

University of Groningen

A Taxonomy for Large-Scale Cyber Security Attacks

Mohsen, Fadi; Zwart, Cornelis; Karastoyanova, Dimka; Gaydadjiev, Georgi

Published in:
EAI Endorsed Transactions on Cloud Systems

DOI:
[10.4108/eai.2-3-2022.173548](https://doi.org/10.4108/eai.2-3-2022.173548)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Mohsen, F., Zwart, C., Karastoyanova, D., & Gaydadjiev, G. (2022). A Taxonomy for Large-Scale Cyber Security Attacks. *EAI Endorsed Transactions on Cloud Systems*, 7, Article e5. <https://doi.org/10.4108/eai.2-3-2022.173548>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

A Taxonomy for Large-Scale Cyber Security Attacks

Fadi Mohsen*, Cornelis Zwart, Dimka Karastoyanova and Georgi Gaydadjiev

Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence
University of Groningen, The Netherlands

Abstract

In an effort to examine the spread of large-scale cyber attacks, researchers have created various taxonomies. These taxonomies are purposefully built to facilitate the understanding and the comparison of these attacks, and hence counter their spread. Yet, existing taxonomies focus mainly on the technical aspects of the attacks, with little or no information about how to defend against them. As such, the aim of this work is to extend existing taxonomies by incorporating new features pertaining the defense strategy, scale, and others. We will compare the proposed taxonomy with existing state of the art taxonomies. We also present the analysis of 174 large cyber security attacks based on our taxonomy. Finally, we present a web tool that we developed to allow researchers to explore exiting data sets of attacks and contribute new ones. We are convinced that our work will allow researchers gain deeper insights into emerging attacks by facilitating their categorization, sharing and analysis, which results in boosting the defense efforts against cyber attack.

Received on 21 February 2022; accepted on 01 March 2022; published on 02 March 2022

Keywords: Taxonomy, Cyber, Security, Attacks, Large-Scale, Defense, Portal

Copyright © 2022 Fadi Mohsen *et al.*, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/eai.2-3-2022.173548

1. Introduction

With the emergence of more and more diverse types of devices into modern networks, e.g., smart phones, the attack surface has been drastically extended. As a result, the number and volume of cyber attacks have increased. According to ENISA, in 2019, 66% of healthcare organizations experienced a ransomware attack [1]. In 2020, the average data breach cost was \$3.86 million, according to IBM [2]. In 2021, this increased to \$4.24m, with individual (mega) breaches costing up to \$401m [3]. The most common causes behind these breaches in 2021 were compromised credentials, phishing, cloud mis-configuration, and vulnerability in third-party software.

In order to protect against such attacks, cyber security experts and researchers need to be up-to-date on what types of attacks take place [4]. They are also in need of standards and methodologies to systematically analyze occurring attacks in order to learn from them. For example, there is a need for a reference cyber security taxonomy. Such taxonomy can be seen as an essential tool to facilitate the classification and

understanding of large-scale cyber attacks and increase the defense capabilities of different organizations and device vendors. Taxonomies give security researchers and practitioners a standardized way of analyzing large-scale cyber attacks and learning from them. A number of cyber security taxonomies have been proposed over the years. A closer look at these taxonomies reveals various shortcomings and indicate the need of an improved one. For instance, some of the reported taxonomies, McCumber Cube [5], focus only on the high-level attributes of the attack and ignores the low-level attributes. Other taxonomies do exactly the opposite, such as AVOIDIT [6] and ATT&CK [7]. Additionally, almost all existing taxonomies overlook the defense part. In our view, it is imperative for taxonomies to include, in some aspect, how cyber security experts can defend against the different cyber attacks. Finally, existing taxonomies do not usually include example data sets that can be used as a reference by researchers and practitioners.

Therefore, in this paper, we are proposing a new taxonomy that overcome all aforementioned limitations. Our proposed taxonomy builds upon a widely accepted taxonomy such that it also covers the characteristics of an attack, as well as the impact

*Corresponding author. Email: f.f.m.mohsen@rug.nl

and how it could have been avoided. We then collect a data set of cyber attacks that are classified using our taxonomy. The data set is made publicly available through a dedicated web application. This allows for low-overhead exploring, analyzing, and exporting current attacks. Additionally, the interface enables security researchers and practitioners to submit their own data sets of cyber attacks.

