Information Security in Education and Practice

Information Security in Education and Practice

Edited by

Kalinka Kaloyanova

Cambridge
Scholars
Publishing

Information Security in Education and Practice

Edited by Kalinka Kaloyanova

This book first published 2021

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-6066-X
ISBN (13): 978-1-5275-6066-6

TABLE OF CONTENTS

Foreword vii

Chapter One 1
Cybersecurity Tools for Threat Intelligence and Vulnerability Monitoring for National and Sectoral Analysis
George Sharkov, Yavor Papazov, Christina Todorova, Georgi Koykov and Georgi Zahariev

Chapter Two 19
CVE Annotation
Vladimir Dimitrov

Chapter Three 29
Model of a Healthcare Integrated System and Solving Security and Privacy Matters in Personal Health Records
Snezana Savoska, Ilija Jolevski, Blagoj Ristevski, Natasha Blazeska-Tabakovska, Andrijana Bocevska, Vassilis Kilintzis, Ivan Chorbev and Boro Jakimovski

Chapter Four 48
Towards Requirements for Cybersecurity Curricula
Kalinka Kaloyanova

Chapter Five 61
Basic Principles of the Planning of Cybersecurity Training for Students of Computer and Software Engineering Specialties
Roumen Trifonov, Ognian Nakov, Georgi Popov, Slavcho Manolov, Georgi Tsochev, and Galya Pavlova

Chapter Six 78
Design of a Bulgarian National Academic Cyber Network in the NATO and EU Context
Georgi Penchev and Velizar Shalamanov

Chapter Seven 99
SSH Security Improvement with Blockchain
Ivaylo Chenchev and Adelina Aleksieva-Petrova

Chapter Eight 114
Requirements for Securing User Data in Android Applications
at a Software Level
Peter Sabev and Milen Petrov

Chapter Nine 130
Analysis of RPKI Adoption from the Perspective of AS5421
Nikolay Nikolov and Stefan Dimitrov

FOREWORD

The growth of cybersecurity issues affects all aspects of our life. A prime example is the rise of cyberattacks, which has increased the political, business, and national interest in finding different ways to prevent and resolve them. To meet the increasing threat of cybercrime and cyberattacks and the growing societal and commercial demand for a safer digital environment, numerous initiatives at a national and international level have been raised.

This book presents extended versions of papers among the contributions of the First Workshop on Information Security (ISec2019), held in Sofia, Bulgaria, in September 2019. The workshop was organised under the umbrella of the 9th Balkan Conference in Informatics (BCI 2019). The main objective of the workshop was to be a forum for scholars, doctoral students, and practitioners interested in the information security sphere, allowing the exchange of ideas and encouraging collaboration among universities and other institutions. Most of the research was supported by the National Scientific Program "Information and Communication Technologies for a Single Digital Market in Science, Education and Security" (ICTinSES).

Chapter One

Cybersecurity Tools for Threat Intelligence and Vulnerability Monitoring for National and Sectoral Analysis

George Sharkov, Yavor Papazov, Christina Todorova, Georgi Koykov and Georgi Zahariev

Abstract

Modern economies and digitised societies are increasingly dependent on the accessibility and availability of information and services within the cyberspace. Meanwhile, we suffer the consequences from malicious activities not only against online services, but against entire sectors as well, due to their interdependence and interconnectivity. This interdependence implies a need for targeted efforts to ensure the integrity and availability of individual systems and complex systems-of-systems (SoS) as well. The purpose of this research is to provide an aggregated picture of the technical profile and the vulnerabilities of a large number of systems, focusing first on the Bulgarian cyberspace and secondly, on a larger scale, on specific sectors or interconnected digital businesses (clusters), while providing the public and private sector with tools for early warning, threat intelligence, and vulnerability monitoring.

In this contribution, we provide a summary of the work and some of the findings from the pilot run of "Cyber Map Bulgaria", a fully-functional product providing and visualising data for the analysis of topics such as chronic vulnerabilities per domain or groups of domains, or identifying critical points within public and private IT infrastructure. We also present the "CyResLab Monitor", a subscription-based product integrating the functionality of “Cyber Map Bulgaria” with a customised alert system, on demand automated web testing, and fully customisable user interface.

Keywords: cyber threat, vulnerability analysis, cyber risk analysis, resilience, early warning, socio-technical approach, cyber picture, systems-of-systems (SoS), situational awareness

Background

Nowadays, businesses, industries, health care, governments, warfare, economies, education and life at the individual level are all dependent on the availability and integrity of the information systems within the cyberspace. Our interconnectivity online enables us to improve our lives substantially, while at the same time, poses significant security challenges (Piccoli, 2018). More often than not, we bear witness to the grave consequences resulting from malicious activities not only against online services, but against entire sectors as well, due to their interdependence with the affected infrastructures (APCERT, 2019).

The rise of cyberattacks (APCERT, 2019) and the interconnectivity between societal sectors and structures implies a need for focused efforts to ensure the availability and integrity of the individual systems (Ren, 2016) and to improve their capabilities to prevent, predict, detect, mitigate and remain resilient against cyber threats. For the purposes of detection and mitigation of vulnerabilities and exploits, both academia and practitioners have long been developing various Intrusion Detection and Prevention Systems (IDS/IPS) with automated response systems (Valdes, 2009). Traditional IDS/IPS are applicable to infrastructures developed by a single organisation. This implies a unified management of the infrastructure and processes within that organisation. More advanced systems aggregate data from various channels and sources (for instance, SIEM [Security Information/Incident Event Management] systems are popular).

