JLIS.it vol. 15, no. 2 (May 2024)
ISSN: 2038-1026 online
Open access article licensed under CC-BY
DOI: 10.36253/jlis.it-591

JLIS.it

Towards a semi-automatic classifier of malware through tweets for early warning threat detection*

Claudia Lanza[(a)], **Lorenzo Lodi**[(b)]

a) University of Calabria, https://orcid.org/0000-0002-3018-1987
b) Zanasi & Partners

Contact: Claudia Lanza, claudia.lanza@unical.it; Lorenzo Lodi, lorenzo.lodi@zanasi-alessandro.eu

Received: 29 November 2023; **Accepted:** 18 March 2024; **First Published:** 15 May 2024

ABSTRACT

This paper presents a method for developing a malware ontology structure by detecting malware instances on Twitter. The ontology represents a semi-automatic classifier fed by the data extracted from tweets. In particular, the automatic part of the presented methodology relies on a pattern-based approach to detect trigger expressions leading to new information about malware, whilst the manual one covers the evaluation of the results by domain-experts, who also validate the reliability of the semantic relationships within the ontology framework. We present preliminary results on the application of our methodology to tweets extracted from MalwareBazaar database showing how the documents' collection analysis, through Natural Language Processing (NLP) tasks, can support the knowledge retrieval and documents' classification procedures for building early warning system of detected malware. Results obtained from this research paper within the time framework of 2023 are referred to the previous version of the current social network *X*.

KEYWORDS

Malware; Classification; NLP; Twitter; Text Mining.

* The authors, although they have jointly worked on the paper, have specifically covered the following sections: Claudia Lanza specifically dealt with sections "Methodology", "Corpus design", "Terminological extraction", "Rule-based pattern recognition", "Classification tool", Lorenzo Lodi sections "Related works" and "Results". Section "Conclusion" is to be attributed to both authors.

This paper is the result of the research activies carried out within the framework of the UE co-financing – FSE REACT-EU, PON Ricerca e Innovazione 2014-2020.

Introduction

This paper will describe a preliminary study on a method to detect new data about malware and structure them in an ontology model. The ontology represents a means through which is possible to build a classifier able to structurally organize the knowledge behind the malicious events within the cyber-sphere. The novelty brought from our approach can be envisaged in the source documentation taken into account to construct the malware classifier. More specifically, not only the set of documents considered, i.e., tweets, but also the techniques applied to retrieve the information about the malware can constitute the originality of this work. In detail, we propose an approach which, through the Natural Language Processing (NLP) tasks over the normalized group of tweets is meant to systematize the informative set of obtained data into an ontology framework. The ontology represents the classifier created in a semi-automatic way and can help cyber analysts in creating a conceptual structure to infer knowledge about malicious events as well as in supporting malware triaging operations from a semantic point of view.

Related Works

The detection of new events from Twitter represents a common research branch and usually is focused on the interpretation of tweets' content from a topic-based approach, as Twitter Stand platform created at Maryland University (Sankaranarayanan 2009) shows by capturing the late breaking news from tweets becoming popular topics per each country. Regarding the cyber threats detection from Twitter, Gaglio (2015) proposed an extension of Soft Frequent Pattern Mining (SFPM) through an improved topic detection algorithm with the presentation of Twitter Live Detection Framework (TLDF) able to face the new incoming data from a topic detection perspective. Cordeiro (2012) presented a work on topic inference events from the social platform by using the Latent Dirichlet Allocation topic inference model based on Gibbs Sampling. Concone et al. (2017) also proposed a methodology to detect, and give an alert on, new malware using the data coming from reliable Twitter's subscribers by means of a Bayes naıve classifier. Specifically, they worked with the "Bayes classifier trained on a set of tweets containing an equal number of i) events related to security attacks, viruses, malware, and ii) generic messages", and realized "groups of tweets discussing the same topic, e.g, a new malware infection, are summarized in order to produce an alert". The authoritativeness of users selected by the authors has been based on an "influence metric" which links the users' interaction with the community in terms of retweets, feelings, answers and number of likes. Another study covering cyber threat detection from Twitter is that of Sabottke (2015), where the authors specifically refer to the exploit detection by creating a Twitter-based exploit detector. This system detects on Twitter the use of exploits against known vulnerabilities by looking within the tweets containing texts mentioning vulnerabilities and comparing, as ground truth, to CVE IDs as well as ExploitDB and classifying them using the SVM classifier.

Given the increase in the variants of malware, a resource able to analyze similarities and gather these features as informative elements in a classification structure becomes a valid means for the enhancement of cybersecurity predictive actions. In the literature review, malware classification has been considered as an urgent and evolving study to foster and a wide range of techniques has been proposed within the scientific community. The most common way to identify increasingly complicated malware is signature-based, (Akhtar and Feng 2022) offer a literary review of new machine learning based techniques which aim at analyzing the efficacy of those approaches in the identification of Polymorphic malwares.

The method we hereby propose may show potential benefits for the malware identification independently from their code, but, conversely, by analyzing the contents of tweets dealing with new malware outbreaks. A comprehensive overview of Deep Learning (DL) tasks used to classify malware events is offered by Mathews (2019). Still referring to DL approaches, Kalash (2018) proves the better performance of Convolutional Neural Networks (CNN) in identifying and properly classifying malware, as well as Tekerek and Yapici (2022), Adem and Yapici (2022), and Habibi (2023). Amongst the main tested malware classification procedures, the call graph clustering approach by Kinable and Kostakis (2011) has been executed for the detection of structural similarities between malicious behaviour samples. Echo state networks (ESNs) and recurrent neural networks (RNNs) have been exploited, among a vast number of experts, by Pascanu (2015) to detect malicious files. Annachhatre (2015) applies hidden Markov models and cluster analysis to discover malware classes, whereas Tang (2023) uses LightGBM to identify malware families. Mirza (2018) proposes a combination of machine learning approaches employed over a group of features extracted from a wide corpus made of benign and malicious files through a bespoke feature extraction tool. A set of studies focus on semantics in malware code detection under the lens of obfuscation used by attackers to hide the actual code and the behaviour of malware (Singh 2018; Sahu 2014; Christodorescu 2005). Against this background one cannot fail to mention the main classification systems which guide the comprehension and representation of the malware families, their features and targets in their attacking processes, both released by the Mitre Corporation, an American association that supports government structures specifically with respect to the cybersecurity area. The first one is the MITRE ATT&CK platform, a web-based tool that helps in enhancing knowledge on threat tactics and techniques applicable to several operative systems: it is subdivided into 14 tactics and 188 techniques representing the ways by which attackers can perform a cyber attack against the infrastructures. This tactics' representation supports either the acquisition of a knowledge base in the cyber adversaries' techniques and a dictionary modeling of this information under a classification perspective to overcome cyber threats (Xiong et al. 2022; Georgiadou 2021; Kwon 2020). The second one is the Common Attack Pattern Enumeration (CAPEC) catalogue that gathers under a tree-like configuration a range of attacks' mechanisms and attacks' domain by merging the patterns according to the common features they share (Kotenko and Doynikova 2015). This structure becomes essential to understand the adversaries' behaviours and to create a common dictionary and a classification taxonomy of the attacks' patterns to be used by analysts or developers working within the cyber defense field (Andrei 2019).