Through this interface, users can get good insights into the most frequent cyber attacks and the best ways to defend against them. The data set currently includes 174 cyber attacks that took place between 2010 and 2020. Finally, we compare between our proposed taxonomy against all existing taxonomies. The comparison is based on whether they provide a defense mechanism or a strategy to compact the attacks, and the availability of data sets and exploration tools. We believe that our work will facilitate better prevention of the steadily growing number of large-scale cyber attacks by means of categorization support, sharing and analysis.

The remainder of this paper will be structured as follows. We start in Section 2, by looking at and comparing existing taxonomies. We further look at the criteria that a good taxonomy should meet. In Section 3 we propose our taxonomy and present our data set. In Section 4 we present the web application that we created to interact with our collection of cyber security attacks. Finally, we conclude the paper in Section 5.

2. Existing Taxonomies

Cyber-attack taxonomies cover several aspects of a cyber attack. Each of these aspects is called a *feature* and might be further subdivided into two or more features. A taxonomy is meant to facilitate the collection, analysis and dissemination of information pertaining to the cyber attacks. Another benefit to the standardized representation of attacks is that they can easily be compared. This gives practitioners and researchers the opportunity to discover various kinds of relationships between involved attacks. In the following section we will give several concrete examples on cyber-attack taxonomies, however, we will first introduce our comparison criteria used to assess these taxonomies.

2.1. Requirements for Usable Taxonomy

Before we dive into any comparison attempt between existing taxonomies and our proposed taxonomy, we first need to define a comparison criteria. The criteria will pinpoint the strengths and limitations of security-attack taxonomies. There had been a number of research efforts towards this goal, for example, the work of J.D. Howard et al [8] in 1998 and the work of D. Lough et al [9] in 2001. Both of these works were used in surveys of existing cyber-security taxonomies such as the one

from R. Derbyshire et al [10]. Our analysis of these efforts resulted in identifying six major requirements for a good taxonomy:

- Accepted: it should build on previous work;
- Mutually exclusive: overlapping between different classes should be impossible;
- Comprehensible: It should be clear what categories mean, for experts and those less familiar;
- Complete: the taxonomy should be able to classify all known attacks;
- Repeatable: repeated classification of the same attack should yield the same result.
- Unambiguous, the meaning of each category should be clear;

In designing our taxonomy, we made sure that the above requirements are met. Additionally, we propose new requirements for an effective taxonomy. A taxonomy would be effective if it helps security experts and researchers gain insights about future defense strategies. Moreover, a taxonomy would have more chance of being widely accepted and used if it comes with a data set and a web application that enables viewing and sharing of cyber attacks.

2.2. Related Taxonomies

Numerous taxonomies have been proposed over the years. In this section, we will give an overview about some of them, the ones that are related to our work. Chuck Easttom et al [5] formulated a taxonomy based on McCumber's cube, initially proposed in 1991. The modified McCumber cube classifies an attack based on whether it compromises the confidentiality, integrity or availability of the data. The taxonomy distinguishes between data in storage, data that is being transferred, for example over the Internet, and data that is being processed. It also classifies an attack based on what impacted the confidentiality, integrity and availability, the technology used, the policy and practices of the organization or the education and awareness of the personnel. Though this taxonomy can be applied to any cyber attack, it is quite broad. This taxonomy is limited in capturing some of the important aspects of an attack. For instance, no information is provided regarding the attacking strategy, e.g., attack model, nor how the attack could have been prevented. Dorottya Papp et al [11] presented a taxonomy for classifying cyber attacks specifically targeted at embedded systems. Their taxonomy is mainly focused on the factors that contribute to a successful attack and how to use these factors to make embedded systems more secure. Though, they did not use any actual cyber attacks. Instead they used a subset of the vulnerabilities listed in the Common Vulnerabilities and Exposures data set [12]. Dennis Kiwia et al [13] designed a taxonomy

for classifying trojans used against banking systems. Their taxonomy goes into a lot of detail with regard to how banking credentials can be stolen from people. Their classification of the attacks is based on the attack evasion technique, attack vector, and data exfiltration tactic. This taxonomy is attack and domain specific, thus, it is not meant to classify other kinds of cyber attacks. K. Harrison [14] looked at cyber attacks from a different perspective and focused on the social aspects of cyber attacks and their effects on the community, such as who carried out the attack, why was this community attacked and what the impact is.

