UNIVERSITI PUTRA MALAYSIA

PRIVACY OPTIMIZATION AND INTRUSION DETECTION IN MODBUS/TCP NETWORK-BASED SCADA IN WATER DISTRIBUTION SYSTEMS

DANIEL JOSÉ DA GRAÇA PECEGUINA FRANCO

FSKTM 2022 14

PRIVACY OPTIMIZATION AND INTRUSION DETECTION IN MODBUS/TCP NETWORK-BASED SCADA IN WATER DISTRIBUTION SYSTEMS

By

DANIEL JOSÉ DA GRAÇA PECEGUINA FRANCO

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

September 2021

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

PRIVACY OPTIMIZATION AND INTRUSION DETECTION IN MODBUS/TCP NETWORK-BASED SCADA IN WATER DISTRIBUTION SYSTEMS

By

DANIEL JOSÉ DA GRAÇA PECEGUINA FRANCO

September 2021

Chair : Abdullah Muhammed, PhD
Faculty : Computer Science and Information Technology

Water Distribution Systems (WDS) are now controlled and monitored by computational systems, constituting the primary management challenge from both an operational and public health standpoint. Modbus/TCP networks in WDS were initially developed to work based on a high availability and under closed-networks, where security was not an issue and communications were performed in clear-text. The need of interoperability and financial reduction, triggered the evolution to opened-standard TCP/IP networks, where clear-text communications are no longer safe and are putting the systems into a highly-vulnerable level.

One of the key essential elements is the privacy of data sets; they can be turned publicly available and has potential to be use for the development of security solutions. Therefore, the first problem to be tackled is the privacy optimization of Modbus/TCP packet fields. In scientific literatures, packet anonymization is performed according to attribute types (numerical, categorical and hierarchical), not taking into consideration the singular characteristics of the Modbus packet fields, using Euclidean distance algorithms that are not capable to deal with binary data and may result in information loss. Another problematic aspect is related to the intrusion detection solutions that are based on machine learning cluster algorithms to learn systems' specifications and extract general state-based rules for attacks identification. Such approach is highly dependable on the clustering algorithm parameterization, and is not capable to deal with the normal system's specification changes. Different parameterizations achieve different results ending in high false positive alarms or miss-identification of real intrusions. Based on these problems, this research objectives are firstly to propose SCADA Modbus/TCP packet fields' privacy optimization using anonymization algorithms, increasing the privacy level and reducing information loss, and, secondly, to propose a State-Based IDS for attacks identification, dedicated to SCADA Modbus/TCP in WDS, capable of

extracting specific rules and deal with the constant system specification changes, while reducing false positive rates and increasing accuracy.

Experimental design and simulations are carried out through a quantitative approach, where the proposed solutions perform the anonymization of Modbus/TCP packet fields to achieve acceptable privacy levels for data sets publication and proposes a state-based IDS tailored to Modbus/TCP networks in WDS, taking advantage of a knowledge database and state-based rules' language to control on systems states and constant specification changes.

Experimental results show that our proposed privacy algorithm is able to work effectively in terms of privacy level (12.01 against 10.48), efficiency (2.74ms against 3.84ms) and scalability (470.15ms against 507.48ms), when dealing with multivariate traffic attributes. In relation to information loss, the proposed solution was able to achieve an average of 12.2% against 18.6% of the benchmark solution. Moreover, state-based IDS experimental results show a higher effectiveness in terms of true (99.50% against 95.75%), false positive rates (1.20% against 1.85%) and accuracy (98.70% against 93.68%), on the identification of attacks and intrusions. Over all, this research proposes a set of solutions to address privacy and security issues related to Modbus/TCP networks in WDS. Research work presented in this thesis is a significant step towards a safer SCADA WDS and public health.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PENGOPTIMUMAN KERAHSIAAN DAN PENGESANAN PENCEROBOHAN DALAM SCADA BERASASKAN RANGKAIAN MODBUS/TCP DALAM SISTEM PENGAGIHAN AIR

Oleh

DANIEL JOSÉ DA GRAÇA PECEGUINA FRANCO

September 2021

Pengerusi : Abdullah Muhammed, PhD
Fakulti : Sains Komputer dan Teknologi Maklumat

Sistem Pengagihan Air (SPA) yang kini dikawal dan dipantau oleh sistem komputasi merupakan suatu cabaran utama kepada pihak pengurusan dalam memastikan tahap operasi dan kesihatan awam berada pada tahap yang terbaik. Rangkaian Modbus / TCP dalam SPA pada awalnya dibangunkan untuk kegunaan rangkaian tertutup dan kebolehsediaan tinggi, di mana keselamatan bukanlah menjadi isu dan komunikasi dapat dilakukan dalam keadaan selamat. Bagaimanapun, keperluan saling kendali dan pengurangan kos kewangan telah memacu evolusi dari rangkaian tertutup ke rangkaian TCP / IP piawaian terbuka di mana komunikasi teks jelas tidak lagi selamat dan meletakkan sistem ke tahap yang sangat rentan.