Although there are many commercial products available on the market, allowing us to monitor individual web platforms, up to this moment, there are no publicly available or commercially distributed products and services that allow the monitoring, collection, processing, aggregating and visualising of information from the availability and technical profiles of large groups of both public and private online services. Within a national context, such aggregated information (a cyber picture) will serve as a key element for the monitoring and the visualisation of the state and the cybersecurity posture of the publicly available IT services in Bulgaria (or more precisely for the "national cyber picture" as envisaged within the cyber resilience model of the Bulgarian National Cybersecurity Strategy (National Cybersecurity Strategy, 2016). There are also no known analyses of the aggregated status of such services in different sectors and sub-sectors, economic dependencies (supply chains), and clusters in key groups of services provided by public and private organisations (e.g. in connection with the implementation of the Network and Information Security Directive, [EU Directive], 2016). Analysing, monitoring and reporting standard the behavioural patterns (availability) and technical metrics of the web platforms of target groups from both the public and the private sector will definitely help identify symptoms of anomalies and deviation from the standard behaviour and will serve for focused improvement and awareness programs, as well as possible warnings for mass cyber-attacks and larger scale crises.

In addition to providing comprehensive intelligence and visibility, we recognise the need for providing public and private organisations with intuitive tools and early warning mechanisms to enable them to take proactive steps to securing their online services. This research focuses on the creation of a two-faceted tool, serving two main purposes; namely 1) obtain, analyse and debrief cyber threat intelligence on a national and sectoral level, and 2) provide accessible and intuitive early warning mechanisms, based on the intelligence obtained (and tailored to the national needs and sectoral context), to enable companies to gain awareness and take proactive measures towards improving their overall cybersecurity posture. In doing so, we seek to improve the overall cyber picture of the country.

In the second section of this chapter (Cyber Threat Map and Cyber Monitor: Static and Dynamic Cyber Threat Pictures), we present the double-phased work on a software instrument developed precisely to achieve accurate and effective situational awareness regarding the Bulgarian cyberspace in general and within the granulation of industrial sectors or individual systems in particular. It provides a subscription-based instrument for individual organisations to gain awareness and improve their overall cybersecurity posture.

The third section (Research Method and Tools: Cyber Map Bulgaria) is dedicated to further discussing how to facilitate analysis, understanding, and reporting by applying various metrics, granulation filters, visualisation and user-customisable views to monitor the accessibility and availability of clusters of services. This section concerns the already successfully completed first phase of the system "Cyber Map Bulgaria". The fourth section (Research Method and Tools: The CyResLab Monitor) discusses the implementation of the second phase of this instrument, tailored to be used by individual organisations to gain intelligence on their online services' overall cybersecurity state. The fifth part of the chapter (Main Results and Findings) reports on the main findings and results of both phases of the development of the instrument. Finally, in the sixth section (Ongoing and Future Development) we discuss the ongoing work on the instrument, its planned future improvements, and identify some promising research directions. The chapter then ends with a summarising section.

Cyber Threat Map and Cyber Monitor: Static and Dynamic Cyber Threat Pictures

In November 2018 under the pilot project “Cyber Map Bulgaria”, funded by Sofia Tech Park JSC and SNIRD (see the Acknowledgements), we developed and implemented a method for the non-intrusive collection of technical, geographical, organisational and other data for the experimental database of more than 55,000 Bulgarian domains and domain groups.

Cyber Map Bulgaria is a fully-functional software system which provides and visualises data that could be further used for conducting multi-faceted analysis on topics such as the chronic vulnerabilities of the Bulgarian cyberspace per domain or the identification of critical points in public and private Bulgarian IT infrastructure.

Cyber Map Bulgaria, however, is just the first phase of our large-scale project, providing only a static picture of the Bulgarian cyberspace. Upon the successful completion of Cyber Map Bulgaria, we were awarded a new grant to initiate the second phase of our project, namely the CyResLab Monitor, a project again funded by Sofia Tech Park JSC and SNIRD (see the Acknowledgements).

The final goal of the CyResLab Monitor is to provide a dynamic cyber threat picture to organisations from the public and the private sector upon subscription and to deliver early warnings to analysts. We would also seek to enable decision making for preventive measures in order to face incoming attacks at a sectoral level and on the level of digital clusters.

The CyResLab Monitor integrates the functionality of Cyber Map Bulgaria, with a monitoring service, tailored to the national and sectoral specifics of the Bulgarian cyberspace, providing a functional module and a common platform for monitoring and early warning against potential mass cyber-attacks on economic and public organisations. The CyResLab Monitor realises new tools, means, and methods for the dynamic monitoring and analysis of the behaviour (availability) of the web systems of specific target groups, by adding mechanisms for monitoring, and for historical retrospectives on as well as identification of symptomatic behaviour models, which will allow for the early warning for mass cyber-attacks and crisis threats.

Currently, the CyResLab Monitor allows users to personalise alerts and notifications via e-mail or SMS for certain preliminary defined events such as service failures or denial of service, and to dynamically monitor various technical metrics related to its availability in multiple geographic regions. The pilot testing of the CyResLab Monitor and its results will be at the core of further research on the design of early warning mechanisms and the overall improvement of cybersecurity and the digital economy in Bulgaria, as well as the resilience of services relevant to the economy and the public sector.

Within the following sections of this paper, we present the research method and tools applied for the implementation of both Cyber Map Bulgaria and the CyResLab Monitor, as well as some key findings from it.

Research Method and Tools: Cyber Map Bulgaria

The project employed both theoretical and empirical methods to approach the creation of a holistic instrument to serve for both the collection of data from web-based services and the visualisation and aggregation of the collected information.

Architecture and realisation. Cyber Map Bulgaria has 3 components: a database management system, a collector component, and web services, which we generally divide in two sub-components, namely API and a web platform, as visualised in Fig. 1-1.