Bonnie Zhu et al [15] describes a taxonomy for cyber attacks on SCADA ¹ systems, which are systems that control physical devices or gather data about physical systems. The taxonomy recognizes the uniqueness of SCADA systems in comparison to conventional IT systems. The attacks against SCADA systems were classified based on the common attack vectors. Additionally, they were also looked at from the point view of a control engineer.

Chris Simmons et all [6] designed AVOIDIT in 2009. They base their taxonomy on the requirements for a complete taxonomy, as described in [9][8]. Their taxonomy, just like the McCumber Cube, is able to classify any cyber attack. However, they include more specific information about how the attacker got into the system and what they did once she is in. They also included a feature called “defense” that specifies how to prevent that type of attack, or how to minimize the damage once the attack has taken place.

Chanchala Joshi et al [16] build on the AVOIDIT [6] taxonomy in 2014. They partially have similar features as AVOIDIT, such as the type of attack and what it does once the system is compromised. However, they expand upon it by asking who did the attack and how it could be prevented.

The MITRE organization [7] developed the ATT&CK taxonomy. Though not described in a paper, its taxonomy is worth looking at. They classify attacks based on 206 techniques, many of which have sub-techniques. This allows for classification of many aspects of an attack. This is shown in their data set of classified attacks. When looking at this data set, it becomes clear that the taxonomy allows for many cyber attacks to be classified. However, many of the attacks in the data set are only partially classified, for

example, only 2 techniques might be classified. This, together with the format in which it is represented, makes making conclusions based on the data set hard.

Summary In Table 1, we summarize the related taxonomies that we discussed above. For each taxonomy, we specify the domain, the sector, whether a data set was published with it, whether there is a GUI interface, and the number of features. As can be seen, only one of the reviewed taxonomies comes with a data set. The AT&CK data set from MIRTRE [7] is maintained by experts that send classified attacks to the MITRE organization. These classified cyber attacks, however, are often partially classified, with only a few features noted. This, together with the JSON format used, makes it harder to compare attacks. Additionally, there are only one taxonomy that has features pertaining preventative defense solutions, which is the AVOIDIT [6]. In the last row of Table 1 we consider the taxonomy we introduce in this work. As we will see in the next section, it includes a data set, a web application, and a defense strategy feature.

3. The Proposed Taxonomy

Our taxonomy is based on the AVOIDIT taxonomy. We added relevant features to allow for more analysis and comparison of contemporary and future cyber attacks. In Table 2, we show the list of features in our taxonomy and their origins. Later in this section we will discuss each one of these features in more details.

3.1. AVOIDIT

AVOIDIT taxonomy is fundamentally designed to classify any cyber attack. It classifies the cyber attacks based on five features. Namely, the attack vector, operational impact, defense, informational impact and target. Each one of these features is further subdivided into subcategories to enable more precise classification of the attacks.

Attack vector. It is the path used by the attacker to gain access to the device. Below is a list of attacking vectors as defined by AVOIDIT:

- Misconfiguration: the attacker uses a flaw in a configuration to gain access to the device;
- Kernel flaw: the attacker uses a flaw in the kernel of the operating system;
- Buffer overflow: a piece of code writes data outside of its allotted memory;
- Insufficient input validation: the application does not sufficiently check its input. An attacker can use this to input arbitrary code, like an SQL injection;

¹MEANING of Abbreviation?

lightgray Tax.	Year	Domain	Sector	Data set	Portal	Defense.	Size
[5]	2019	General	Academic	No	No	No	18
[11]	2015	Embedded systems	Academic	No	No	No	5
[13]	2017	Banking Trojans	Academic	No	No	No	8
[14]	2011	General	Academic	No	No	No	9
[15]	2011	SCADA systems	Academic	No	No	No	4
[6]	2009	General	Academic	No	No	Yes	5
[16]	2014	General	Academic	No	No	No	5
[7]	2021	General	Industry	Yes	No	No	206
<i>Our</i>	<i>2021</i>	<i>General</i>	<i>Academic</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>15</i>

Table 1. Overview of existing taxonomies in comparison to our taxonomy.