Salah satu elemen penting keselamatan adalah kerahsiaan set data; ianya boleh didedah kepada umum dan berpotensi digunakan untuk pembangunan penyelesaian keselamatan. Oleh itu, masalah pertama yang perlu diselesaikan adalah pengoptimuman kerahsiaan bidang paket Modbus/TCP. Dalam kajian saintifik, pendekatan anonimisasi paket dilakukan mengikut jenis atribut (berangka, berkategori dan berhierarki), tidak mengambil kira ciri tunggal medan paket Modbus, menggunakan algoritma jarak Euclidean yang tidak mampu menangani data penduaan dan boleh mengakibatkan kehilangan maklumat. Aspek permasalahan lain adalah berkaitan dengan penyelesaian kepada pengesanan pencerobohan yang berasaskan algoritma gugusan pembelajaran mesin untuk mempelajari spesifikasi sistem dan mengekstrak peraturan berasaskan keadaan umum untuk mengenal pasti serangan. Pendekatan sedemikian sangat bergantung kepada pemparameteran algoritma penggugusan dan ianya tidak mampu menangani perubahan spesifikasi sistem biasa. Pemparameteran yang berlainan boleh mengakibatkan hasil yang berlainan dan ianya boleh berakhir dengan penghasilan penggera positif palsu

yang tinggi atau ketidakupayaan untuk mengenalpasti pencerobohan sebenar. Berdasarkan permasalahan ini, objektif penyelidikan pertama adalah untuk mencadangkan algoritma anonimisasi untuk pengoptimuman kerahsiaan bidang paket SCADA Modbus/TCP bagi meningkatkan tahap kerahsiaan dan mengurangkan kehilangan maklumat; dan keduanya mencadangkan IDS berasaskan keadaan bagi mengenalpasti serangan khusus untuk SCADA Modbus/TCP dalam SPA yang mampu mengekstrak peraturan khusus dan menangani perubahan spesifikasi sistem yang berterusan, disamping mengurangkan kadar positif palsu dan meningkatkan ketepatan.

Reka bentuk eksperimen dan simulasi dilaksanakan melalui pendekatan kuantitatif di mana penyelesaian yang dicadangkan melakukan anonimisasi ke atas bidang paket Modbus/TCP bagi mencapai tahap kerahsiaan yang boleh diterima untuk publikasi set data dan mencadangkan suatu IDS berasaskan keadaan yang disesuaikan dengan rangkaian Modbus/TCP dalam SPA dengan memanfaatkan pangkalan data pengetahuan dan bahasa peraturan berasaskan keadaan untuk mengawal keadaan sistem dan perubahan spesifikasi berterusan.

Keputusan eksperimen menunjukkan algoritma kerahsiaan yang kami cadangkan boleh bekerja dengan lebih efektif dari segi tahap kerahsiaan (12.01 berbanding 10.48), kecekapan (2.74ms berbanding 3.84ms) dan kebolehskalaan (470.15ms berbanding 507.48ms) apabila berhadapan dengan atribut trafik variat berbilang. Berkaitan dengan kehilangan maklumat, penyelesaian yang dicadangkan mampu mencapai purata 12.2% berbanding 18.6% daripada penyelesaian penanda aras. Selain itu, keputusan eksperimen IDS berasaskan keadaan menunjukkan keberkesanan yang lebih tinggi dari segi benar (99.50% berbanding 95.75%), kadar positif palsu (1.20% berbanding 1.85%) dan ketepatan (98.70% berbanding 93.68%) dalam pengenalpastian serangan dan pencerobohan. Secara keseluruhan, penyelidikan ini mencadangkan satu set penyelesaian untuk mengatasi dua masalah kerahsiaan dan keselamatan yang berkaitan dengan rangkaian Modbus / TCP dalam SPA. Hasil kerja penyelidikan yang dikemukakan dalam tesis ini adalah merupakan suatu langkah penting ke arah penghasilan SPA SCADA dan kesihatan awam yang lebih selamat.

ACKNOWLEDGEMENTS

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

{أَمَّنْ هُوَ قَانِتٌ آنَاءَ اللَّيْلِ سَاجِدًا وَقَائِمًا يَحْذَرُ الْآخِرَةَ وَيَرْجُو رَحْمَةَ رَبِّهِ قُلْ هَلْ يَسْتَوِي الَّذِينَ يَعْلَمُونَ وَالَّذِينَ لَا يَعْلَمُونَ إِنَّمَا يَتَذَكَّرُ أُولُو الْأَلْبَابِ}

صدق الله العظيم

[الزمر : 9]

In the name of God, The Most Gracious, The Most Merciful

{ Say: "Are those equal, those who know and those who do not know? It is those who are endued with understanding that receive admonition.**}**

