially explain the dominance of maliciousness as a motive. However, it is also perceived that this use of specific motivational patterns as a differentiator brings considerable insights to the understanding of attackers and their assumed motives.

IV. Natural Experiment

A. Methodology

In the second part of this contribution, we investigate the motivation of attackers and their contribution to cyber risks by exploiting the COVID-19 pandemic as a natural experiment. We use explorative data analysis to highlight the effects of the COVID-19 pandemic.

The COVID-19 pandemic has led to new behavior patterns due to the new conditions (home office, contacts with authorities, etc.). This offers a unique opportunity to consider it as a natural experiment. Specifically, we compare the attack patterns before and during the pandemic in order to draw conclusions about a change in attacker types and the pattern of attacker motivations.

"A natural experiment is a study in which the treatment assignment mechanism is neither designed nor implemented by the researcher, is unknown to the researcher, and is probabilistic by means of an external event or intervention that is outside of the control of the units, which are the subject of the intervention" [42].

Definition I (natural experiment): *Assume that Z is a treatment, Y(0) is the outcome that attains under control, Y(1) vice versa the outcome that attains under treatment Z, and X defines a vector of k covariates determined before the treatment is assigned. Then a natural experiment can be defined as proposed by Titiunik [40] according to the following conditions (I) – (III):*

(i.) Pr (Z\X,Y(0),Y(1)) is neither designed nor implemented by the researcher,

(ii.) is unknown to the researcher,

(iii.) is probabilistic by virtue of an external event or intervention that is outside the experimental units' direct control.

Condition (I) defines that a natural experiment as "a research design where the researcher is neither in charge of the design of the treatment assignment mechanism nor of its implementation" [42]. Condition (II) states that "the treatment assignment mechanism is unknown (…) (hence,) the researcher does not know and has no way of knowing the probabilities associated with each possible treatment allocation" [40]. Last, (III) describes a natural experiment as an "observational study where the mechanism that allocates treatment is known to depend on an external factor" [42].

To determine the feasibility of using the COVID-19 pandemic as a natural experiment and for differentiating the chosen methodology from an observational study, a tentative verification of these characteristics is done. Within doing so, conditions (I) and (II) can be verified. For Condition (III), we assume that the external force of natural processes produced a probabilistic assignment. In the case of COVID-19, this is highly probable; although, the origin of the virus is not unequivocally clarified yet. "This is a heuristic rather than a formal argument, as the units' lack of control of their own assignment is not by itself sufficient to ensure a probabilistic assignment. Rather, the lack of control introduced by the external factor is simply used as the basis for assuming that the assignment was governed, at least partly, by chance" [40].

B. Data

Data sharing practices regarding cyber-attacks are limited. This is because information about cyber threats is perceived as being critical by affected companies. Hence, comprehensive datasets of cybercriminal issues are rare, and often only delayed information is available. In addition, the data sets available to the public and researchers are often highly aggregated. As a result, only highly granular data records are available, which do not allow systematic and comprehensive analyses. For this reason, we have decided to use data from Hackmageddon.com for this research. Hackmageddon.com provides near-time monthly "Cyber Attack Timelines" in which individual attacks are listed. We have used this data to generate a dataset, which includes cyber-attacks from the period 01.10.2019 to 31.05.2020. This period is determined by including a sufficiently long period of time before the outbreak of the COVID-19 pandemic, as well as a sufficiently long period of time after the outbreak of the pandemic. The final dataset includes 244 days and 1,373 attacks. Fig. 5 presents the distribution of cyber-attacks included in our dataset over time.

As a differentiation criterion of the post COVID-19 emergence period (during the pandemic: PCOVID-19) and the pre-COVID-19 period (ACOVID-19), we use the classification as a pandemic (11.03.2020). Within PCOVID-19, we observe an average of 6.2 attacks per day compared to 5.3 attacks per day during ACOVID-19.

We further differentiated those attacks that are related to COVID-19 from those which are not. In doing so, we defined COVID-19 related attacks as those attacks were one of the search terms COVID-19, SARS-CoV2, or corona* could be found in the descriptions. After applying this filter criterion, we came up with 109 COVID-19 related attacks. Those are presented in Fig. 6, according to their date of occurrence. It can be noted that 16 attacks related to COVID-19 were observed before the classification as a pandemic. It can be seen

that there was a spike around the announcement of COVID-19 as a pandemic, which supports the assumption of attacks spiking through the need for information. If COVID-19 related attacks are removed from the record, an average number of attacks per day of 5.1 PCOVID-19 and 5.2 ACOVID-19 can be observed. This shows that, principally, COVID-19 did not change the overall threat landscape but rather offered new opportunities for attacks, mainly due to increased exposure and higher vulnerability of rather inexperienced IT users.