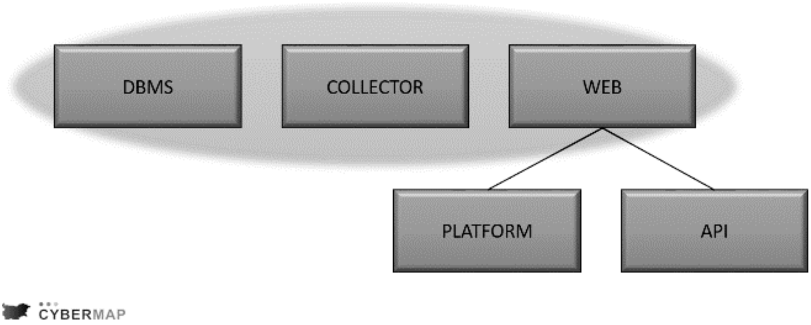

Fig. 1-1. General Architecture of Cyber Map Bulgaria

The DBMS component is a non-relational database as it allows for the collection of data without a previously established structure. By the same token, a non-relational DBMS allows us to work with incomplete data as the data collected is harvested only through non-invasive methods. We then use only the data which the monitored services publicly provide. The data is not always complete and does not always follow the recommendation for the structuring of this sort of information (IETF Tools, 2018).

The Collector component serves to collect, process and store the data from the analysed websites and allows us to add a new website or group of websites (a given industry) for monitoring, and to update the data.

The web component consists of a web server (platform) which provides the entire GUI, including the ability to personalise the visualisation of data on different maps, graphs, etc.

The other subcomponent, the API, serves for the provision of the data in order to enable visualisation and filtration, as well as authentication and access control within Cyber Map Bulgaria. Within this component are the security controls of the system, which support user roles to provide the opportunity for differentiated levels of access, based on data sensitivity.

Note that this is only the partial architecture for Cyber Map Bulgaria, the first phase of the bigger system. The final integrated product architecture is available in Fig. 1-2.

Model of Collected Data. The data model has the goal of clearly and effectively describing the observed unit—i.e. a given website. To briefly overview the functionality of the model, the required input of the system is a Fully Qualified Domain Name (FQDN).

An initial requirement for the data model was to allow the collection of historical information and its enrichment with time. Another was that the system was to be designed so as to handle incomplete data obtained from publicly available sources. Following those requirements (and a careful analysis), the research team created a data model that provides a standardised representation which is conveniently stored and maintained and which is easy to be processed, visualised and analysed.

A simplified representation of the data model would include the following:

- FQDN
- Historical data, collecting objects under the following simplified structure:
- Timestamps for the raw data collection
- Network features:
  - IP
  - ASN
  - DNS

  - GeoIP
- TLS / HTTPS information
  - availability
  - detailed certificate information
- Data regarding software features
  - OS information
  - Web technologies information
  - returned "raw" data from the server
  - algorithmically extracted information based on the "raw" data
- CPE
- Others

The modular design of the system allows for an important additional feature: the distinction between collected raw data and parsed and structured information. This distinction creates the opportunity to retroactively apply improvements in data parsing algorithms to older raw data, improving the quality of the collected information.

Reports and Visualisations. The methods used for the visualisation of the information are graphs, tables and maps. All of those methods could be applied to different domain groups, filtered based on technical, geographical, and other metrics, and then further customised. The visualisation and recording of physical location are realised using a geographic information system (GIS).

The choice to deploy a GIS for this project was motivated by the research conducted on the stability of cyber-physical systems security. The physical clusterisation of the hosted web platforms is of key importance for the availability of those services in case of hybrid attacks or natural disasters, for instance as related by Esri (2015).

The purpose of integrating a GIS in this software solution is the opportunity to analyse and rapidly identify location-based dependencies as well as to provide geospatial recommendations and solutions for risk reduction and flexibility of the web-based systems. The geospatial analysis of the maintenance of the web-based services could be further related to the ability to assess the potential for creating a sustainable cyber-geographical model for perimeter protection, as well as for an impact assessment of potential vulnerabilities.

Research Method and Tools: The CyResLab Monitor

The CyResLab Monitor aims to provide a flexible, robust and scalable monitoring platform. Particular attention is paid to the processes of metric collection, processing, storage and querying.

Architecture and realisation. To provide a scalable base for the Monitor's implementation, our team chose the Kubernetes platform (more commonly transcribed as "k8s"), which builds upon existing container infrastructure, but provides a different, newer set of abstractions, designed to empower modern development / IT operations.

This design choice has some important advantages: Several design concerns were addressed entirely within the Kubernetes platform, in particular scalability (a core feature for k8s), loose coupling (another staple of the microservice / container-oriented approach), as well as security and availability to a certain degree (k8s has an in-built enterprise access control toolkit, as well as simple monitoring tools).

To achieve the other main aim of its architectural design, namely flexibility, the team utilised an innovative computational paradigm: Function-as-a-Service (FaaS). With FaaS technologies, the user can provide custom source code to be executed in the designated runtime environment, usually in the form of a container, running dynamically mounted code. This code must run in a timely manner (standard limitations are between 5 and 15 minutes) and should have similar granularity to that of a function (ergo the name of the technology). Due to the utilisation of the FaaS paradigm, the CyResLab Monitor can perform custom availability checks for different types of infrastructure, such as various black-box, grey-box, and white-box availability checks/metrics.

Some notable drawbacks of the current architectural approach are:

- **Cost**: the current platform has a high "at rest" runtime cost. That is, under a 0% external load, the platform still costs a significant sum to run (on a public cloud provider).
- **Untested**: due to the innovative architecture, the informal knowledge base of the technologies involved, as well as the proper way to integrate them, are still under active development.