In this paper we propose a new method to detect in a predictive way the new malware denomination and classify them by a hierarchical structure through an ontological configuration . The proposed method will provide institutions working within the cybersecurity strategy plans with a classification tool to manage their knowledge base when exposed to cyber threats. This will be empowered by an updated classification system, i.e., an ontology, covering the connections with the malware families' attributes starting from a tweets analysis.

With respect to other malware ontologies existing and largely used in the scientific community (see for instance those of Rastogi 2020; Zareen 2016; Huang 2010), the one proposed in our work will aim at potentially including the zero-day cyber-attack events through the detection of the semantic information within tweets in a father-son classes' relationship and object properties, by exploiting the advantages of OWL language (McGuinness and Van Harmelen 2004; Antoniou and Van Harmelen 2004; Wang 2004). The main contributions of this paper can be summarized as follows: (i) a tweets dataset taken into consideration to realize the classification tool, (ii) the terminological analysis of the texts extracted from Twitter (iii) that leads to (iv) the employment of NLP methods, particularly based on a patterns-based approach (León-Araúz 2013; Auger and Barrière 2008) to find trigger expressions in tweets, used to retrieve in a predictive way the upcoming malware classes within the cybersecurity spectrum.

Methodology

Our methodology relies on a source corpus containing a set of documents constituted by tweets, which, for their intrinsic nature, are marked by a regular configuration as well as by an unstructured way to formulate texts from a linguistic perspective (Arora and Kansal 2019; Kumar and Das 2013). The reason why tweets have been chosen as source corpus is related to the predictive goal of this work towards early malware detection and their subsequent classification: tweets are a real-time information that can allow a punctual acquisition of knowledge to be studied. Indeed, "Tweets are free text micro blogging posts of no more than 140 characters, used by millions of people around the world; with one important characteristic, its real-time nature. Although their length per post is limited, the variety of words that can be used is high. If we take in account that each single word represents a different variable, a tweet is considered a high dimensional data."(-Gutierrez 2014, 168)

Moreover, as observed by Barnard (2016), each tweet owns an inner narrative form of communication that synthetically highlights salient information to be retrieved. Thanks to the character limit behind the posts' publication the information within tweets can be spread in an immediate way without facing the semantic noise that generally implies a massive removal of unnecessary text (Gupta and Rao 2020). On the basis of using the Twitter microblogging as a major source of information (Bakliwal 2012) our approach that leads to a malware predictive detection and their classified configuration can be described by the steps depicted in Figure 1.

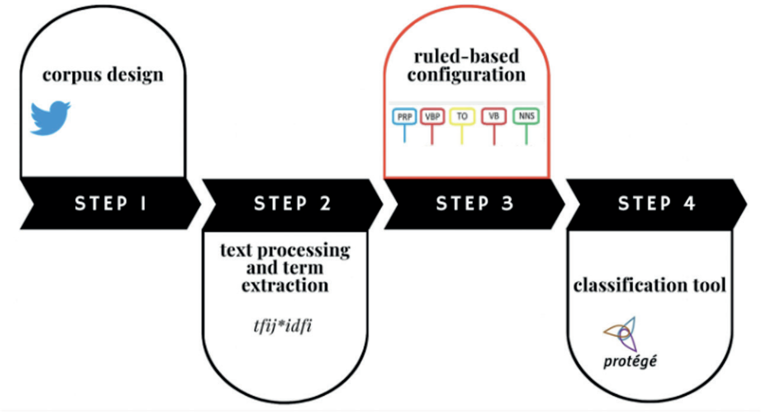


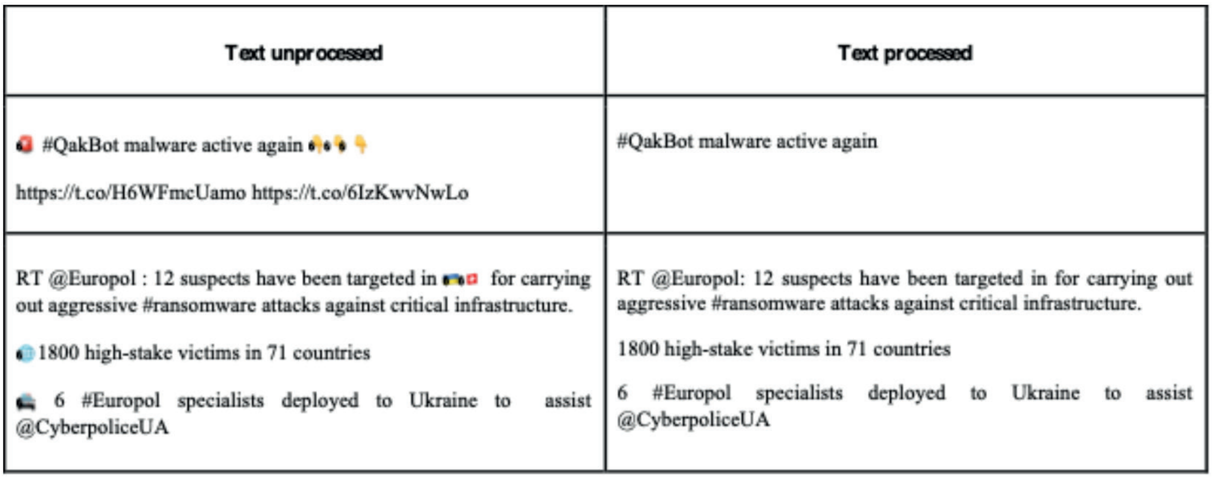
Figure 1. Our methodology steps for tweets processing, extraction and classification