God Almighty has spoken the truth
[Az-Zumar: 9**]**

First and foremost, I would like to thank God, who has given me the power to believe in myself and pursue my dreams, I couldn't have ever done this research without God's support. I take immense pleasure to express my sincere and deep sense of gratitude to my supervising guides and mentors, Prof. Dr. Abdullah Muhammed, Y. Bhg. Prof. Dato' Dr. Shamala A/p K Subramaniam, Prof. Dr. Azizol Hj Abdullah, Prof. Dr. Omar Khasro Akram (UrbPD) and Prof. Dr. Rui Silva, who have been an excellent source of support. Also, I would like to show my extreme thankfulness to the Mayor of Alcácer do Sal, Dr. Vitor Proença, Dr. Ana Mendes and Filipa Gonçalves, as well as to the entire IT team, Ana Rodrigues, Rui Santos, António Lameira, João Zurrapa, Miguel Gonçalves, Dr. Teresa de Sousa, Dr. Carla Mota and Dr. Martinho for their tireless and unstoppable support during all my research and trips to Malaysia. Moreover, I can't stop thanking to Eng. Andreia Graça, Eng. Palmira Martins, Dr. Ricardo, Dr. Ana Galaio and Therapist Maria Romão, for their encouragement, help, friendship, and support. Words cannot describe how grateful I am for the generosity and encouragement of my parents, José, Rosária, Deolinda, Dr. Khasro and Dr. Nada, as well as my sister Rute, my niece Carolina, my nephew Salvador, my brothers Omar and Hassan, and my friends, Margarida, Ana, Rita, Jorge, Luís, Dania, Luís, Filipa, Chico, Gustavo, Elisa and Mariana. There are so many others who I may have inadvertently left out and I sincerely thank all of them for their help.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Abdullah bin Muhammed, PhD
Associate Professor, Ts.
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Azizol bin Hj Abdullah, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Shamala A/p K Subramaniam, PhD
Professor Dato'
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

_____
ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 10 February 2022

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: Daniel José da Graça Peceguina Franco, GS45715

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____
Name of Chairman of Supervisory Committee: _____

Signature: _____
Name of Member of Supervisory Committee: _____

Signature: _____
Name of Member of Supervisory Committee: _____

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Chapter one provides an overview of the study and outlines the structure of thesis. It starts with the research background focusing on the Modbus/TCP communication protocol and its privacy and security issues, highlighting the need of protection of its communications in water distribution systems through the use of privacy algorithms and intrusion detection, problem statement, main research question, research objectives, scope of the study and its significance, ending with the limitation of research and thesis structure. This general structure will guide the study in the subsequent chapters.

1.2 Study Background

Since the beginning of the universe, water has always been an extremely important element for all living creatures. Along the centuries, Man has been giving high attention to water in their society activities, not just because it is a scarce element on the planet, but also because human activities influence this vital resource. The absence of a proper water treatment makes it dangerous for human health, resulting in many different diseases caused by bacteria or chemical substances. In developed countries, governments and other national entities give main attention to the water used for human consumption, specially to avoid human health threats.

Based on this concern, there were created tools and mechanisms able to make sure that the water is presenting a high quality and is free from elements and substances that can affect human beings and other animals (Martins, 2014). In a deeper approach, modern water distribution systems are based on cyber components for computation and communication, including not just physical components, like sensors and actuators but also the entire process control and communication networks. Here, vulnerabilities and other cyber-attacks are also a reality to this type of systems, since their supervisory control and data acquisition (SCADA) components are based on workstations, human machine interfaces and programable logic controllers, highly susceptible to attacks when connected to opened-standard TCP/IP networks and based on outdated operating systems, working 24h a day, 7 days a week. Also, because of the communication protocols standardization and interoperability, allowing multi-vendor components, the system is once more exposed to cyber threats (Adepu, Palleti, et al., 2019; Cherdantseva et al., 2015).

In a deeper analysis of the communication protocols, they are still based on the previous implementations, being embedded in the TCP payload and keeping the same lack of security as in closed-proprietary networks. In the specific case of the Modbus (Figure 1.1), it is considered one of the most popular and used protocols to manage and control water distribution systems, using not just traditional serial communications, but also Ethernet links, while allowing different devices to use this protocol as their main communication method (Fovino, 2014; Li et al., 2014).

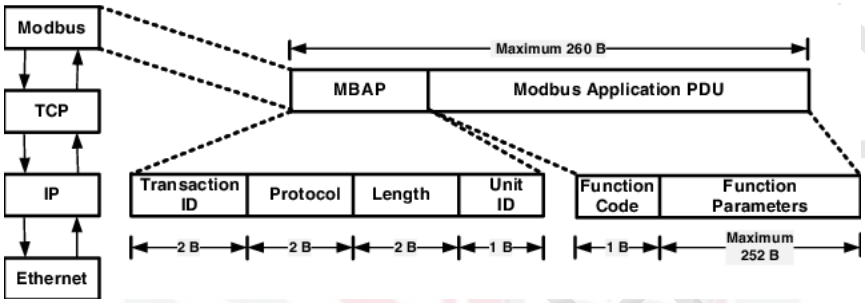

Figure 1.1: Modbus/TCP Communication Packet Shape (Edmonds et al., 2008)

The Modbus/TCP does not have any security mechanism on itself (Figueroa-Lorenzo et al. 2019), in order to protect the communications between the master device and its slaves. A good example on this lack of security, is the fact that it is not possible to find if a delivered message is indeed the original message sent from a master unit to its slave devices, or if it has been modified by an attacker. Moreover, there is also not included any anti-repudiation nor anti-replay mechanisms on Modbus/TCP, making it possible for attackers to easily damage the system or even take full control of the entire SCADA network (Parian et al. 2020).