Implementation. The team implemented the above architecture in a (mostly) minimal configuration, foregoing some operational concerns, such as High Availability (HA), backup and multi-region availability. These can be trivially (although costly) addressed at the Kubernetes implementation level.

The platform also necessitated the development of a (somewhat) standard three-tier web application that serves as the UI for the project. This web application was developed in Python (due to team preference) and deployed as an FaaS application on the Kubernetes cluster, allowing for the high utilisation of the available resources, thus decreasing the 'at rest' cost.

The platform builds on top of the basic monitoring (micro)service software Prometheus. It is the native monitoring solution for Kubernetes, which provides further architectural and implementational cohesion, due to the number of pre-developed plugins/extensions that can be used to attach Prometheus to many standard infrastructure setups and products.

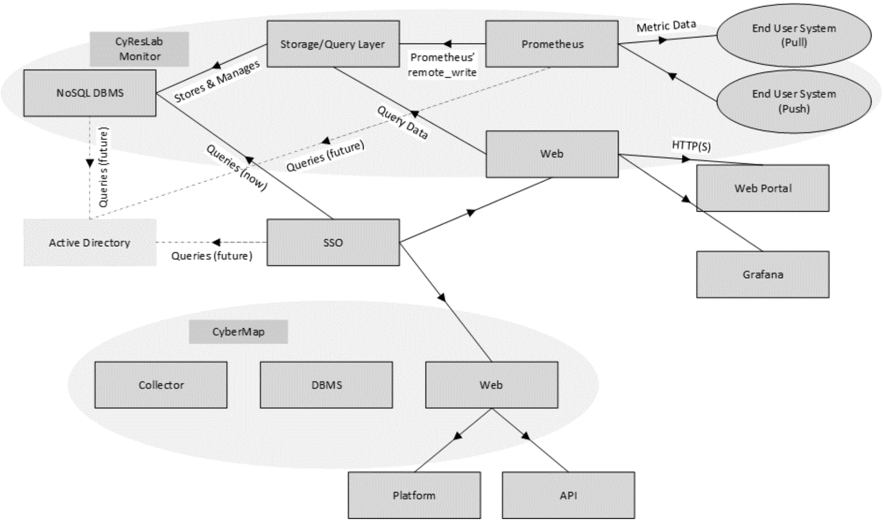

Fig. 1-2. The general architecture of the CyResLab Monitor and the integration of Cyber Map Bulgaria within it.

The Cyber Map Bulgaria component and its sub-components remain as per Fig. 1-1.

In terms of other integrations, the team developed several new tests; these can be attached to Prometheus:

- Selenium

- API tests
- WordPress

These integrations allow the product to test a wider variety of systems, and, therefore, increase the delivered value to the end user.

Main Results and Findings

The project employed both theoretical and empirical methods to approach the creation of a holistic instrument to serve for both the collection of data from web-based services and the visualisation and aggregation of the collected information.

Upon the onset of the project, the working hypothesis of the research was that such an aggregated picture would uncover key and critical points in the Bulgarian IT infrastructure. To be able to test this hypothesis in practice, the team of the Cybersecurity Laboratory created an experimental database with more than 55,000 Bulgarian domains. We further entered metrics and reference values for several domain groups including, but not limited to, hospitals and dispensaries, the IT sector, banks, insurance companies, small and large administrative institutions, pharmaceutical companies, schools, kindergartens and a few others.

Within Cyber Map Bulgaria, passive tests were carried out on different domains and domain groups and several mechanisms allowing information filtration and sorting based on different technical metrics were created.

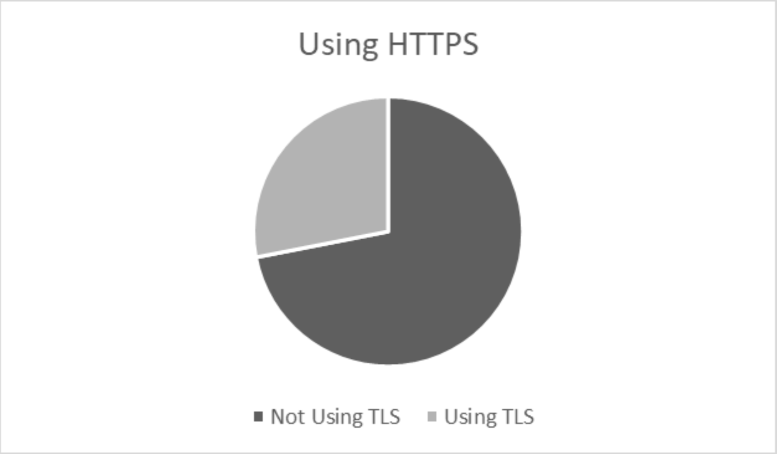

Fig. 1-3. Web sites using HTTPS (TLS), Bulgaria, July 2019

"Cyber Map Bulgaria" provides **a static visualisation** of the availability, for instance, of a secure connection, as seen in Fig. 1-3. The ratio between secure and insecure websites could be clearly visualised and further filtered through other metrics.

As visualised in Fig. 1-3, we can see that the overall distribution of Bulgarian web platforms, as of July 30th, 2019, that do not provide a secure connection from our database is quite worrying. Within the sectors that mostly use a secure connection (HTTPS) are the IT companies and private companies in general, along with banks and pharmaceutical companies.

A revealing discovery was that a lot of the smaller administrative web-platforms, such as those of municipalities or smaller administrative agencies are still not using HTTPS. Along with them, web-platforms of schools, kindergartens and some hospitals, especially those of services that are physically located in smaller cities in Bulgaria, are not using HTTPS as well.