Corpus design

The collected corpus related to the predictive discovery of new malware denominations is composed by a set of tweets' contents published by few reporters identified as a test set to experiment the methodology. Ranking the social activity of Twitter users has represented a research interest in several areas of study and it mainly focuses on the computation of the most influential profiles on Twitter given certain topics. For instance, amongst the others (e.g., Lo 2016; Bartoletti 2016; Drakopoulos 2016; Noro 2013; Subbian and Melville 2011; Das Sarma 2010), Montangero and Furini (2015) proposed a method based on the algorithm TRank that connects the user's activity, i.e., tweets, and the profile itself in order to reveal the user's level of influence; Cappelletti and Sastry (2012) set out a technique based on the IARank ranking model that orders the information about Twitter users in a real-time span, the logic behind it is to compute the average of users' influence by taking into account the retweet and mentions as information amplification sources being both the features proving how a user is likely to be retweeted and mentioned; Yamaguchi (2010) published a work presenting an algorithm called TURank (Twitter User Rank) based on the connection existing between users and tweets, both of them represented by a user-tweet graph. For the purpose of our initial study on how the new malware occurrences are publicly shared on Twitter, alongside the support of experts in the cybersecurity field of knowledge, the platform MalwareBazaar has been chosen as a first resource from which to begin to test the extraction of Twitter users' profiles. These latter usually share information about the new malware generation through their posts on social media. This platform offers to cyber analysts statistical means through which it is possible to be informed about the latest cyber threat reported by determined users identified as Top Reporters. Therefore, the first task addressed the crawling of tweets published by the main profiles indicated as 'reporters' on the MalwareBazaar portal, which are in total 10597. The crawling executed through a custom Twitter API client gave in output a list of files, each one of them including a list of metadata about users' activity in a tabular format. The columns within each file contain: tweet Id, Text, Name, Screen Name, Date, Favorites, Retweets, Language, Client, Tweet type (e.g., retweet, reply, tweet), URL, number of Hashtags, number of Mentions, Media type (e.g., photo) and Media URLs. The generated files are then parsed in a next step. In the extraction phase, just the column referred to the tweets' content (Text) from the crawling output has been performed. Each column extracted

JLIS.it vol. 15, no. 2 (May 2024)
ISSN: 2038-1026 online
Open access article licensed under CC-BY
DOI: 10.36253/jlis.it-591

has represented a separate file considered as a single document to put into the source corpus to be semantically analyzed. For instance, the tweet text column of a selected user, e.g., tolisec user, has represented a single document containing the 121 tweets published by this user. Successively, through Python, specifically with NLKT and SPACY packages, the texts have been cleaned, this step specifically addressed the removal of stopwords as well as of symbols and emoticons in order to make the documents processable for the term extraction tool, as shown in Table 1.



Text unprocessed	Text processed
🚨 #QakBot malware active again 👇👇👇 https://t.co/H6WFmcUamo https://t.co/6IzKwvNwLo	#QakBot malware active again
RT @Europol : 12 suspects have been targeted in 🇺🇦🇨🇭 for carrying out aggressive #ransomware attacks against critical infrastructure. 🌐1800 high-stake victims in 71 countries 🚓 6 #Europol specialists deployed to Ukraine to assist @CyberpoliceUA	RT @Europol: 12 suspects have been targeted in for carrying out aggressive #ransomware attacks against critical infrastructure. 1800 high-stake victims in 71 countries 6 #Europol specialists deployed to Ukraine to assist @CyberpoliceUA

Table 1. Details of text processing of collected tweets

Figure 2 depicts the steps followed after the compilation of the source corpus and the employment of the text included in the tweets to be used as a starting point from which to begin the classification process.

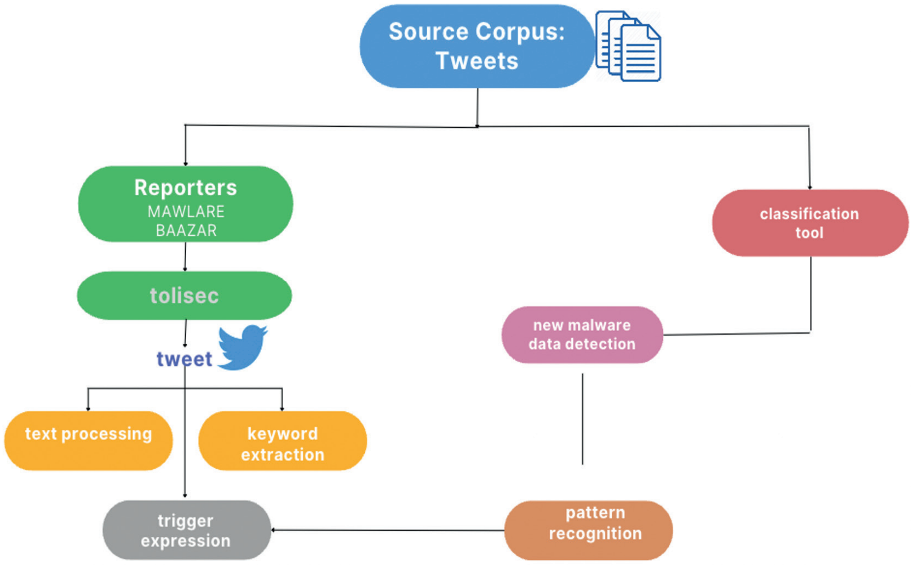


Figure 2. A comprehensive view of the processing and classification steps of the tweets[1]

[1] Tolisec is a Twitter user included in the graph as to provide a starting use-case.

Terminological extraction

The term extraction has been realized through the software SketchEngine (Kilgarriff 2008; Jakubíček et al. 2014), a corpus analytic tool which gives in output a ranked list of the most representative terms included in the source corpus. Indeed, by using a semantic extractor it is possible to analyze the knowledge domain information under a terminological perspective and see which are the most frequent terms in the documents selected to reflect the information under study and apply reasoning techniques. For what concerns the frequency and the relevance of terms with respect to a specialized corpus, we used the Term Frequency Inverse Document Frequency (TF/IDF) (Qaiser 2018) measure. This formula allows to have in the first position terms that are very specific to the domain under study, in this case the cybersecurity one, and in the last one those most commonly used in the general language. This measure supported the identification of the most representative lexical units and the building of a network of co-occurrences that have guided the systematization of the trigger expressions. Indeed, these latter have constituted the semantic means to discover new information about malware denominations in tweets meant to represent the new entities within the classification system.
Table 2 shows an extract of the lemmas retrieved in the source documentation through the integration of a stopword list in English and the definition of the minimum frequency threshold at 1 in order to have as much terms as possible to detect.