Furthermore, Table II and Fig. 7 present the distribution of COVID-19 related cyber-attacks over specific targets. Table III shows that the most vulnerable target for cyber-attacks related to COVID-19 are individuals (nearly half of all attacks related to COVID-19). This is reasonable for the high number of people searching for new information regarding the virus and how to best cope with it in daily life. Hence, hackers have been seen to increasingly use the fear of individuals as a basis for their actions.

Added to this are the consequences of the lockdown, as a result of which private and business communications are exclusively IT-based. This has led to new types of attacks, which were observable during COVID-19. For example, the increasing use of video conferencing platforms invited a new form of attack called Zoom bombing. We reviewed the attacks presented in our dataset and found that these were observed at least 30 times.

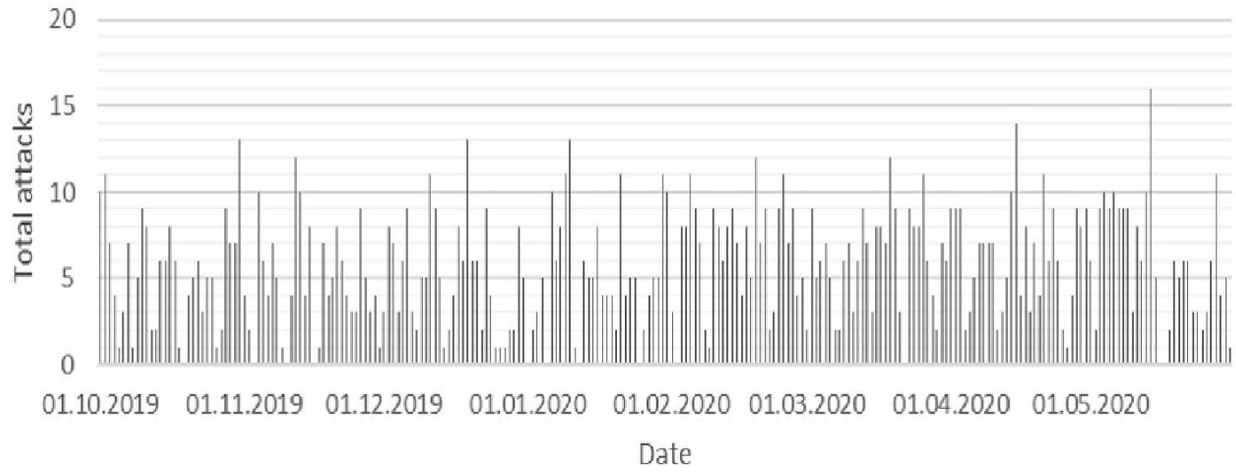

Fig. 5 Total attacks in the dataset between 01.10.2019 and 31.05.2020.

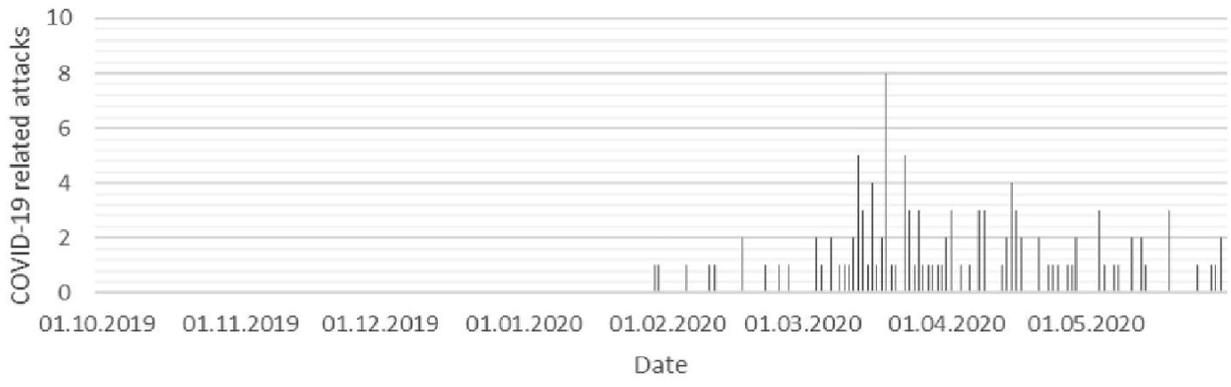

Fig. 6 COVID-19 related attacks within the dataset.