The versions of a given component (Apache, Nginx, PHP, etc.) could also be visualised to showcase a ratio between up-to-date product versions and outdated ones. Older versions of components are likely to have known vulnerabilities that could be easily exploited even by people without extensive experience and knowledge in cybersecurity.

On a higher level, Cyber Map Bulgaria provides opportunities to visualise clusterisations of domains in different geographical regions. This is of importance to analyse vulnerabilities and risks on a geographical level. As expected, most of the web-services in Bulgaria are physically hosted within several locations in Sofia (the capital and major commercial city).

The conclusions from the data and its filtered visualisation provide opportunities for more in-depth research in the future to support a better-looking overall picture, an improved competitiveness and awareness, and the undertaking of adequate corrective measures.

For the purposes of a dedicated report by the Institute of Public Administration (IPA), "Cybersecurity and opportunities for the application of innovative technologies in the work of the state administration", a special sample of public administration web-platforms (central agencies, ministries and local administration, such as municipalities) was selected and scanned, on the basis of which a detailed analysis was carried out concerning the common weaknesses and vulnerabilities of their websites and Internet-accessible systems. Using Cyber Map Bulgaria for the collection and visualisation of data, special recommendations were defined in the report (Polimirova et al., 2018).

Within the section of the report on "Identifying Critical Points and Prospects: Presenting a Clear Vision for the Improvement the Cybersecurity

Posture of the Public Sector in Bulgaria", a proof-of-concept sectoral analysis was performed using Cyber Map Bulgaria to compare three "sectors" of the Bulgarian cyberspace:

- A large group of state administration websites (abbreviated to "muni" from "municipalities"): an aggregated group of websites of central and local administration (municipalities)
- A smaller group of state administration websites: only those based centrally
- A group of "business organisations": a representative sample of industry organisations (a similar number to the "muni" group)

The results of some very basic and accumulated metrics (performed in July 2018 and November 2019, respectively) for these three sectors can be seen below in Figs. 1-4 and 1-5:

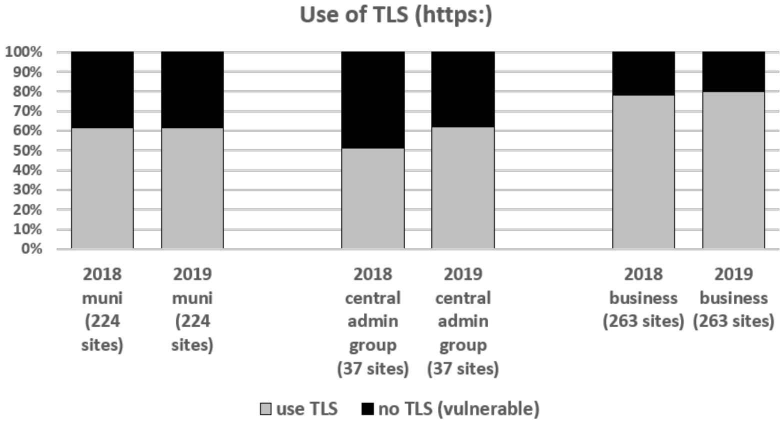

Fig. 1-4. Comparison of the distribution of TLS usage among local administration ("muni" from "municipalities"), central public administration and business organisations in 2018 and 2019.

Here in Fig. 1-4 it is evident that the web-platforms of the central and local administration are mostly still not using even TLS, to ensure a secure connection to their websites. With the group of the web-platforms of the central state administration, the situation is arguably not much better. From all analysed web-platforms of the central state administration, about half do not ensure a secure connection to their services. Within the group of private business organisations, we could see a better-looking overall picture, showing that the majority of the analysed web-platforms are using TLS. A comparison of the aggregated results of the identical groups between 2018 (July) and 2019 (November) also shows a slight improvement both in the central administration group and the business organisations' websites.

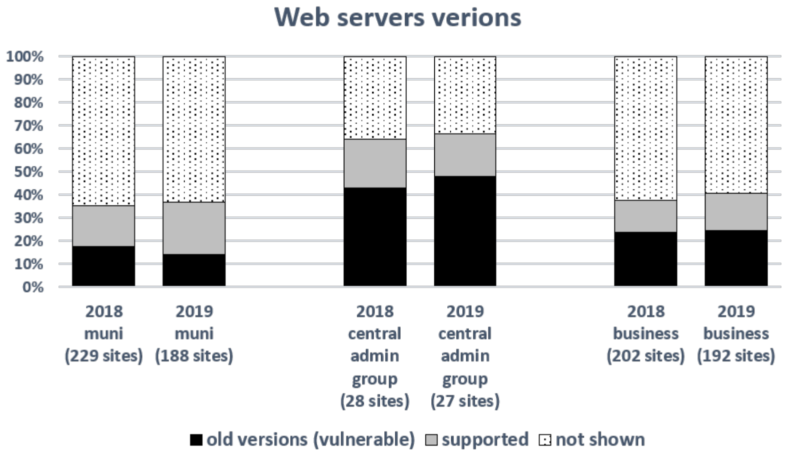

Fig. 1-5. Comparison of the distribution of old (not supported), new, and not-displayed server versions supported by web platforms of local administration ("muni"), central public administration and business organisations in 2018 and 2019.