Based on the previous described example, it is perceptible that Modbus/TCP traffic data may open the system to many types of attacks, where a previous analysis of that traffic data gives a large knowledge and advantage to attackers to properly exploit the system and gain control over it. It is a recognized fact that sharing logs is important and one of the major tools for security research and improvement, however, it is also a fact that accomplishing such logs is a difficult task, even among small groups (Al-Malawi et al. 2016).

Studies like (Werling, 2014), (Fahad et al., 2014) and (De Lima et al., 2018) focused on the need for protecting Modbus/TCP logs and traffic data, by applying anonymization and cryptographic algorithms, highlighting the problem of sensitive data exposure and the singular characteristics of Modbus/TCP packets.

Working over TCP/IP, Modbus/TCP networks are highly vulnerable, where attacks and vulnerability's exploitation have become a possibility and securing networks is now a necessity. The nature of SCADA in WDS, being time-critical, makes regular updates and patching a really difficult or even impossible operation and the inexistence of test environment and patching may introduce new unknown vulnerabilities or system failure. SCADA systems relying on old operating systems and old software are commonly found nowadays (Cherdantseva et al., 2015). Mechanisms, such as access control, VPN connections and Firewall appliances are already known to be efficient on SCADA in WDS, while cryptography and authentication mechanisms need special attention and must be used with special care, due to possible disruptive effects. Recently, intrusion detection systems, or IDSs, are being proposed to help network administrators to analyze the security risks and detect attacks against their SCADA networks and SCADA systems (Fahad et al. 2014; Nivethan and Papa 2016). Though, there are also some limitations in the use of IDS systems, making them not a generalized solution. SCADA main components, such as PLCs (Programmable Logic Controllers) and RTUs (Remote Terminal Units) have usually low computational and memory capabilities, making them not suitable to allocate a HIDS (Host-Based IDS) that must be installed on the host itself for it to be analyzed. On the other hand, NIDS (Network-Based IDS) sensors can be installed in a separated machine connected to the network to be monitored. Such approach can be easily integrated with the SCADA system, where it is necessary to understand and analyze communication protocols.

Scholars including (Al-Malawi et al., 2016; Kaouk et al., 2019; S. J. Kim et al., 2013; Monzer et al., 2019; Robles-Durazno et al., 2020; Schuster & Paul, 2012; Waagsnes, 2017) are identifying a network-based IDS with an anomaly-based approach as a good solution to be implemented on Modbus/TCP networks in water distribution systems, installing it on the network and basing its analysis on the network traffic itself, not overloading the communications with more IDS information packets. Also, this IDS type does not need high computational power on the monitored devices, since it may be installed on a separated server, connected to network and analyzing its network traffic data.

The examples of (Al-Malawi et al., 2016), (Monzer et al. 2019) and (Robles-Durazno et al., 2020), focus on the development of NIDS systems specific to Modbus/TCP networks in WDS, highlighting the lack of benchmark data sets and the need of simulation. Also, the authors emphasize the automatic extraction of state-based rules and the need of real-time anomaly monitoring against intrusions and attacks.

1.3 Problem Statement

SCADA Modbus/TCP networks in WDS were developed to work based on a high availability and under closed-networks. However, since their evolution to opened-standard networks, security became a necessity, where their clear-text

communications, without any security mechanism, play a major role on the systems' protection (Al-Malawi et al., 2016; Fahad et al., 2014). Logged network data is directly connected to the development of security solutions, though, because communications are performed in clear-text, datasets are not easily obtained due to the lack of privacy and the existing ones are simulated with a short number of attributes. Many studies were carried out based on the privacy and anonymization of Modbus/TCP packets and intrusion detection solutions for SCADA in WDS, however, many challenges are left unsolved. The main gaps that motivate this research are:

- **Segmentation of Packet Fields:** (Fahad et al., 2014) focus on the segmentation of the original Modbus/TCP packet into three different cluster attributes, separating them according to their nature (numerical, categorical and hierarchical), though, field's individual characteristics are not taken into consideration and data may be lost when applying anonymization;
- **Transformation of Data:** Also, the transformation of data (anonymization) is performed using the Euclidean distance clustering algorithm that is not adapted to binary data, used, for instance, in Modbus/TCP coil messages;
- **Minimizing the Information Loss:** Furthermore, information loss is presented as a objective, however, when applying the same anonymization algorithm to all numeric Modbus/TCP fields, it, once again, may result in data loss or not anonymized;
- **Absent System's Specifications:** In the solution presented by (Al-Malawi et al., 2016), the system's specifications are not taken into consideration and IDS rules are created in a general way, making them not suitable for all water distribution systems, since each one of them have their own specifications and configurations;
- **Extraction of State-Based Rules:** Rules are extracted using a clustering algorithm that bases its identification on the radius of the neighborhood of a given point and the maximum number of points required to form a cluster. Being totally dependent on these parameters, results are different when the parameters are changed and may result in high false positive rates;
- **Changes in System Specifications and New Devices:** Once implemented, the system doesn't have any mechanism to deal with systems' constant specification changes, demanding a total parameterization of the IDS.