#	Feature	Taxonomy
1	Attack vector	[6]
2	Operational impact	[6]
3	defense	[6]
4	Informational impact	[6]
5	Target	[6]
6	Date of attack	New
7	Date of discovery	New
8	Date of announcement	New
9	Breached before	New
10	Sector	New
11	defense strategy	New
12	Inside job	New
13	Countries affected	New
14	Money loss	New
15	People affected	New

Table 2. The list of all features that we included in our taxonomy

- Symbolic links: a file points to another file;
- File descriptor: the file does not have a name but relies on numbers from the system to track files. The attacker can use this to gain elevated privileges;
- Race condition: The program accesses a process, but between references the object changes;
- Incorrect file/directory permission: A file or directory has an incorrect permission associated with it;
- Social engineering: The attacker uses interaction with a person to gain access to the system.

Operational Impact. It describes what the attacker does once inside the system. It can be any combination of the followings:

- Misuse of resources: the attacker uses system resources that it should not be able to use;
- User compromise: the attacker gains user privileges on a system;

- Root compromise: the attacker gains root or administrator privileges on the system;
- Web compromise: a website or web-application is used to spread the attack;
- Installed malware: the attacker installs malware, allowing him to take full control of the system;
- Denial of service: the attacker prevents the attacked from accessing certain resources;
- Unknown: if either the operational impact is not known or not one of the earlier options this is used.

Defense. It describes whether there are steps that can prevent the attack from happening or lessen the damages caused when it is happening. The steps can be best classified into:

- Mitigation: there are steps that can lessen damages by the attack;
- Remediation: there are steps that can be taken to prevent an attack, or remove the attack when it is happening;
- Both: for when there are both Mitigation and Remediation steps available.

Informational Impact. It describes what the attacker does to the data that is stored in the system. Which can be one or more of the followings:

- Distort: the data is altered;
- Disrupt: the inability to access data or services on the system, like in a Denial of Service attack;
- Destruct: sensitive data is removed from the system;
- Disclosure: data is leaked to others, this would be a data-leak;
- Discovery: the attacker gains information about the attacked system or network. This information could be used to launch further attacks.

Target. It specifies that part of the system that is being attacked. Which can be one or more of the followings:

- Operating system: the attack is made to affect a specific operating system;
- Network: the attack is made to use a vulnerability in a particular network;
- Local: a user's local computer is attacked;
- User: a specific user of a system is targeted;
- Application: the attack uses a vulnerability in a specific application.

3.2. The Proposed Feature Extensions

In addition to the AVOIDIT features, we propose to add the following features:

- Date of attack: at what date was the system breached;
- Date of discovery: at what date did the organization discover that their system was breached;
- Date of announcement: when did the organization reveal to the world there had been an attack;
- Breached before: whether this was the first or a later breach: whether the attack was the result of information or access gained in a previous attack;
- Sector: whether the organization is a government or civilian;
- Defense strategy: the defense strategy that could have been deployed to prevent the attack: options will be listed below;
- Insider job: whether it was an inside job, someone in the organization willingly helped the attacker or acted on their own;
- Countries affected: list of affected countries.
- Money loss: how much money was lost as a result of the attack;
- People affected: how many people were affected by the attack.

Defense Strategy. This feature describes the defense strategy that could have been implemented to prevent from a specific attack. Each strategy entails implementing various kinds of security techniques. Therefore, this feature is meant to cover more techniques than the *defense* feature of AVOIDIT [6] taxonomy. We therefore believe this to be a valuable addition to the taxonomy. For this feature, we took inspiration from T. Shimeall and J. Spring's book [17]. The authors described a layered-defense approach to counter cyber security attacks. Their approach is comprised of four layers as explained below:

- Deception: prevent the attacker from knowing what part of the system to attack. If it is not known in what part of the network or system sensitive data or services are located they cannot attack it;
- Frustration: prevent the attacker from gaining access to the system. For example, routers and

firewalls prevent access to certain resources for unauthorized people;

- Resistance: prevent the attacker from gaining further access to the network after initial access to a system;
- Recognition and recovery: this strategy focuses on recognizing when an attack is happening and then stopping it. As this is reactive, it can only start once the attack is happening, it should not be the only defense used.

This is a crucial feature that provides a strategy that can be employed to prevent similar attacks from happening in the future. The security practitioners may then choose one or more of the measures that fall under that particular strategy.