Item	Frequency
malware	735
email	710
url	642
sample	628
com	601
new	587
exe	579
use	545
dll	531
payload	508
file	399
domain	399
report	386
zip	357
host	317
see	307
abuse	298
ransomeware	294

Table 2. Term extraction output

Rule-based pattern recognition

The next step addresses the definition of the trigger expressions by relying on the terminological analysis, followed by their normalization through a rule-based pattern recognition (Babicm 2008; Anicic 2010). In detail, the objective of this phase is strictly connected to the one of realizing a classification tool on malware because it is aimed at establishing a set of rules to be applied over the corpus made up of tweets to retrieve new information about malware to be included into the classification tool, i.e., ontology, through the terms in the list obtained by the extraction (relying on the TF/IDF scores). The process continued by checking the co-occurrences identified within the terms in the source corpus. In detail, each term in SketchEngine can be analyzed according to the concordance terms show in the semantic distribution thanks to the syntactic connective structures within the source corpus through the Word Sketch function.

For instance, according to the output given by the semantic tool, the first two terms identified to generate a first list of regular expressions, alongside the support of domain experts, have been malware and ransomware. Therefore, the collocations represented the key combinations that could give back a networked kind of knowledge information through which to detect new names or to empower the cyber attack classification. A list of few collocations used in this preliminary stage is the following:

- Modifiers of malware: infect victims with – using a, the most resilient, the most dangerous/interesting malware
- Malware + verbs: attack (malware attacking), call (malware (also) called), identify (malware identified as)

These collocations have been used as trigger expressions to run the automatic identification of unknown malware and information about them meant to be included in the classification tool. Each of these expressions has been transformed into a syntactic structure following the rule-based pattern recognition (Xu and Cai 2021). In this way the automatic extractor tool is able to detect the morpho-syntactic structures to be mapped with the new information to be retrieved and populate the malware classifier in the form of ontology. Once understood that terms are accompanied by expressions that can lead to knowledge-domain discovery (Ervert 2008), the following phase of this research activity covered the creation of the trigger expressions to be implemented in the tool in order to represent the training set for the alert on new malware denominations sorted by time. The definition of the rule-based pattern recognition system is based on the exploitation of the regular expressions, created with the use of SpaCy library (Vasiliev 2020) in Python. This configuration has been based on the discovery of new denominations of malware starting from the information contained in the documents (the tweets of the users selected through the Malware Bazaar platform) interpreted under the lens of clues to be considered as alerts of new data to import. In contrast to the fixed pattern matching of regular expressions, this method allows us to match tokens according to some pre-set patterns. Additionally, it includes features such as parts-of-speech analysis, entity types, dependency parsing, lemmatization, and a great deal more. In addition to this, this further bolsters regular expression patterns. The token pattern matcher provided by the SpaCy library takes advantage of the word level features proper to this linguistic toolkit such as LOWER, LENGTH, LEMMA, and SHAPE as well as flags, such as IS PUNCT, IS DIGIT, LIKE URL. An input text may be given and rules can be defined in order to parse the text and determine whether or not it includes the appropriate morpho-syntactic objects in the appropriate sequence. In order to provide some practical examples of results obtained by our method, we present the outputs from the execution of the first group included in the list of regular expressions presented in Section 2.4. The sentence "infect victims with... malware" represents the starting point from which to discover new information about the malware being discussed. The resulting information retrieved by executing the tool over the corpus will be the key entity to be integrated in the classification system. In detail, regarding the case of the aforementioned sentence, the pattern-base code instructs the SpaCy library to recognize sentences that begin with one or more verbs, followed by one or more nouns, and by one or more prepositions or postpositions, possibly a determiner, the "malware" lemma, and then end with a noun that will be identified as the new potential malware name in this case. This enables us to identify sentences that contain different words but are constructed in the same syntactic manner.

Results

As a consequence of our work, various instances of patterns as well as their representation in the form of spacey regexes are hereby presented. These regexes have the potential to be utilized in the filtering of pre-processed and normalized tweets. The following short list represents an overview of the source expressions, selected under the basis of the terminological analysis outcomes, used to construct the patterns to be employed over the crawled tweets for the detection of new malware-related data.

Expression	**RegEx**
(malware) infect victims with…	*NOUN+ VERB (infect) + NOUN+ADP+NOUN*
(malware) known as... / (malware) formerly known as…	*NOUN+ VERB|ADV+VERB (know) +ADP (as) +PROPN*
identified as…	*VERB (Identify)+ADP*
encrypted by	*VERB (encrypt) + ADP + NOUN*

Classification tool

In the literature many malware classification schemes have been configured, such as the Common Taxonomy for law Enforcement and the National Network of CSIRTs published by the Europol Public Information, which describes a range of incidents according to their class and type and then support the "mapping each type of incident with the pertinent article of the international legislative framework" (Euripol Public Information 2017, 5); MISP Taxonomies ; CIRCL taxonomy schemes of classification in Incident response and detection ; OSINT Open Source Intelligence; The VERIS Framework, Vocabulary for event recording and Incident sharing; Kaspersky which reports the types of malware by behaviors.

Our proposal relies on the construction of an ontology structure as a classification tool starting from the entities discovered by executing the pattern-based approach to the tweets' contents. Ontology is a: "[...] hierarchically structured set of concepts describing a specific domain of knowledge that can be used to create a knowledge base." (Blomqvist and Sandkuhl 2005, 1)

With reference to the meaning of ontology in the informatics area, Gaurino (2009, 2) gives a clear definition by stating that "Computational ontologies are a means to formally model the structure of a system, i.e., the relevant entities and relations that emerge from its observation, and which are useful to our purposes. An example of such a system can be a company with all its employees and their interrelationships. The ontology engineer analyzes relevant entities and organizes them into concepts and relations, being represented, respectively, by unary and binary predicates. The backbone of an ontology consists of a generalization/specialization hierarchy of concepts, i.e., a taxonomy."