1.4 Research Objectives

This study focuses on SCADA Modbus/TCP packet fields anonymization and IDS State-Based rules, assuming the following main research question (Main-RQ): *How to develop a State-based Intrusion Detection System and Privacy Optimization for SCADA Modbus/TCP Networks in Water Distribution Systems?*

As referred in problem statement, WDS network security is still not robust enough, where the use of open-standard TCP/IP networks is putting systems in

a highly vulnerable risk. To address this main research question, it is necessary to deeply study similar solutions on privacy optimization and intrusion detection systems, using them as reference and base to propose a way for Modbus/TCP data sets publication and a State-Based IDS for identification of intrusions and attacks.

Research Objective 1: To propose SCADA Modbus/TCP packet fields' privacy optimization using anonymization algorithms, aiming to increase the privacy level and reducing information loss in SCADA network.

Aiming to optimize Modbus/TCP packet fields' privacy, offering a way for SCADA system's owners to feel more comfortable and safer on publishing their network traffic data for analysis and helping on the improvement of their network security levels, the first research objective (RO1) focuses on the development of a privacy algorithm capable to anonymize Modbus/TCP packets, while keeping them able to be use in researches and studies. In order to assure its optimization, the developed algorithm must be compared to existing privacy methods and anonymization tools on SCADA Modbus/TCP networks for water distribution systems.

Research Objective 2: To propose a State-Based IDS for attacks identification, dedicated to SCADA Modbus/TCP networks in water distribution systems, capable of extracting specific automatic rules and deal with the constant system specification changes, while reducing false positive rates and increasing accuracy.

Focusing on the development of State-Base IDS rules applied to SCADA Modbus/TCP networks used in water distribution systems, the second research objective (RO2) aims to improve SCADA networks' security by the identification of intrusions and attacks.

1.5 Motivation

Water distribution systems are critical structures, spread all around the globe, representing the vast majority of physical infrastructure for water supplies, and thus constituting the primary management challenge from both an operation and public health standpoint (Cantelmi et al., 2021). Controlled by SCADA Modbus/TCP lacking of security mechanisms and working over opened-standard TCP/IP networks, water distribution systems are put at a high-risk level of vulnerabilities and attacks. Cyber-attacks to SCADA systems have increased 100% in 2014, when compared to 2013, and the presence of a large number of vulnerabilities in SCADA systems, demonstrates that they are still not strong enough as it would be expected (Singh & Kekatos, 2021). Securing water distribution systems is a major factor to properly protect human health.

1.6 Research Scope and Limitation

The proposed privacy algorithm focus on the Modbus/TCP packets fields, used in SCADA WDS networks, which are the most challenging and relevant part still working without any security mechanisms and in plain-text. At the communication layer, plain-text information can be hidden either by cryptographic or privacy algorithms or through VPN connections, thus, the main focus of this study is privacy optimization by the use of anonymization algorithms to allow Modbus/TCP data set's publication.

Intrusion detection systems are pointed as a good countermeasure to increase security on SCADA Modbus/TCP networks, where the most challenging part is to address the singular specifications and characteristics of a water distribution system. The implementation of an IDS in a water distribution system Modbus/TCP network may be done based on signatures or anomalies and at the host or networks levels. However, this research work focuses on the network level, through an anomaly-based approach to provide the identification of critical system states and recognize intrusions and attacks.

1.7 Research Contributions

The main contributions of this thesis are summarized as follows:

1. A WDS Modbus/TCP privacy algorithm for packet fields anonymization has been proposed. It was formulated an optimization problem to improve packet's data privacy, reducing the amount of plain-text within the packet, while keeping it usable for researches and other studies and prevent information loss;
2. A behavioral network intrusion detection system for SCADA Modbus/TCP networks in water distribution systems has been proposed to improve network security by identifying and monitoring system's critical states. The new solution focuses on a knowledge-based approach, taking advantage of the well-documented and described WDS and a new state-based rules' language to raise alerts in case of intrusions and attacks.

1.8 Thesis Structure

This research is divided into six different chapters:

Chapter one focuses an overview and outlines the main approached topics, including the study background, problem statement, research objectives, study scope, significance and limitations and provides a general structure to the subsequent chapters.

Chapter two highlights the three main key constructs of the study, focusing on SCADA Modbus/TCP networks, development of a state-based intrusion detection system for SCADA in water distribution systems and the privacy of Modbus/TCP packets, giving a vision of the literature review on existing anonymization and security solutions and contributing for the base of this research through its theoretical hypothesis.

Chapter three highlights the methodology and the adopted methods, focusing on a quantitative approach through different experimental procedures, defining them and how can they be applied.

Chapter four focuses on the results obtained by the privacy optimization process, comparing them with other studies and tools, as well as, describing the developed solution and its adopted anonymization algorithms.

Chapter five also focuses on results, but this from the state-based IDS analysis, highlighting the IDS architecture, state-based rules' language and knowledge database. Once again, results are compared, in a way of validation, to other similar solutions.

Chapter six highlights the summary of findings and emphasizes the discussion between main key findings, summary of the study, recommendations for future works and conclusions.