TABLE II
SAMPLE SIZES

Branch/ Target	ACOVID-19	PCOVID-19 [a]
Manufacturing	34	15
Electricity gas steam and air conditioning supply	19	8
Water supply sewerage waste management, and remediation activities	2	0
Wholesale and retail trade	30	12
Transportation and storage	6	7
Accommodation and food service activities	16	6
Information and communication	23	20
Financial and insurance activities	62	33
Real estate activities	2	2
Professional scientific and technical activities	51	35
Administrative and support service activities	3	1
Public administration and defence, compulsory social security	121	46
Education	66	16
Human health and social work activities	84	45
Arts entertainment and recreation	25	18
Other service activities	18	10
Activities of extraterritorial organizations and bodies	9	3
Fintech	17	9
Individual	136	98
Multiple Industries	131	118
Unknown	7	7

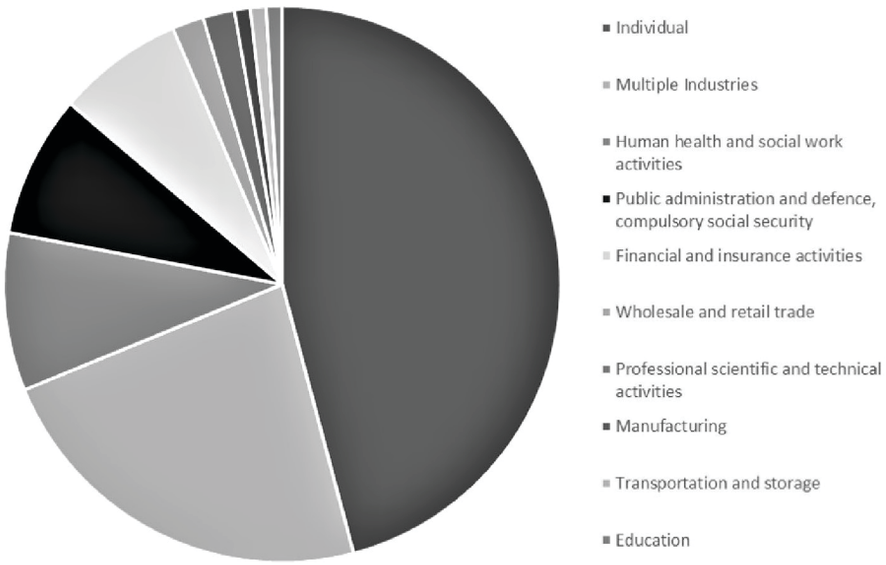

Fig. 7 Shares of COVID-19 related attack targets.

TABLE III
COUNT OF COVID-19 RELATED ATTACKS AND THEIR TARGETS

Branch/ Target	Total	Quote a
Individual	50	0.4587
Multiple Industries	25	0.2294
Human health and social work activities	10	0.0917
Public administration and defence, compulsory social security	9	0.0826
Financial and insurance activities	8	0.0734
Wholesale and retail trade	2	0.0183
Professional scientific and technical activities	2	0.0183
Manufacturing	1	0.0092
Transportation and storage	1	0.0092
Education	1	0.0092

C. Hypotheses

To investigate whether the COVID-19 pandemic has had a statistically significant impact on the cyberattack landscape, we formulated the following hypothesis:

RQ2: H_0^2: The observed mean number of cyber-attacks in PCOVID-19 does not systematically deviate upwards from the observed mean in ACOVID-19 across all branches/targets.

Furthermore, we investigated whether the distribution of cyber-attacks over the different sectors and targets changed. Therefore, we formulated the following hypothesis:

RQ3: H_0^3: The variable "number of attacks during PCOVID-19" follows the same distribution as "number of attacks during ACOVID-19".

RQ4: H_0^4:The variable "attacks by type during PCOVID-19" follows the same distribution as "attacks by type during ACOVID-19".