An interesting study, specifically oriented to the cybercrime field using ontologies, has been reported by Donalds and Osei-Bryson (2019) who, besides offering an extensive overview of the cybercrime classification schemes existing in the literature, describe in a practical way the realization of a high-level ontology for the cybercrime events, specifically called cybercrime classification ontology (CCO) through the use of Protégé platform (Sivakumar and Arivoli 2011) as our study will do. The author starts by isolating the main cybercrime-related concepts (i.e., attack event, vulnerability, tool and technique, objective, offence, location, complainant, victim, target, impact, attacker) and continues by enhancing the parent-child relationships through the use of the object properties which help in improving the attack events classification.

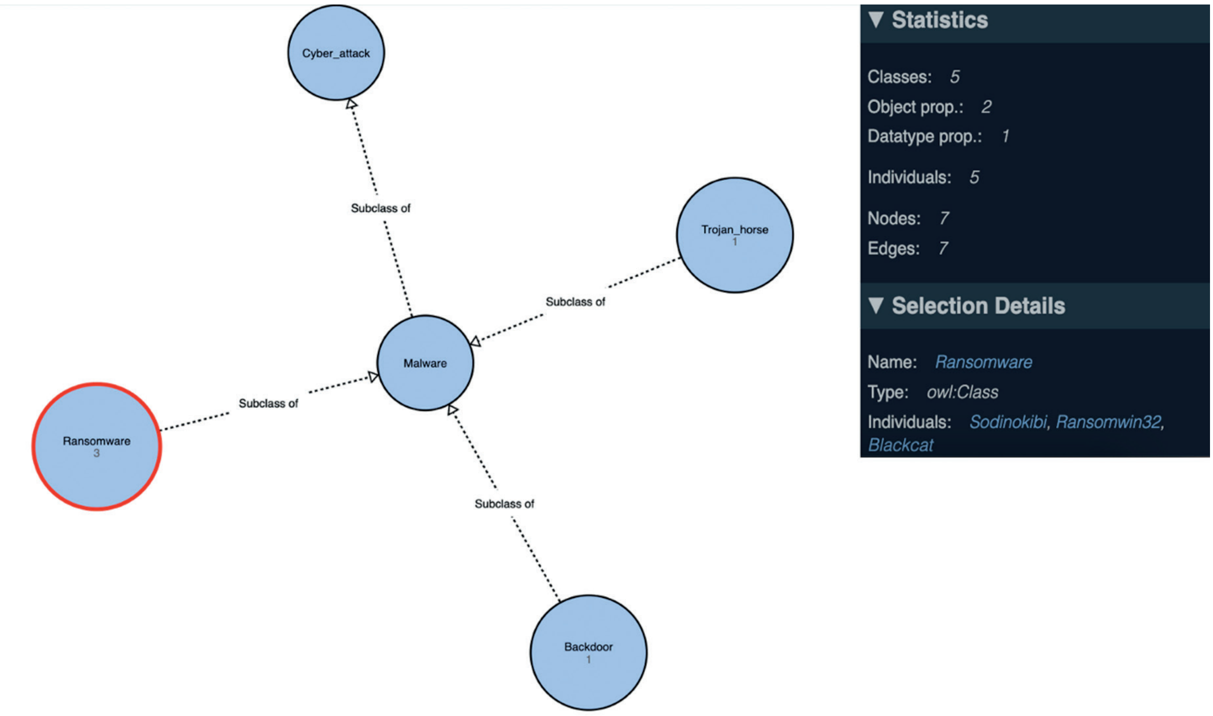
Our work regarded the construction of the classification scheme on the basis of tweets' analysis where the structure of the ontology follows the hierarchical configuration of classes and sub-classes and the association of each new malware discovered as an individual, as the following example demonstrates: Cyber_attacks hasSubclass Malware; Malware hasSubclasses Zombie, Crypro_miner_malware, Trapdoor, Trojan_horse, Banking_malware, Virus, Logic Bomb, Worm, Ransomware. In this regard, we targeted the inclusion of new denominations of malware in the classification scheme as the extreme leaf of this tree-like configuration, hence the matching will be between, for instance, Banking_malware hasIndividual: new name of worm. This process has been executed by running the semantic analysis over the tweets and confirmed by cyber defense experts as well as by relying on the information included in the aforementioned malware classification tools.

Indeed, through the implementation of the regular expressions we have obtained encouraging results, some of which are the following:

pattern	tweet	pattern retrieved	new information	Ontology
NOUN+ VERB|ADV+VERB (know) +ADP (as) +PROPN	@msftsecintel: we have detected and are **now** blocking a **new family** of ransomware being used after an initial compromise of unpatched on-premises exchange servers. microsoft protects against this threat known as ransom:win32/doejocrypt.a, and also as dearcry.	threat known as ransomwin32	ransomwin32	*Class: Cyber_attacks* *Sub-Class*: Malware *Sub-sub-Class*: Ransomware *Individual:* Ransomwin32
NOUN+ VERB|ADV+VERB (know) +ADP (as) +PROPN	rt @craiu : while looking at the #solarwinds #sunburst backdoor, **we discovered** several features that overlap with a previously identified backdoor known as kazuar, used by turla. our analysis: https://t.co/3ef6y2g5ly #unc2452 #darkhalo	backdoor known as kazuar	kazuar	*Class*: *Cyber_attacks* *Sub-Class*: Backdoor *Individual:* Kazuar
NOUN+ VERB (infect) + NOUN+ADP+NOUN	malspam uses national bank of argentina (banco de la nación argentina) as a lure to infected user with #nanocore rat exe:https://t.co/i3pusmwnws nanocore rat c2:185.140.53.11:6532 inetnum: 185.140.53.0 - 185.140.53.255 netname: freedom_of_speech_vpn https://t.co/ng7yee4ajx	malspam infect user with #nanocore rat	nanocore rat	*Class: Cyber_attacks* *Sub-Class*: Malware *Sub-sub-Class*: Spam *seeAlso* Malspam[2] *Class: Cyber_attacks* *Sub-Class*: Malware *Sub-sub-Class*: Trojan *Individual:* Nanocore rat Object Property: Malspam *USES* Nanocore rat
VERB (encrypt) + ADP + NOUN	@govcert_ch : don't open fax messages pretending to come per email within a zip file, or you will have your data encrypted by ransomware family sodinokibi (e.g. email from hellofax with subject "sie haben ein fax"). https://t.co/ai6unad2bv	encripted by ransomware family sodinokibi	sodinokibi	*Class: Cyber_attacks* *Sub-Class*: Malware *Sub-sub-Class*: Ransomware *Individual:* Sodinokibi Object Property: Data *EncriptedBy* Sodinokibi
VERB (identify)+ADP	@cyb3rops @n0le_ptr your yara rule **hits** on 10 samples on malware bazaar by @abuse_ch all identified as blackcat before :+1: : https://t.co/bl6duhhuwi	10 samples identified as blackcat	blackcat	*Class: Cyber_attacks* *Sub-Class*: Malware *Sub-sub-Class*: Ransomware *Individual:* Blackcat