REFERENCES

Adepu, S., Kang, E., & Mathur, A. P. (2019). Challenges in secure engineering of critical infrastructure systems. *Proceedings - 2019 34th IEEE/ACM International Conference on Automated Software Engineering Workshops, ASEW 2019*. https://doi.org/10.1109/ASEW.2019.00030

Adepu, S., & Mathur, A. (2016). Distributed detection of single-stage multipoint cyber attacks in a water treatment plant. *ASIA CCS 2016 - Proceedings of the 11th ACM Asia Conference on Computer and Communications Security*. https://doi.org/10.1145/2897845.2897855

Adepu, S., & Mathur, A. (2021). Distributed Attack Detection in a Water Treatment Plant: Method and Case Study. *IEEE Transactions on Dependable and Secure Computing*, *18*(1). https://doi.org/10.1109/TDSC.2018.2875008

Adepu, S., Palleti, V. R., Mishra, G., & Mathur, A. (2019). Investigation of cyber attacks on a water distribution system. In *arXiv*.

Ahmad, Z., & Durad, M. H. (2019). Development of SCADA Simulator using Omnet++. *Proceedings of 2019 16th International Bhurban Conference on Applied Sciences and Technology, IBCAST 2019*. https://doi.org/10.1109/IBCAST.2019.8667158

Akram, O. K., Franco, D. J., Lee, A. J., & Mohammed Jamil, N. F. (2019). *Research Methods and Methodology for Undergraduate and Postgraduate Studies - Linking practical approach to theory* (v1.0). ONDAS Framework.

Al-Malawi, A., Fahad, A., Tari, Z., Alamri, A., Alghamdi, R., & Zomaya, A. Y. (2016). An Efficient Data-Driven Clustering Technique to Detect Attacks in SCADA Systems. *IEEE Transactions on Information Forensics and Security*, *11*(5). https://doi.org/10.1109/TIFS.2015.2512522

Arulselvi, S., Hemalatha, B., & Balaji, S. (2019). Security for industrial communication system using encryption/decryption modules. *International Journal of Engineering and Advanced Technology*, *8*(6 Special Issue 2), 269–273. https://doi.org/10.35940/ijeat.F1072.0886S219

Ayodeji, A., Liu, Y. kuo, Chao, N., & Yang, L. qun. (2020). A new perspective towards the development of robust data-driven intrusion detection for industrial control systems. In *Nuclear Engineering and Technology* (Vol. 52, Issue 12). https://doi.org/10.1016/j.net.2020.05.012

B & B Electronics. (2017). *Introduction to Modbus*. 1. http://www.bb-elec.com/Learning-Center/All-White-Papers/Serial/Introduction-to-Modbus.aspx

Barbosa, R. R. R., Sadre, R., & Pras, A. (2012). Towards Periodicity Based Anomaly Detection in SCADA Networks. *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*, 0–3. https://doi.org/10.1109/ETFA.2012.6489745

Bhalla, M., Pandey, N., & Kumar, B. (2015). Security Protocols for Wireless Sensor Networks. *International Conference on Green Computing and Internet of Things (ICGCIoT)*, 1005–1009. https://doi.org/10.1109/ICGCIoT.2015.7380610

Bild, R., Kuhn, K. A., & Prasser, F. (2018). A Truthful Data Anonymization Algorithm With Strong Privacy Guarantees. *Proceedings on Privacy Enhancing Technologies*, *2018*(1), 67–87. https://doi.org/10.1515/popets-2018-0004

Bradley, T. (2018). *What is an IDS and Why Do You Need It?* Alert Logic.

Cantelmi, R., Di Gravio, G., & Patriarca, R. (2021). Reviewing qualitative research approaches in the context of critical infrastructure resilience. In *Environment Systems and Decisions*. https://doi.org/10.1007/s10669-020-09795-8

Chen, T. M., & Walsh, P. J. (2014). Guarding Against Network Intrusions. In *Network and System Security: Second Edition*. https://doi.org/10.1016/B978-0-12-416689-9.00003-4

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., & Stoddart, K. (2015). A Review of Cyber Security Risk Assessment Methods for SCADA Systems. *Computers & Security*, *56*, 1–27. https://doi.org/10.1016/j.cose.2015.09.009

Cyburt, B. (2017). *Introduction to Modbus*. Automation.Com. http://www.automation.com/library/articles-white-papers/fieldbus-serial-bus-io-networks/introduction-to-modbus

De Lima, D. F., Bezerra, J. R., & De Macedo, J. F. (2018). Privacy and integrity protection of data in SCADA systems. *LANC 2018 - Proceedings of the 10th Latin American Networking Conference*. https://doi.org/10.1145/3277103.3277136

Edmonds, J., Papa, M., & Shenoi, S. (2008). Security Analysis of Multilayer SCADA Protocols: A Modbus TCP Case Study. *IFIP International Federation for Information Processing*, *253*, 205–221. https://doi.org/10.1007/978-0-387-75462-8

Environmental Protection Agency, U. S. (2017). *EPANET*. https://www.epa.gov/water-research/epanet

Fahad, A., Tari, Z., Almalawi, A., Goscinski, A., Khalil, I., & Mahmood, A. (2014). PPFSCADA: Privacy Preserving Framework for SCADA Data Publishing. *Future Generation Computer Systems*, *37*, 496–511.

https://doi.org/10.1016/j.future.2014.03.002

Farah, T. (2013). *Algorithms and Tools for Anonymization of the Internet Traffic*. School of Engineering Science.