D. Results from the natural experiment

The RQ2 requires a pairwise comparison of the two states, PCOVID-19 and ACOVID-19. Hence, we performed a T-Test (whole test statistics are listed in the appendix) to investigate whether the observed increase of cyber-attacks in PCOVID-19 is statistically significant. H_0^2 can be rejected (p=0.0217).

We performed a Chi-Square Goodness of Fit test to test RQ3 and RQ4, investigating whether the categorical variables follow the hypothesized distributions (distributions from ACOVID-19). Hypothesis H_0^3can be rejected (p=5,6673·10^{-7}). Hence, we conclude that COVID-19 has had an impact on the distribution of cyber-attacks over the considered industries/targets. These changes are due to the different effects of the COVID-19 pandemic on the attractiveness of these targets for potential attackers.

Furthermore, we investigated whether there are effects of COVID-19 on the different sectors/targets. In doing so, we focused on those industries where the total number of attacks was larger or equal to thirty (whole test statistics are listed in the appendix). This threshold enables the use of normal distributions assumptions for analytical purposes. We observed significant effects of the COVID-19 pandemic on public administration and defence, compulsory social security (decrease in the number of cyber-attacks; p=0.0459, one-tailed), individuals (increase in the number of cyber-attacks; p=0.0183, one-tailed), education (decrease in the number of cyber-attacks; p=0.0027, one-tailed), and multiple industries (increase in the number of cyber-attacks; p=0.0001, one-tailed). Furthermore, we observe a tendency for increased cyber-attacks on the information and communication branch (p=0.0617, one-tailed). In contrast, no significant effects of the changed situation during COVID-19 could be found for the industries/targets of manufacturing, wholesale and retail trade, financial and insurance activities, professional scientific and technical activities, human health, and social work activities, and arts, entertainment, and recreation.

Hypothesis H_0^4 has to be rejected (p=0.0265). Hence, we conclude that the pandemic has had an impact on the employed attacks (attack techniques).

Regarding the attacks, we observe a statistically significant increase of unknown attack techniques during the pandemic (p=0.0204, one-tailed) and a statistically significant decrease of malicious script injection (p=0.0154, one-tailed).

Finally, a further noteworthy finding sheds light on attackers' motivations: Attacks on the human health and social work activities sector, which is one of the most vulnerable sectors during the COVID-19 pandemic, have not increased significantly. Attackers, therefore, do not seem to exploit every possible vulnerability to the maximum at any price. One possible explanation for this could be ethical concerns and norms within the hacker scene. The results of the natural experiment thus suggest the need for a differentiated approach in attack modelling, using disjunctive attributes. Such an approach could make a significant and valuable contribution to the understanding of the attackers and the cybercriminal landscape.

V. Conclusion and Discussion

We reviewed and analyzed research articles in the area of cyber security to investigate which attacker motivations are considered in scientific literature. Furthermore, we investigated the importance of these motivations within the literature based on the number of occurrences in different scientific papers. In doing so, a strong focus of articles on malicious motives was revealed. It was assumed that this strong focus on maliciousness could hamper our understanding of attackers. Highlighting the oversimplification in attacker motivations, we aimed to build awareness about this problem in cyber security research. Therefore, this work should improve our understanding of attackers to overcome the narrow, undifferentiated view that cybercriminals are inherently bad or abnormal.

Furthermore, we highlighted that the oversimplified attacker motivation modelling leads to misunderstanding cyber-attacks. Therefore, we employed a natural experiment to understand whether what we would assume from a malicious attacker modelling comes close to reality. We observed within this natural experiment various targets that should be more at risk from what we would expect (e.g. human health and social services, public administration and defence, compulsory social security). However, these targets show a different behavior in the natural experiment. This is perceived as a further hint (besides the overrepresentation of malicious motives in scientific literature) that our approach to model attacker motivation can significantly improve cyber risk assessment. Including e.g. moral considerations of hacking a target could explain the behavior revealed within our natural experiment (e.g. no statistically significant increases of attacks on the human health and social work branch or statistically significant decreases of attacks on public administration and defence, compulsory social security). Furthermore, the currently used attacker classifications may be a burden for cyber risk quantification using threat centric approaches, as these are not actionable for computational (game theoretic) models. Through including the variety of motivations that may influence an attacker, it may be possible to quantify risks more accurately and make cyber risk management more efficient. It is hence of utmost importance to model attackers correctly to increase cybersecurity capabilities, quantify cyber risks and engineer secure systems.