[2] SeeAlso stands for the synonymy relationship.

s demonstrated in these examples another element that can contribute to the retrieval of novelty trait can be the attention to some verbs, e.g., discover or hit, adjectives, e.g., new, or adverbs such as ‘now’ held up by gerundive constructions (‘are now blocking’), which can lead to the inference of current data meant to be included in the ontology. The ontology offers a comprehensive way to represent the classified information by structuring it according to the generic-specific principle and associating to each class a set of instances, in our case study the new malware denominations, as Figure 3 depicts.



▼ Statistics

Classes: 5
Object prop.: 2
Datatype prop.: 1

Individuals: 5

Nodes: 7
Edges: 7

▼ Selection Details

Name: Ransomware
Type: owl:Class
Individuals: Sodinokibi, Ransomwin32, Blackcat

Figure 3 Ontology structure for malware instances[3]

The object properties expressed through OWL language relate pairs of entities (Glim 2014) and are the means of organizing the informative data by specific connections that explicit the conceptualization of the knowledge base. This will be the focus of our next research activity which will be performed using the additional data we crawled from Twitter.

Conclusion

This paper develops a new method to predict malware appearance through the analysis of tweets shared by active Twitter users identified in the MalwareBazaar database and spreading information on new malware instances. In this work we conducted a semantic analysis of the isolated users’ tweets after a crawling operation and the configuration of a set of identified trigger expressions, normalized in the form of regular expressions, that are used as a knowledge inference task

[3] Realized with WebVowl platform (Lohmann 2015) https://service.tib.eu/webvowl/

to tune data. The establishment of the trigger expressions followed a terminological selection of terms to be used as morphosyntactic units within the regular expression rules. We then exploited the entities retrieved by the trigger expressions over the tweets dataset to be used to define a classification tool. The classifier is considered to represent the connections of new malware, detected by implementing the above-mentioned steps, show within a hierarchical structure.

One of the future perspectives will address the continuous enhancement of the malware classifier (ontology instances) which will be also fed by the CSIRT *Settimana cibernetica* enabling the mapping with the new attacks included in the MITRE official framework of CAPEC and the associated vulnerabilities. This triangular interconnection could support the association of new types of malwares with existing networked semantic flow of information related to the vulnerabilities present in the hardware, software or protocols infrastructures. This activity could represent a forecastable knowledge platform to be used by companies when it comes to considering the elements meant to be analyzed to reduce the risk of being exposed to cyberthreats.

References

Adem, Tahir, and Muhammed Mutlu Yapici. 2022. "A Novel Malware Classification and Augmentation Model Based on Convolutional Neural Network." *Computers & Security* 112. https://doi.org/10.1016/j.cose.2021.102515.

Akhtar, Muhammad Shoaib, and Tao Feng. 2022. "Malware Analysis and Detection Using Machine Learning Algorithms." *Symmetry* 14 (11): 2304.

Andrei, Brazhuk. 2019. "Semantic Model of Attacks and Vulnerabilities Based on CAPEC and CWE Dictionaries." *International Journal of Open Information Technologies* 7 (3): 38-41.

Anicic, Darko, Paul Fodor, Sebastian Rudolph, Roland Stühmer, Nenad Stojanovic, and Rudi Studer. 2010. "A Rule-Based Language for Complex Event Processing and Reasoning." In *Web Reasoning and Rule Systems. RR 2010. Lecture Notes in Computer Science*, edited by Pascal Hitzler, and Thomas Lukasiewicz, vol 6333, 4: 42–57. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-15918-3_5.

Annachhatre, Chinmayee, Thomas H. Austin, and Mark Stamp. 2015. "Hidden Markov Models for Malware Classification." *Journal of Computer Virology and Hacking Techniques* 11: 59–73. https://doi.org/10.1007/s11416-014-0215-x.

Antoniou, G., van Harmelen, F. (2004). "Web Ontology Language: OWL". In *Handbook on Ontologies. International Handbooks on Information Systems*, edited by Steffen Staab, and Rudi Studer. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-24750-0_4.

Arora, Monika, and Vineet Kansal. 2019. "Character Level Embedding with Deep Convolutional Neural Network for Text Normalization of Unstructured Data for Twitter Sentiment Analysis." *Social Network Analysis and Mining* 9: 12. https://doi.org/10.1007/s13278-019-0557-y.

Auger, Alain, and Caroline Barrière. 2008. "Pattern-based Approaches to Semantic Relation Extraction: A State-of-the-Art." *Terminology* 14 (1). https://doi.org/10.1075/term.14.1.02aug.

Babic, Bojan, Nenad Nesic, and Zoran Miljkovic. 2008. "A Review of Automated Feature Recognition with Rule-based Pattern Recognition." *Computers in Industry* 59 (4): 321–337.

Akshat Bakliwal, Piyush Arora, Senthil Madhappan, Nikhil Kapre, Mukesh Singh, and Vasudeva Varma. 2012. "Mining Sentiments from Tweets." In *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*, 11–18. Jeju, Korea: Association for Computational Linguistics.

Barnard, Josie. 2016. "Tweets as Microfiction: On Twitter's Live Nature and 140-Character Limit as Tools for Developing Storytelling Skills." *New Writing* 13 (1): 3–16. https://doi.org/10.1080/14790726.2015.1127975.

Bartoletti, Massimo, Stefano Lande, and Alessandro Massa. 2016. "Faderank: An Incremental Algorithm for Ranking Twitter Users." In *Web Information Systems Engineering–WISE 2016: 17th International Conference, Shanghai, China, Proceedings, Part II 17*, 55–69. Springer International Publishing.