3.3. Data set

In this work, we have collected a data set of recent large-scale cyber attacks. All of these attacks are classified based on our taxonomy. These cyber attacks will be used as the initial seed for the interface we have implemented. In collecting these attacks, we have employed two criteria; an attack must be large-scale and it must be recent. The attacks must have been taken place between 2010 and 2020. The large-scale criteria is determined by either the money loss resulting from the attack or the number of people affected by it. As such, the data set includes cyber attacks on high profile government targets. Currently the data set consists of 174 large scale cyber attacks. These attacks have a total damage of more than 3.3 billion dollars and more than a billion, non-unique, people affected.

4. Implementation

As described earlier, most of the existing taxonomies do not have large public data sets available. For example, the ATT&CK taxonomy, even though it has a large data set; yet, it has the problem that many of these attacks are only partially classified. Moreover, the ATT&CK taxonomy does not provide any option for visualizing and querying the data set. As such, we decided to develop a web application for security researchers to view, query and visualize our collection of cyber attacks. In the next section, we will give an overview of this web application.

4.1. Technology stack

The web application consists of two primary parts; the database part and the visualization part. The cyber attacks are stored in a MySQL database managed by phpMyAdmin 4.9.5. The visualization part is implemented using PHP 8.0. The website is currently hosted on *000webhostapp.com/*.

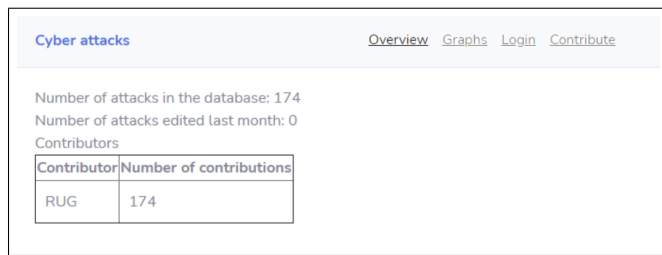


Figure 1. The summary of the exiting large-scale data set is displayed at the top-left corner of the overview page.

4.2. Interface features

Our website enables dynamic view of the attacks. Security researchers can choose to view, query and generate graphs for all attacks or a subset of these attacks. The researchers can then download the results of their queries as CSV files or PNG images. Most importantly, the website allows researchers to upload their own collections of attacks based on the template that we provide them with. The template includes all the features of our taxonomy.

Overview of cyber attacks. The main page of the website shows a summary of the existing data set at the top-right corner, Figure 1. To the right of the summary, there is a menu for navigating to other pages such as the *Graphs* page, *Login* page and *Contribute* page. Below the summary, there is the search form, which allows security researchers to locate certain attacks that meet certain conditions, Figure 2. The list of all attacks or part of them, depending on the selections made through the search form, will be displayed at the bottom of the main page, Figure 3.

The download button allows researchers to download all the attacks or only the results of the search process as .CSV file. The choice for .CSV file is made because it is an open standard, anyone can use it and no special software is needed to access the file. An Excel, .xlsx file, for example, does not have this advantage.

Add or edit an attack. Since the collection and the classification of cyber attacks is a time-consuming process. Our website allows researchers to contribute their own collection of cyber attacks. For that, we created two options that will appear in the top-right menu once the user is logged. In case there is a few numbers of attacks that needs to be added, a contributor can simply use the *Add Attack* page. The page has a simple HTML form that contains all the features of the taxonomy. The categorical features of the taxonomy are represented by drop-down menus. This is intuitive and prevents the possibility of incorrect input. In addition to the taxonomy specified categories, each dropdown has an "Unknown" option in case a certain feature is missing. The second option is more appropriate

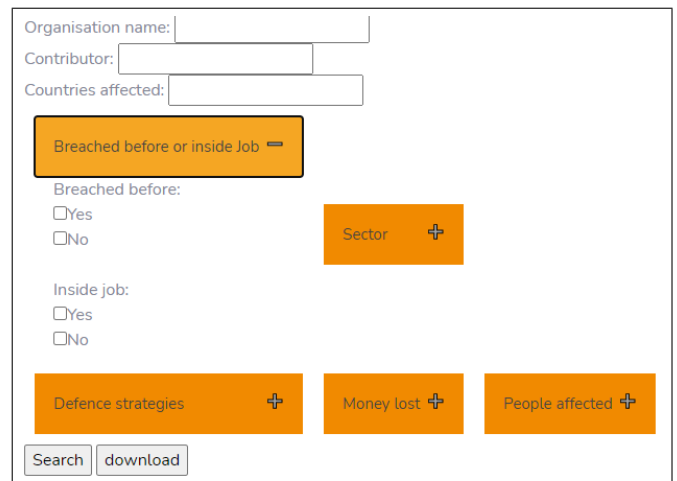


Figure 2. The search form has numerous options for filtering the attacks.