Figueroa-Lorenzo, S., Añorga, J., & Arrizabalaga, S. (2019). A role-based access control model in modbus SCADA systems. A centralized model approach. *Sensors (Switzerland)*, *19*(20). https://doi.org/10.3390/s19204455

Finnan, K. (2016). *Water Security: The Role of the SCADA System*. Automation.Com. http://www.automation.com/library/articles-white-papers/hmi-and-s?software-technologies/water-security-the-role-of-the-scada-system

Fovino, I. N. (2014). SCADA system cyber security. In *Secure Smart Embedded Devices, Platforms and Applications* (Vol. 9781461479154). https://doi.org/10.1007/978-1-4614-7915-4_20

Fovino, I. N., Carcano, A., De Lacheze Murel, T., Trombetta, A., & Masera, M. (2010). Modbus/DNP3 state-based intrusion detection system. *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, 729–736. https://doi.org/10.1109/AINA.2010.86

Goldenberg, N., & Wool, A. (2013). Accurate Modeling of Modbus/TCP for Intrusion Detection in SCADA Systems. *International Journal of Critical Infrastructure Protection*, *6*(2), 63–75. https://doi.org/10.1016/j.ijcip.2013.05.001

Javed Butt, U., Abbod, M., Lors, A., Jahankhani, H., Jamal, A., & Kumar, A. (2019). Ransomware Threat and its Impact on SCADA. *Proceedings of 12th International Conference on Global Security, Safety and Sustainability, ICGS3 2019*. https://doi.org/10.1109/ICGS3.2019.8688327

Kali. (2019). *Kali Linux | Penetration Testing and Ethical Hacking Linux Distribution*. https://www.kali.org/

Kaouk, M., Flaus, J. M., Potet, M. L., & Groz, R. (2019). A review of intrusion detection systems for industrial control systems. *2019 6th International Conference on Control, Decision and Information Technologies, CoDIT 2019*. https://doi.org/10.1109/CoDIT.2019.8820602

Kentucky, U. of. (2015). *Communication Network - General Overview of SCADA Communications*. http://www.uky.edu/WDST/PDFs/[23] Communications network.pdf

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, *2*(1). https://doi.org/10.1186/s42400-019-0038-7

Kim, H. (2012). Security and Vulnerability of SCADA Systems Over IP-based Wireless Sensor Networks. *International Journal of Distributed Sensor Networks*, *2012*. https://doi.org/10.1155/2012/268478

Kim, S. J., Kim, B. H., Yeo, S. S., & Cho, D. E. (2013). Network anomaly detection for m-connected SCADA networks. *Proceedings - 2013 8th International Conference on Broadband, Wireless Computing, Communication and Applications, BWCCA 2013*, 351–354. https://doi.org/10.1109/BWCCA.2013.61

Koo, D., Piratla, K., & Matthews, C. J. (2015). Towards Sustainable Water Supply: Schematic Development of Big Data Collection Using Internet of Things (IoT). *Procedia Engineering*, *118*, 489–497. https://doi.org/10.1016/j.proeng.2015.08.465

Li, J. L., Zheng, W. G., Shen, C. J., & Wang, K. W. (2014). Application of modbus protocol based on μC /TCPIP in water saving irrigation in facility agricultural. *IFIP Advances in Information and Communication Technology*, *419*. https://doi.org/10.1007/978-3-642-54344-9_34

Loucks, D. P., Beek, E. Van, Stedinger, J. R., Dijkman, J. P. M., & Monique T. Villars. (2005). Water Resourses Systems Planning and Management. In *Water Resources Systems Planning and Management An Introduction to Methods, Models and Applications*. UNESCO.

Madnick, S., Sayfayn, N., & Madnick, S. (2017). *Cybersafety Analysis of the Maroochy Shire Sewage Spill Cybersafety Analysis of the Maroochy Shire Sewage Spill* (Issue May).

Martinelli, F., Mercaldo, F., & Santone, A. (2019). Real-Time SCADA Attack Detection by Means of Formal Methods. *Proceedings - 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2019*. https://doi.org/10.1109/WETICE.2019.00057

Martins, T. J. C. (2014). *Sistemas de Abastecimento de Água para Consumo Humano – Desenvolvimento e Aplicação de Ferramenta Informática para a sua Gestão Integrada* [Polytechnic Institute of Bragança, Portugal]. https://bibliotecadigital.ipb.pt/bitstream/10198/9311/1/Sistemas de Abastecimento de A?gua para Consumo Humano_versa?o final.pdf

Monzer, M. H., Beydoun, K., & Flaus, J. M. (2019). Model based rules generation for Intrusion Detection System for industrial systems. *2019 International Conference on Control, Automation and Diagnosis, ICCAD 2019 - Proceedings*. https://doi.org/10.1109/ICCAD46983.2019.9037882

Neha, N., Priyanga, S., Seshan, S., Senthilnathan, R., & Shankar Sriram, V. S. (2020). SCO-RNN: A behavioral-based intrusion detection approach for cyber physical attacks in SCADA systems. In *Lecture Notes in Networks and Systems* (Vol. 89). https://doi.org/10.1007/978-981-15-0146-3_88

Nivethan, J., & Papa, M. (2016). Dynamic Rule Generation for SCADA Intrusion Detection. *2016 IEEE Symposium on Technologies for Homeland Security, HST 2016*, 1–5. https://doi.org/10.1109/THS.2016.7568964

Ozturk, M., & Aubin, P. (2011). *SCADA Security : Challenges and Solutions* (Issue June).