Blomqvist, Eva, and Kurt Sandkuhl. 2005. "Patterns in Ontology Engineering: Classification of Ontology Patterns." *ICEIS* 3: 413–416.

Brazhuk, Andrei. 2019. "Semantic Model of Attacks and Vulnerabilities Based on CAPEC and CWE Dictionaries." *International Journal of Open Information Technologies* 7(3): 38–41.

Cappelletti Rafael, and Sastry Nishanth. 2012. "IARank: Ranking Users on Twitter in Near Real-Time, Based on Their Information Amplification Potential." *International Conference on Social Informatics,* 70–77. Alexandria, VA, USA. https://doi.org/10.1109/SocialInformatics.2012.82.

Christodorescu, Mihai, Sanjit Jha, Sanjit A. Seshia, Dawn Song, and Randal E Bryant. 2005. "Semantics-Aware Malware Detection." *IEEE Symposium on Security and Privacy (S&P'05)*, Oakland, CA, USA, 2005, 32–46. https://doi.org/10.1109/SP.2005.2032–46.

Concone, Mário. 2012. "Twitter Event Detection: Combining Wavelet Analysis and Topic Inference Summarization." DSIE'12, Doctoral Symposium on Informatics Engineering, 1: 11–16.

Das Sarma, Anish, Atish Das Sarma, Sreenivas Gollapudi, and Rina Panigrahy. 2010. "Ranking Mechanisms in Twitter-Like Forums." In *Proceedings of the Third ACM International Conference on Web Search and Data Mining WSDM'10, 21–30, February 4-6*. New York City, New York, USA: Association for Computer Machinery.

Das, Tushar Kant, and P. Mohan Kumar. 2013. "BIG Data Analytics: A Framework for Unstructured Data Analysis." *International Journal of Engineering and Technology* 5: 153–156.

Donalds, Charlette, and Kweku-Muata Osei-Bryson. 2019. "Toward a Cybercrime Classification Ontology: A Knowledge-Based Approach." *Computers in Human Behavior* 92: 403–418.

Drakopoulos, Georgios, Andreas Kanavos, and Athanasios K Tsakalidis. 2016. "Evaluating Twitter Influence Ranking with System Theory." *WEBIST* 1: 113–120.

Europol Public Information. 2017. "Common Taxonomy for Law Enforcement and The National Network of CSIRTs." https://www.europol.europa.eu/cms/sites/default/files/documents/common_taxonomy_for_law_enforcement_and_csirts_v1.3.pdf.

Evert, Stefan. 2008. "Corpora and Collocations." In *Corpus Linguistics: an international handbook 2*, 1212–1248. Berlin, New York: De Gruyter Mouton.

Gaglio, Salvatore, Giuseppe Lo Re, and Marco Morana. 2016. "A Framework for Real-Time Twitter Data Analysis." *Computer Communications* 73: 236–242.

Georgiadou, Anna, Spiros Mouzakitis, and Dimitris Askounis. 2021. "Assessing MITRE AT-T&CK Risk Using a Cyber-Security Culture Framework." *Sensors* 21(9): 3267.

Glimm, Birte, Ian Horrocks, Boris Motik, Rob Shearer, and Giorgos Stoilos. 2012. "A Novel Approach to Ontology Classification." *Journal of Web Semantics* 14: 84–101.

Guarino, Nicola, Daniel Oberle, and Steffen Staab. 2009. "What Is an Ontology?." *Handbook on Ontologies* 1–17. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-92673-3.

Gupta, Rishabh, and Rajesh N Rao. 2020. "Towards Semantic Noise Cleansing of Categorical Data Based on Semantic Infusion." https://doi.org/10.48550/arXiv.2002.02238.

Gutierrez, Carlos Enrique, Mohammad Reza Alsharif, Katsumi Yamashita, and Mahdi Khosravy. 2014. "A Tweets Mining Approach to Detection of Critical Events Characteristics Using Random Forest." *Int J Next-Gener Comput* 5(2): 167–176.

Habibi, Omar, Mohammed Chemmakha, and Mohamed Lazaar. 2023. “Performance Evaluation of CNN and Pre-trained Models for Malware Classification.” *Arabian Journal for Science and Engineering*: 1–15.

Huang, Hsien-Der, Tsung-Yen Chuang, Yi-Lang Tsai, and Chang-Shing Lee. 2010. “Ontology-based Intelligent System for Malware Behavioral Analysis.” In *International Conference on Fuzzy Systems*, 1–6, Barcelona, Spain. doi: 10.1109/FUZZY.2010.5584325.

Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý, and Vít Suchomel. 2014. “Finding Terms in Corpora for Many Languages with the Sketch Engine.” In *Proceedings of the demonstrations at the 14th conference of the european chapter of the association for computational linguistics*, 56-56. Gothenburg, Sweden: Association for Computational Linguistics. https://doi.org/10.3115/v1/E14-2014.

Kang, Boojoong, KimTaekeun, Heejun Kwon, Yangseo Choi, and Eul Gyu Im. 2012. “Malware Classification Method via Binary Content Comparison.” In *Proceedings of the 2012 ACM Research in Applied Computation Symposium*, 316–321, New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2401603.2401672.

Kalash, Mahmoud, Mrigank Rochan, Noman Mohammed, Neil D.B. Bruce, Yang Wang, and Farkhund Iqbal. 2018. “Malware Classification with Deep Convolutional Neural Networks.” In *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, 1–5. Paris, France. https://doi.org/10.1109/NTMS.2018.8328749.

Kilgarriff, Adam, Pavel Rychlý, Pavel Smrž, and David Tugwell. 2008. “The Sketch Engine.” *Practical lexicography: a reader*: 297–306.

Kinable, Joris, and Orestis Kostakis. 2011. “Malware classification based on call graph clustering.” *Journal in Computer Virology* 7(4): 233–245. https://doi.org/10.1007/s11416-011-0151-y.

Kotenko, Igor, and Elena Doynikova. 2015. “The CAPEC based generator of attack scenarios for network security evaluation.” In *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, 436–441. Warsaw, Poland. https://doi.org/10.1109/IDAACS.2015.7340774.

Kwon, Roger, Ashley Travis, Jerry Castleberry, Penny Mckenzie, and Sri Nikhil Gupta Gourisetti. 2020. “Cyber Threat Dictionary Using MITRE ATT&CK Matrix and NIST Cybersecurity Framework Mapping.” *Resilience Week (RWS)*, 106–112.