when the number of attacks are large. In this case, a contributor can use the *Import* page. On this page, an explanation for each feature is given as well as a CSV template. The contributor can use this template to fill out her attacks then upload it. It is worth noting that a contributor account can only be created by the admin. A contributor who wishes to have an account on the website needs to submit a request from an official email address, e.g., university email. Aside from the contribution part of the website, the other parts of the website do not require an account.

Due to the possibility of mistakes in entering the attacks or missing information at the time of entry, our website allows updating existing records. An update of a record creates a new log entry with a specific timestamp. This allows researchers to track the changes and reference the correct version of the data set in their works. Researchers can also filter exiting attacks based on this timestamp, giving them the opportunity to either download the most recent copy of the data set or older versions.

Graphs. Our website allows researchers to use the *Graphs* page to generate and download numerous graphs. For example, a researcher might be interested to look at the average money loss per sector as shown in Figure 5, average money loss per attack vector as shown in Figure 6, or number of attacks per attack vector as shown in Figure 7. In Figure 4, the first drop-down menu determines what should be on the y-axis. This could be, for example, the number of attacks on a target or the average money lost because of an attack. The second drop-down menu determines the x-axis. The feature that is selected here will determine over which groups the query from the first drop-down will be done. For instance, it will show the average money lost because of an attack for each of the possible targets.

Organisation name	Contributor	Date of attack	Date of discovery	Date of announcement	Sector	Breached before
Reddit	RUG	2018-06-14	2018-06-19	2018-08-01	Civilian	0
TicketFly	RUG	2018-05-30	2018-05-31	2018-05-31	Civilian	0
Government of The Phillipines	RUG	2018-04-01		2018-04-24	Government	1
MyFitnessPal	RUG	2018-02-01	2018-03-25	2018-03-30	Civilian	0

Figure 3. A snippet of the attacks displayed at bottom of the main page. Note that not all columns are shown here.

per

Figure 4. The graph generator menu.

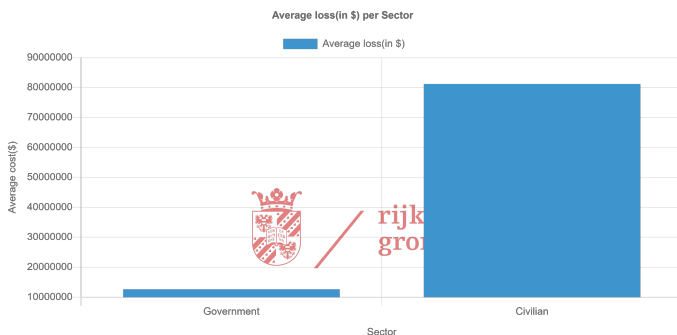


Figure 5. Average money loss per sector.

For example, selecting "Sector" yields the groups "Government" and "Civilian". The first drop-down determines what information should be shown for each of these groups, e.g., the average money lost per group. The result of this query is shown in Figure 5.

Interpreting Graphs. With the ability to make graphs we shall explore some interesting results from our data set. In order to determine the attack vector that had caused the most money loss, we generate Figure 6. As can be seen, excluding attacks where the attack vector is unknown, social engineering and unsecured hardware are the most financially damaging attacking vectors.

A single social engineering attack cost, on average, around 90 million dollars. Unsecured hardware attacks cost about 119 million dollars. To compare, the other attack vectors result in damages between 5

and 9 million dollars. The result suggests that if an organization wish to limit monetary damages due to an attack, they would need to focus on defending against social engineering and unsecured hardware attacks. This because though there are way more misconfiguration attacks, see Figure 7, those attacks cost only 10% of what a social engineering attack cost.

4.3. Future Improvements

The system at the moment has no way of identifying duplicate records. This is because there is not a single identifying characteristic. If wanted an approximation could be used. For example, a record could be marked a duplicate if its date of attack falls within a certain time-range relative to another attack on the same organization. Administrators would then need to manually check whether the attacks are duplicates. This method might not identify all duplicates or alternatively identify too many duplicates when the time-range is too large. To reduce the number of false positives, the criteria might be made stricter, either with a narrower time-range or by including more features in the comparison. Stricter criteria however, have the disadvantage of catching fewer duplicates. The other threat to validity is that we did not get any validation by users. Such validation can be obtained through conducting user studies. Finally, we believe that the visualisation and search features can be further improved.