A similarly troublesome picture can be observed with relation to the distribution of server versions. Although a small percentage of the local administration (municipalities or related) web-platforms are running on clearly outdated web servers with known vulnerabilities, there might be a significant hidden number of vulnerable platforms in the not-displayed versions. On the other hand, the increasing number of "supported" versions in 2019 compared to 2018 is a positive sign. The situation with the central state administration looks worse with over 40% of the platforms running on outdated web server versions and a signal for an increased level of threat in 2019. With the business organisations, the number of platforms known to be running on outdated servers is less than 20%, however we can clearly see a tendency for the web server version information to be hidden. Contrary to expectation, there is no visible improvement between 2018 and 2019, and deeper analysis of a larger group within this category would provide correlation with organisation size (with expected concerns about SMEs), be they sectoral or territorial dependencies.

These figures are just a portion of the results from the analysis run on those groups, the rest are available in the report by the IPA (Polimirova et al., 2019).

Such information as that presented above is very useful for receiving a static picture of the overall state and well-being of the Bulgarian cyberspace, as it enables experts to come up with sectoral recommendations, actions and supporting programs. Periodical scanning (at least yearly) and tracking various groups (clustered by organisation type or size, sector, geography or other parameters) would definitely allow warnings and predictions about potential cyber threats and known vulnerabilities to be exploited and would serve as a baseline for further dynamic monitoring of the state by the CyResLab Monitor. These results would also serve as quantitative indicators of regional and national cybersecurity (web) improvement measures and programs.

Ongoing and Future Development

The project employed both theoretical and empirical methods to approach the creation of a holistic instrument to serve both for the collection of data from web-based services and for the visualisation and aggregation of the collected information.

Cyber Map Bulgaria is only the first phase of the ongoing bigger project, the CyResLab Monitor. The CyResLab Monitor is the intuitive future development of Cyber Map Bulgaria which aims to integrate the functionality of Cyber Map Bulgaria (the “static picture”) within a larger platform, containing tools for the dynamic monitoring of standard behavioural patterns of the monitor web-services and the availability of their web-based systems. The CyResLab Monitor currently implements mechanisms for the historical analysis of this standard behaviour to come up with definitions for deviation. Furthermore, it provides a continuous and non-destructive monitoring and collection of technical data concerning the availability of the web systems and implements tools for personalised notifications when access to key services is denied or if changes in the standard service availability line occur.

The future development of the CyResLab Monitor envisages the adaptation and personalisation of the instrument to the context of specific services, sectors and economic segments. This cybersecurity situational awareness is expected to raise competitiveness, contribute to resilience against cyber threats, and improve the overall cybersecurity of the end-client (Caralli, 2011; CERT, 2016). It could also further allow for tailored sector or region oriented cybersecurity improvement programmes and initiatives (National Cybersecurity Strategy, 2016).

In recent years, an increase in the scale and intensity, as well as the hybrid and complex nature of cyber threats, attacks, and actors (addressed publicly as Advanced Persistent Threats, APTs) has been observed. Adversaries are using different methods and means of Artificial Intelligence (and accordingly machine learning techniques, such as AI / ML). This makes attacks much more powerful and difficult to detect, such as malicious code or operations directly (with a prolonged "stealth" period). Therefore, systems behaviour monitoring and deviation alarm techniques are increasingly required as a basic approach (Hall, 2019) and thus with the CyResLab Monitor we aim to trace the overall underlying service and dependency layers, not just the technical metrics of availability, so as to predict a potential “domino effect” along the 15-layered cyber terrain (Riley, 2014) and real complex cyber-physical systems (including complex systems-of-systems).

Nevertheless, the precondition for useful dynamic behaviour monitoring is a sufficient level of massive and sector-specific cyber hygiene. The Cyber Map is proven and helpful tool for this initial maturity assessment at sectoral or cluster level. The recently observed (July 2019) large personal data leak from the National Revenue Agency (nearly 100% of all tax-payers details and sensitive income and tax related info with historical data, including also for death people) clearly demonstrates the importance of sector-specific measures to massively meet the minimum cybersecurity requirements (as required by the NIS Directive and GDPR, as well as the National Cybersecurity Strategy and the Bulgarian Cybersecurity Act, adopted in November 2018 and the follow-up ordinances in effect).

As it was demonstrated by the results of Cyber Map pilot studies described herein, a substantial part of public administration sites and essential services providers are suffering from the basic vulnerabilities (like lack of proper TLS, old unsupported versions of servers and software with number of known exploits, etc.).

Conclusion

This paper examined the cybersecurity problem of early warning, prognosis, vulnerability analysis and threat prevention in Bulgaria and informed on the development and functionalities of an instrument providing an aggregated picture of the technical profile, standard behaviour and vulnerabilities of a large number of independent systems that maintain services within the Bulgarian cyberspace.

The motivation behind the development of this instrument and this paper is the need for improving the limited capability for applying cybersecurity controls, analysis and preventive measures on a national scale.

The development of the Cyber Map Bulgaria and the ongoing larger project, CyResLab Monitor, embeds multiple research methods and state-of-the-art technical platforms to develop a series of interconnected instruments that work together to provide a working product, that aims at allowing an end client to use tools, means and methods for the dynamic monitoring and analysis of the behaviour (availability) of the web systems of specific target groups. This is achieved by adding mechanisms for monitoring, and for the historical retrospective and identification of symptomatic behaviour models, which will permit an early warning to mass cyber-attacks and large-scale crisis threats. Its native quick scalability and customisation makes the integrated CyResLab Monitor platform usable at a higher level, for national or sectoral cybersecurity picture monitoring and resilience, as well as for any complex system-of-systems, including internal enterprise and industrial systems, not necessarily Internet or web based. In addition, the customisable and scalable platform allows a custom services testing approach in such complex systems-of-systems as global supply chains, where only probing and checking the availability of entire composite services could be an indicator of hidden dependencies, subject to the increasing interest of modern APTs (Schauer, 2019).