References

[1] A. Elitzur, R. Puzis, and P. Zilberman, "Attack Hypothesis Generation," in *2019 European Intell. Secur. Inf. Conf. (EISIC)*, Oulu, Finland, pp. 40-47, doi: 10.1109/EISIC49498.2019.9108886.

[2] V. Mavroeidis and S. Bromander, "Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence," in *2017 European Intell. Secur. Inf. Conf. (EISIC)*, Athens, Greece, pp. 91-98, doi: 10.1109/EISIC.2017.20.

[3] R. Stillions, "The DML Model", 2014, [Online]. Available: http://ryanstillions.blogspot.com/2014/0 4/the-dml-model 21.html

[4] J. Lickiewicz, "Cyber Crime psychology-proposal of an offender psychological profile," *Probl. Forensic Sci.*, vol. 87 (LXXXVII), pp. 239-252, 2011.

[5] L. J. Stalans and C. M. Donner, "Explaining Why Cybercrime Occurs: Criminological and Psychological Theories," in *Cyber Criminology*, H. Jahankhani, Ed., in Advanced Sciences and Technologies for Security Applications, Cham, Switzerland: Springer, 2018, pp. 25-45.

[6] S. M. Mandelcorn, "An explanatory model of motivation for cyber-attacks drawn from criminological theories" Ph.D. dissertation, Mech. Eng., Maryland Univ., MD, USA, 2013.

[7] L. Allodi, F. Massacci and J. M. Williams, "The workavers cyber attacker model: Theory and evidence from two million attack signatures," June 2017. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2862299

[8] B. Thomson, "The COVID-19 pandemic: A global natural experiment," *Circulation*, vol. 142, no. 1, Apr. 23, 2020, doi: 10.1161/CIRCULATIONAHA.120.047538.

[9] X. Li, "A Review of Motivations of Illegal Cyber Activities," *Kriminologija & socijalna integracija*, vol. 25, no. 1, pp. 110-126, 2017, doi: 10.31299/ksi.25.1.4.

[10] A. Warikoo, "Proposed Methodology for Cyber Criminal Profiling," *Inf. Secur. J.*, vol. 23, no. 6, pp. 172-178, Oct. 22, 2014, doi: 10.1080/19393555.2014.931491.

[11] C. S. Fötinger and W. Ziegler, "Understanding a Hacker's Mind–a psychological insight into the hijacking of identities," Danube-University, Krems, Austria, 2004.

[12] H. Woo, Y. Kim and J. Dominick, "Hackers: Militants or merry pranksters? A content analysis of defaced web pages," *Media Psychol.*, vol. 6, no. 1, pp. 63-82, Nov. 17, 2009, doi: 10.1207/s1532785xmep0601_3.

[13] Thycotic Software Ltd., "Thycotic Black Hat 2014 Hacker Survey Executive Report," 2014.

[14] R. Madarie, "Hackers' Motivations: Testing Schwartz's Theory of Motivational Types of Values in a Sample of Hackers," *Int. J. Cyber Criminol.*, vol. 11, no. 1, pp. 78-97, 2017, doi: 10.5281/zenodo.495773.

[15] R. Van Holsteijn, "The Motivation of Attackers in Attack Tree Analysis," M.S. thesis, Inf. Commun. Technol., TU Delft Univ., Delf, Netherlands, 2015.

[16] M. R. Moosavi, J. M. Fazaeli, M. H. Sadreddini, and J. M. Zolghadri, "Entropy Based Fuzzy Rule Weighting for Hierarchical Intrusion Detection," *Iran. J. Fuzzy Syst.*, vol. 11, no. 3, pp. 77-94, 2014, doi: 10.22111/IJFS.2014.1571.

[17] V. Brandstätter, J. Schüler, R. M. Puca, and L. Lozo, *Motivation und Emotion*. Berline, Germany: Springer, 2017.

[18] D. C. McClelland, J. W. Atkinson, R. A. Clark and E. L. Lowell, *The Achievement Motive*. New York, NY, USA: Appleton-Century-Crofts, Inc, 1953.

[19] H. A. Murray, Exploration in personality. New York, NY, USA: Oxford Univ. Press, 1938.