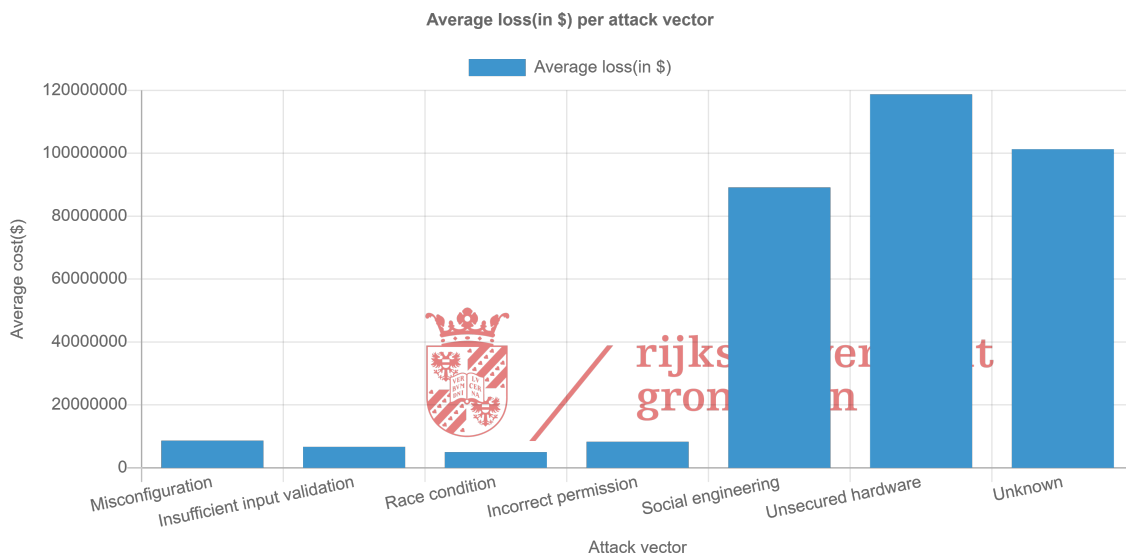


Figure 6. Average money loss per attack vector.

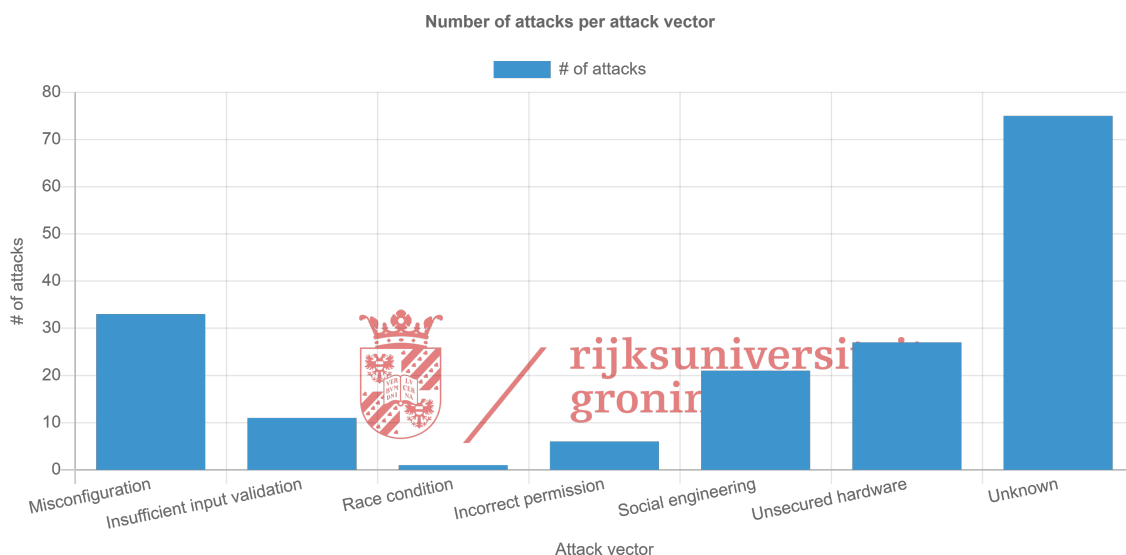


Figure 7. Number of attacks per attack vector.