Finally, the modern technological platform and tools of the CyResLab Monitor (and Cyber Map Bulgaria) have been successfully piloted for use in practical research and educational platforms: i.e. for the purposes of monitoring simulation environments to support research and cyber exercises infrastructure, following a dedicated research management architecture methodology (Tagarev, 2017). The platform will benefit training and education activities at universities and professionals following the recommendations of skills and competences in the area of cybersecurity (Orozova, 2019).

Our belief is that with such instruments and services we will be able to improve the efficacy of predicting, preventing and handling cybersecurity incidents and improve the overall cybersecurity of Bulgaria or any other cyber ecosystem.

Acknowledgements

The pilot version of "Cyber Map Bulgaria" was a project financed under agreement № D-060-2018 / 20.06.2018 by Sofia Tech Park JSC and SNIRD. The activities were carried out within the period July-November 2018 by the Cybersecurity Lab of STP with the support of the European Software Institute-Center Eastern Europe (ESI CEE) and the CyResLab of ESI CEE. The "CyResLab Monitor" is the second-phase of the ongoing project, funded by Sofia Tech Park JSC and SNIRD under agreement № D-094-2019 / 27.09.2019.

References

APCERT. (2019). *Asia Pacific Computer Emergency Response Team (APCERT) Report 2019.* White Paper, APCERT. http://www.apcert.org/documents/pdf/APCERT_Annual_Report_2019.pdf.

Caralli, Richard, Julia Allen, and David White. (2011). C*ERT Resilience Management Model (CERT-RMM).* SEI Series in Software Engineering, Addison-Wesley Professional.

CERT. (2016). *Resilience Management Model (Software Engineering Institute, Carnegie Mellon University).* http://www.cert.org/resilience/products-services/cert-rmm/ .

Directive (EU) 2016/1148. (2016). "*EU NIS Directive" of the of concerning measures for a high common level of security of network and information systems across the Union.* Directive, European Parliament.

Esri. (2015). *The Geospatial Approach to Cybersecurity: Implementing a Platform to Secure Cyber Infrastructure and Operations.* Esri® White Paper.

Hall, Patrick. (2019). *Proposals for model vulnerability and security.* O'Reilly. Accessed July 2019. https://www.oreilly.com/ideas/proposals-for-model-vulnerability-and-security.

IETF Tools. (2018). October 25. Accessed July 31, 2019. https://tools.ietf.org/html/rfc7231#section-7.4.2.

National Cybersecurity Strategy. (2016). *National Cybersecurity Strategy "Cyber Resilient Bulgaria 2020".* Sofia, Bulgaria. http://www.cyberBG.eu.

Orozova, Daniela, Kalinka Kaloyanova and Magdalina Todorova. (2019). "Introducing Information Security Concepts and Standards in Higher Education." *TEM Journal* 8(3): 1017-1024.

Piccoli, Gabriele, and Federico Pigni. (2018). *Information Systems for Managers with Cases (4.0)*. Prospect Press.

Polimirova, Dimitrina, Velizar Shalamanov, Nikolay Stoianov, Todor Tagarev, Yantsislav Yanakiev, George Sharkov, Yavor Papazov, Vasil Rizov, and Krasimira Ivanova. (2018). *Cybersecurity and opportunities for applying innovative technologies in Bulgarian public services*. Sofia: Institute for Public Administration. Available in Bulgarian language under the title “Киберсигурност и възможности за приложение на иновативни технологии в работата на държавната администрация в България”

Ren, Shu Qin, Benjamin Hong Meng Tan, Sivaraman Sundaram, Taining Wang, Yibin Ng, Victor Chang, and Khin Mi Mi Aung. (2016). “Secure searching on cloud storage enhanced by homomorphic indexing.” *Future Generation Computer Systems* 65: 102–110.

Riley, Shawn. (2014). *Cyber Terrain: A Model for Increased Understanding of Cyber Activity.* Conference. Accessed March 2020. http://cyber-analysis.blogspot.com/2014/10/cyber-terrain-model-for-increased.html.

Schauer, Stefan, Despina Polemi and Haralambous Mouratidis. (2019). “MITIGATE: a dynamic supply chain cyber risk assessment methodology.*” Journal* of Transportation Security 12: 1-35.

Tagarev, Todor, George Sharkov, and Nikolay Stoianov. (2017). *Cybersecurity and Resilience of Modern Societies: A Research Management Architecture.* Information & Security 38: 93-108. doi:10.11610/isij.3807.

Valdes, Alfonso, and Steven Cheung. (2009). *Intrusion Monitoring in Process Control Systems.* 42nd Hawaii International Conference on System Sciences. Big Island, HI. 1-7. doi:10.1109/HICSS.2009.273.

Chapter Two

CVE Annotation

Vladimir Dimitrov

Annotation

The leading idea in knowledge extraction from text systems is that knowledge exists in the text. It is not so simple—the text must be understood before the knowledge can be extracted. More external knowledge must be used in order to understand the text. In this chapter, the texts of the CVE (common vulnerabilities and exposures) are annotated. These texts are semi-structured. The aim is to generate a knowledge base from the annotated texts.

Keywords: text annotation, knowledge base, ontology, computer security, vulnerabilities, CVE.

Introduction

The MITRE Corporation (MITRE Corporation, CWE, 2020) defines weaknesses and vulnerabilities as follows:

- "Weakness: a type of mistake in software that, in proper conditions, could contribute to the introduction of vulnerabilities within that software. This term applies to mistakes regardless of whether they occur in implementation, design, or other phases of the SDLC."
- "Vulnerability: an occurrence of a weakness (or multiple weaknesses) within software, in which the weakness can be used by a party to cause the software to modify or access unintended data, interrupt proper execution, or perform incorrect actions that were not specifically granted to the party who uses the weakness."