[20] J.W. Atkinson, "Motivational determinants of risk-taking behavior," *Psychol. Rev.*, vol. 64, no. 6 Pt. 1, pp. 359–372, 1957, doi: 10.1037/h0043445.

[21] O. C. Schultheiss, "Implicit motives," in *Handbook of Personality: Theory and Research*, in O. P. John, and R. W. Robins, New York, NY, USA: The Guilford Press, 2008, pp. 385-410.

[22] R. Baumeister, and M. R. Leary, "The need to belong: Desire for interpersonal attachments as a fundamental human motivation," *Psychol. Bull.*, vol. 117, no. 3, pp. 497–529, 1995, doi: 10.1037/0033-2909.117.3.497.

[23] B. Landreth, *Out of the Inner Circle: A Hacker's Guide to Computer Security*, Bellevue, WA, USA: Microsoft Press, 1985.

[24] R. Hollinger, "Computer hackers follow a guttman-like progression", in Social Sciences Review, 72, 1988, pp. 199-200.

[25] N. Chantler, *Profile of a Computer Hacker*, FL, USA: InterPact Press, 1997.

[26] R. Barber, "Hackers profiled—who are they and what are their motivations?," *Comput. Fraud Secur.*, vol. 2001, no. 2, pp. 14-17, Feb. 2001, doi: 10.1016/S1361-3723(01)02017-6.

[27] A. D. Smith and W.T. Rupp, "Issues in cybersecurity; understanding the potential risks associated with hackers/crackers," *Inf. Manag. Comput. Secur.*, vol. 10, no. 4, pp. 178-183, 2002, doi: 10.1108/09685220210436976.

[28] T. Casey, "Threat Agent Library Helps Identify Information Security Risks," Intel Corporation, USA, 2007.

[29] T. Casey, "Understanding Cyberthreat Motivations to Improve Defense," Intel Corporation, USA, 2015.

[30] M. K. Rogers, "A two-dimensional circumplex approach to the development of a hacker taxonomy," *Digit. Investig.*, vol. 3, no. 2, pp. 97-102, June 2006, doi: 10.1016/j.diin.2006.03.001.

[31] J. Bässmann, "Täter im Bereich Cybercrime-Eine Literaturanalyse," March 2, 2017. [Online]. Available: https://www.bka.de/SharedDocs/Downloads/DE/Publikationen/Publikationsreihen/Forschungsergebnisse/2015TaeterImBereichCybercrime.html

[32] S. Barnum, "Standardizing cyber threat intelligence information with the Structured Threat Information eXpression (STIX)," Mitre Corporation, McLean, VA, USA, Jan. 2012. [Online]. Available: https://www.mitre.org/sites/default/files/publications/stix.pdf

[33] J. Wiggen, "The impact of COVID-19 on cyber crime and state-sponsored cyber activities," Konrad-Adenauer-Stiftung, Berlin, Germany, No 391/June 2020, 2020.

[34] K. Okereafor and O. Adebola, "Tackling the Cybersecurity impacts of the coronavirus outbreak as a challenge to internet safety," *Int. J. IT Eng.*, vol. 8, no. 2, pp. 1-14, Feb. 16, 2020.

[35] Interpol, "Global Landscape on Covid-19 cyberthreat," Apr. 2020. [Online]. Available: https://www.interpol.int/en/Crimes/Cybercrime/COVID-19-cyberthreats

[36] Interpol, "Cybercrime: Covid-19 impact," INTERPOL General Secretariat, Lyon, France, 2020. [Online]. Available: https://www.interpol.int/content/download/15526/file/COVID-19%20Cybercrime%20Analysis%20Report-%20August%202020.pdf

[37] R. Chiesa, S. Ducci, and S. Ciappi, *Profiling hackers*, Boca Raton, FL, USA: CRC Press, 2008.

[38] T. Jordan and P. Taylor, "A sociology of hackers," *Sociol. Rev.*, vol. 46, no. 4, pp. 757-780, Nov. 1, 1998, doi: 10.1111/1467-954X.00139.

[39] M. Bock, *Kriminologie: für Studium und Praxis*, Munich, Germany: Vahlen Franz GmbH, 2017.

[40] R. Titiunik, "Natural experiments," *arXiv:2002.00202*, Feb. 2020, [Online]. Available: https://arxiv.org/abs/2002.00202