León-Araúz, Pilar, Antonio San Martín, and Pamela Faber. 2016. “Pattern-based Word Sketches for the Extraction of Semantic Relations.” In *Proceedings of the 5th International workshop on Computational Terminology (Computerm2016)*, 73–82. Osaka, Japan.

Lo, Siaw Ling, Raymond Chiong, and David Cornforth. 2016. “Ranking of High-value Social Audiences on Twitter.” *Decision Support Systems* 85: 34–48.

Lohmann, Steffen and Vincent Link, Eduard Marbach, and Stefan Negru. 2015. “WebVOWL: Web-based Visualization of Ontologies.” In *Knowledge Engineering and Knowledge Management: EKAW 2014 Satellite Events, VISUAL, EKM1, and ARCOE-Logic, Linköping, Sweden, November 24-28, 2014. Revised Selected Papers,* 19: 154–158. Springer International Publishing.

Mathews, Sherin Mary. 2019. "Explainable Artificial Intelligence Applications in NLP, Biomedical, and Malware Classification: A Literature Review." *Intelligent Computing. CompCom 2019. Advances in Intelligent Systems and Computing*, 998. Cham: Springer. https://doi.org/10.1007/978-3-030-22868-2_90.

Mirza, Qublai K. Ali., Irfan Awan, and Muhammad Younas. 2018. "CloudIntell: An Intelligent Malware Detection System." *Future Generation Computer Systems* 86: 1042–1053.

Montangero, Manuela, and Marco Furini. 2015. "Trank: Ranking Twitter Users According to Specific Topics." In *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*, 767–772. Las Vegas, NV, USA. https://doi.org/10.1109/CCNC.2015.7158074.

Noro, Tomoya, Fei Ru, Feng Xiao, and Takehiro Tokuda. 2013. "Twitter User Rank Using Keyword Search." *Information Modelling and Knowledge Bases XXIV. Frontiers in Artificial Intelligence and Applications* 251: 31–48.

Pascanu, Razvan, Jack W. Stokes, Hermineh Sanossian, Mady Marinescu, and Anil Thomas. 2015. "Malware Classification with Recurrent Networks." In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP*), 1916-1920. South Brisbane, QLD, Australia. https://doi.org/10.1109/ICASSP.2015.7178304.

Qaiser, Shahzad, and Ramsha Ali. 2018. "Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents." *International Journal of Computer Applications* 181(1): 25–29.

Rastogi, Nidhi, Sharmishtha Dutta, Mohammed J. Zaki, Alex Gittens, and Charu Aggarwal. 2020. "Malont: An Ontology for Malware Threat Intelligence." In *International Workshop on Deployable Machine Learning for Security Defense*, 28–44. Cham: Springer International Publishing.

Sabottke, Carl, Octavian Suciu, and Tudor Dumitraş. 2015. "Vulnerability Disclosure in the Age of Social Media: Exploiting Twitter for Predicting {Real-World} Exploits." In *24th USENIX Security Symposium (USENIX Security 15)*, 1041–1056.

Sahu, Manish Kumar, Manish Ahirwar, and A. Hemlata. 2014. "A Review of Malware Detection Based on Pattern Matching Technique." *International Journal of Computer Science and Information Technologie*s (*IJCSIT)* 5 (1): 944–947.

Sankaranarayanan, Jagan, Hanan Samet, Benjamin E. Teitler, Michael D. Lieberman, and Jon Sperling. 2009. "Twitterstand: News in Tweets." In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 42–51. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1653771.1653781.

Singh, Jagsir, and Jaswinder Singh. 2018. "Challenge of Malware Analysis: Malware Obfuscation Techniques." *International Journal of Information Security Science* 7(3): 100–110.

Sivakumar, Ramakrishnan, and P.V. Arivoli,. 2011. "Ontology Visualization PROTÉGÉ Tools–A Review." *International Journal of Advanced Information Technology (IJAIT)* 1: 1-11. http://dx.doi.org/10.5121/ijait.2011.1401.

Subbian, Karthik, and Prem Melville. 2011. "Supervised Rank Aggregation for Predicting Influencers in Twitter." In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust* *and 2011 IEEE Third International Conference on Social Computing*, 661–665. Boston, MA, USA. https://doi.org/10.1109/PASSAT/SocialCom.2011.167.

Vasiliev, Yuli. 2020. *Natural Language Processing with Python and spaCy: a practical introduction*. San Francisco, California, USA: No Starch Press.

Tang, Yonghe, Xuyan Qi, Jing Jing, Liu Chunling, and Weiyu Dong. 2023. "BHMDC: A Byte and Hex N-gram Based Malware Detection and Classification Method." *Computers & Security* 103118.

Tekerek, Adem, and Muhammed Mutlu Yapici. 2022. "A Novel Malware Classification and Augmentation Model Based on Convolutional Neural Network*." Computers & Securit*y 112: 102515.

Zareen, Syed, Padia Ankur, Tim Finin, Lisa Mathews, and Joshi Anupam. 2016. "UCO: A Unified Cybersecurity Ontology." In *Workshops at the Thirtieth AAAI Conference on Artificial Intelligence*. Palo Alto, California, USA: AAAI Press.

Wang, Xiao Hang, D. Qing Zhang, Tao Gu, and Hung, Keng Pung. 2004. "Ontology-Based Context Modeling and Reasoning Using OWL." In *IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second*, 18–22. Orlando, FL, USA. https://doi.org/10.1109/PERCOMW.2004.1276898.

Xiong, Wenjun, Emeline Legrand, Oscar Åberg, and Robert Lagerström. 2022. "Cyber Security Threat Modeling Based on the MITRE Enterprise ATT&CK Matrix." *Software and Systems Modeling* 21.1: 157–177.

Xu, Xin, and Hubo Cai. 2021. "Ontology and Rule-Based Natural Language Processing Approach for Interpreting Textual Regulations on Underground Utility Infrastructure." *Advanced Engineering Informatics* 48, 101288.

Yamaguchi, Yuto, Tsubasa Takahashi, Toshiyuki Amagasa, and Hiroyuki Kitagawa. 2010. "Turank: Twitter User Ranking Based on User-Tweet Graph Analysis." In *Web information systems engineering–WISE 2010: 11th International Conference, Hong Kong, China, December 12-14, 2010. Proceedings, 11*, 240–253. Springer Berlin Heidelberg.