5. Conclusion

In this paper we looked at different taxonomies for classifying cyber attacks and compared them based on a list of crucial factors. We noted that more emphasis should be put on how to defend against cyber attacks. We therefore extended an existing taxonomy to not only include technical details about a cyber attack, but also how to defend against it. This was done to make it more practically useful for security experts and researchers, as insights about future defense strategies can be gained from the classified cyber attacks.

We then collected a data set of 174 cyber attacks and published them in a web interface to allow easy

interaction with the data set. The web interface is available at: cybersecurityrug.000webhostapp.com/overview.php and the data set is available here [18]. Researchers can interact and download this data set to conduct their own studies. The conclusion of such studies shall be focused on minimizing the risk or choosing the most effective defense strategy.

References

- [1] ENISA, Enisa threat landscape - 2020], url = <https://www.enisa.europa.eu/topics/threat-risk-management/threats-and-trends>, urldate = 2021-02-05.
- [2] IBM, Data breach. URL <https://www.ibm.com/security/data-breach>.

- [3] IBM, Data breach report 2021. URL <https://www.ibm.com/>.
- [4] LI, Y. and LIU, Q. (2021) A comprehensive review study of cyber-attacks and cyber security; emerging trends and recent developments. *Energy Reports* 7: 8176–8186. doi:<https://doi.org/10.1016/j.egy.2021.08.126>, URL <https://www.sciencedirect.com/science/article/pii/S2352484721007289>.
- [5] EASTTOM, C. and BUTLER, W. (2019) A modified mccumber cube as a basis for a taxonomy of cyber attacks. In *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*: 0943–0949.
- [6] SIMMONS, C., ELLIS, C., SHIVA, S., DASGUPTA, D. and WU, C. (2009) Avoidit: A cyber attack taxonomy .
- [7] CORPORATION, M., Mitre atck taxonomy. URL <https://attack.mitre.org/>.
- [8] HOWARD, J.D. and LONGSTAFF, T.A. (1998) A common language for computer security incidents doi:10.2172/751004, URL <https://www.osti.gov/biblio/751004>.
- [9] LOUGH, D.L. A taxonomy of computer attacks with applications to wireless networks : 1–348April 2001.
- [10] DERBYSHIRE, R., GREEN, B., PRINCE, D., MAUTHE, A. and HUTCHISON, D. (2018) An analysis of cyber security attack taxonomies. In *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*: 153–161. doi:10.1109/EuroSPW.2018.00028.
- [11] PAPP, D., MA, Z. and BUTTYAN, L. (2015) Embedded systems security: Threats, vulnerabilities, and attack taxonomy. In *2015 13th Annual Conference on Privacy, Security and Trust (PST)*: 145–152.
- [12] CORPORATION, M., Common vulnerabilities and exposures. URL <https://cve.mitre.org/>.
- [13] KIWIA, D., DEGHANTANHA, A., CHOO, K.K.R. and SLAUGHTER, J. (2018) A cyber kill chain based taxonomy of banking trojans for evolutionary computational intelligence. *Journal of Computational Science* 27: 394–409. doi:10.1016/j.jocs.2017.10.020, URL <http://dx.doi.org/10.1016/j.jocs.2017.10.020>.
- [14] HARRISON, K. and WHITE, G. (2011) A taxonomy of cyber events affecting communities. In *2011 44th Hawaii International Conference on System Sciences*: 1–9.
- [15] ZHU, B., JOSEPH, A. and SASTRY, S. (2011) A taxonomy of cyber attacks on scada systems. In *2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing*: 380–388.
- [16] JOSHI, C. and SINGH, U. (2014) Admit- a five dimensional approach towards standardization of network and computer attack taxonomies. *International Journal of Computer Applications* 100: 30–36. doi:10.5120/17524-8091.
- [17] TIMOTHY SHIMEALL, J.S. (2014) *Introduction to Information Security*, 1st edition (Elsevier). Chapter 2.
- [18] MOHSEN, F. (2022), A data set of 174 cyber attacks that took place between 2010 and 2020 manually classified based on 15 features. doi:10.34894/DAYT8A, URL <https://doi.org/10.34894/DAYT8A>.