The MITRE Corporation lists more elaborate definitions of vulnerabilities in their publication on CVEs (MITRE Corporation, CVE, 2020):

- "A 'vulnerability' is a weakness in the computational logic (e.g., code) found in software and some hardware components (e.g., firmware) that, when exploited, results in a negative impact to confidentiality, integrity, OR availability. Mitigation of the vulnerabilities in this context typically involves coding changes, but could also include specification changes or even specification deprecations (e.g., removal of affected protocols or functionality in their entirety)."
- "An 'exposure' is a system configuration issue or a mistake in software that allows access to information or capabilities that can be used by a hacker as a stepping-stone into a system or network.

  CVE considers a configuration issue or a mistake an exposure if it does not directly allow compromise but could be an important component of a successful attack, and is a violation of a reasonable security policy.

  An 'exposure' describes a state in a computing system (or set of systems) that is not a vulnerability, but either:
  - allows an attacker to conduct information gathering activities;
  - allows an attacker to hide activities;
  - includes a capability that behaves as expected, but can be easily compromised;
  - is a primary point of entry that an attacker may attempt to use to gain access to the system or data; or
  - is considered a problem according to some reasonable security policy"

The MITRE Corporation maintains a public database for weaknesses, namely CWE (MITRE Corporation, CWE, 2020) and a public database for vulnerabilities, known as CVE (MITRE Corporation, CVE, 2020).

The focus of this chapter is on the vulnerabilities, i.e. CVEs. Here, the weaknesses (CWEs) are vulnerability types.

CWEs

The CWE catalogue contains several views intended for different auditoria like developers, researchers, architects, etc. Each view can be organised into categories—sets of weaknesses with common characteristics. Categories are conceptual elements structuring weaknesses. Some views are not structured, but the main CWE views for researchers, for developers, and for architects are structured by concepts (*categories*). Each category can contain other categories (*subcategories*).

The classes, bases and variants are simple weaknesses. They are the abstraction levels of CWE. The *class* is an abstract weakness that is not associated with any platform or technology. *Bases* are more specific than classes. Usually, the base is not associated with any platform or technology but contains enough details to be detected. The *variant* is more specific than the base and is usually associated with a specific platform or technology.

The classes, bases and variants are not organised in inheritance hierarchies. They are organised in generalisation levels. A class can be more general in the CWE structure than other classes, bases and variants. A base can be more general than other bases and variants. A variant can be more general than other variants. Each *simple weakness* can be more general than another weakness that has at least the same level of abstraction as its own.

Apart from simple weaknesses, there are also *compound weaknesses* (*composites and chains*) that combine several other weaknesses. The chain weaknesses are ordered, and composite weaknesses are simply sets.

The weaknesses (simple and compound), naturally participate in more than one view, but it is possible for a weakness to participate within a view more than once. The last case is a taxonomy problem.

Weaknesses are organised by structure and abstraction levels, but there are many relations among them than are not included in this paper.

Each CWE entry contains the following information:

- CWE ID and name;
- description;
- alternate terms;
- description of the behaviour;
- description of the exploit;
- likelihood of the exploit;
- description of the consequences of the exploit;
- potential mitigations;
- node relationships;
- source taxonomies;
- code samples for the languages/architectures;
- CVEs (vulnerabilities) for which that type of weakness exists; and
- references.

Weaknesses are vulnerability types. Each CWE references CVEs of its type. Vulnerabilities are classified by their types (weaknesses).

The process of vulnerability investigation starts with vulnerability registration. Usually, the vulnerability type is not clear. After some investigation a type (or types) is assigned to this new vulnerability. If there are no suitable CWEs, a new one is created.

During the investigation process of CWEs, information about conducted attack types can be identified. Attack types are classified by the MITRE Corporation as templates in CAPEC—Common Attack Pattern Enumeration and Classification (MITRE Corporation, CAPEC, 2020). If a new attack pattern is discovered, then a new CAPEC entry is created.

Sometimes, it is impossible to identify the vulnerability type or its attack pattern. In these cases, corresponding references are not created in the CWE.

CWEs are the cornerstone for cybersecurity activities. They contain information for a vulnerability and possibly, how to identify, protect, detect, respond, and recur from it.

CVEs

The vulnerabilities (CVEs) database is very simple. Every CVE entry has a name, description, references to external sources, and some maintenance information.

The NVD (National Vulnerability Database) is based on its CVE counterpart. CVE entry in the NVD contains some additional metrics.

A CVE description must follow the next two patterns:

- [VULNTYPE] in [COMPONENT] in [VENDOR] [PRODUCT] [VERSION] allows [ATTACKER] to [IMPACT] via [VECTOR].
- [COMPONENT] in [VENDOR] [PRODUCT] [VERSION] [ROOT CAUSE], which allows [ATTACKER] to [IMPACT] via [VECTOR].

The product ([PRODUCT]) can be identified in the following combinations:

- [PRODUCT_NAME]
- [VENDOR_NAME] [PRODUCT_NAME]
- with keywords (the product has no name)
- the product name is written as the vendor names it
- [PRODUCT_NAME] (aka [ALT_NAME])
- [PRODUCT_NAME] ([ACRONYM])
- [PRODUCT_NAME (formerly [OLD_NAME])
- [PRODUCT_NAME] and [OTHER_PRODUCT_NAME]
- [PRODUCT_NAME], as used in [BUNDLING_PRODUCT]
- [PRODUCT_NAME] [COMPONENT_TYPE] for [PLATFORM]

The version ([VERSION]) can be represented in several variants:

- The version 1.2.3
- The versions 1.2.3, 2.3.1, and 3.1.2