Parian, C., Guldimann, T., & Bhatia, S. (2020). Fooling the Master: Exploiting Weaknesses in the Modbus Protocol. *Procedia Computer Science*, *171*. https://doi.org/10.1016/j.procs.2020.04.265

Pentair - Water Solutions. (2018). *How do I know If There is Too Much Chlorine in Water?* https://aquascienceaz.com/blog/warning-signs-of-high-chlorine-levels-in-your-water/

Phillips, B., Gamess, E., & Krishnaprasad, S. (2020). An evaluation of machine learning-based anomaly detection in a scada system using the modbus protocol. *ACMSE 2020 - Proceedings of the 2020 ACM Southeast Conference*, 188–196. https://doi.org/10.1145/3374135.3385282

Pidikiti, D. S., Kalluri, R., Kumar, R. K. S., & Bindhumadhava, B. S. (2013). SCADA Communication Protocols: Vulnerabilities, Attacks and Possible Mitigations. *CSI Transactions on ICT*, *1*(June), 135–141. https://doi.org/10.1007/s40012-013-0013-5

Premnath, A. P., Jo, J. Y., & Kim, Y. (2014). Application of NTRU Cryptographic Algorithm for SCADA Security. *ITNG 2014 - Proceedings of the 11th International Conference on Information Technology: New Generations*, 341–346. https://doi.org/10.1109/ITNG.2014.38

Rakas, S. V. B., Stojanovic, M. D., & Markovic-Petrovic, J. D. (2020). A review of research work on network-based SCADA intrusion detection systems. In *IEEE Access* (Vol. 8). https://doi.org/10.1109/ACCESS.2020.2994961

Rezai, A., Keshavarzi, P., & Moravej, Z. (2013). Secure SCADA Communication by Using a Modified Key Management Scheme. *ISA Transactions*, *52*(4), 517–524. https://doi.org/10.1016/j.isatra.2013.02.005

Rezai, A., Keshavarzi, P., & Moravej, Z. (2016). Key Management Issue in SCADA Networks: A review. *Engineering Science and Technology, an International Journal*, *August*. https://doi.org/10.1016/j.jestch.2016.08.011

Robles-Durazno, A., Moradpoor, N., McWhinnie, J., & Russell, G. (2020). Real-time anomaly intrusion detection for a clean water supply system, utilising machine learning with novel energy-based features. *Proceedings of the International Joint Conference on Neural Networks*. https://doi.org/10.1109/IJCNN48605.2020.9207462

Sanchez, G. (2017). Man-In-The-Middle Attack Against Modbus TCP Illustrated with Wireshark. In *SANS Institute Information Security Reading Room*.

Sathish, R., Kumar, D. V., & Senthilkumar, C. (2020). Design and implementation IOT based industrial sub station monitoring and control system. *Journal of Advanced Research in Dynamical and Control Systems*, *12*(7 Special Issue). https://doi.org/10.5373/JARDCS/V12SP7/20202291

Schneider Electric. (2017). *What is the difference between Modbus and Modbus Plus?* http://www.schneider-electric.ca/en/faqs/FA198221/

Schuster, F., & Paul, A. (2012). A Distributed Intrusion Detection System for Industrial Automation Networks. *Proceedings of 2012 IEEE 17th International Conference on Emerging Technologies & Factory Automation (ETFA 2012)*, 1–4. https://doi.org/10.1109/ETFA.2012.6489703

Shahzad, A., Lee, M., Kim, H., Woo, S., & Xiong, N. (2015). New Security Development and Trends to Secure the SCADA Sensors Automated Transmission during Critical Sessions. *Symmetry*, *7*(4), 1945–1980. https://doi.org/10.3390/sym7041945

Shahzad, Aa., Guangzhi, Z., Chae, H., Lee, M., & Irfan, M. (2015). The Approximate Approaches in Addressing of SCADA Security Issues. *Journal of IT and Economic Development*, *6*(2), 31–35.

Shirazi, S. N., Gouglidis, A., Syeda, K. N., Simpson, S., Mauthe, A., Stephanakis, I. M., & Hutchison, D. (2016). Evaluation of anomaly detection techniques for SCADA communication resilience. *Proceedings - 2016 Resilience Week, RWS 2016*, 140–145. https://doi.org/10.1109/RWEEK.2016.7573322

Simi, M. S., Nayaki, K. S., & Elayidom, M. S. (2017). An Extensive Study on Data Anonymization Algorithms Based on K-Anonymity. *IOP Conference Series: Materials Science and Engineering*, *225*(1). https://doi.org/10.1088/1757-899X/225/1/012279

Simply Modbus. (2015). *About Modbus TCP*. http://www.simplymodbus.ca/TCP.htm

Singh, M. K., & Kekatos, V. (2021). On the flow problem in water distribution networks: Uniqueness and solvers. *IEEE Transactions on Control of Network Systems*, *8*(1). https://doi.org/10.1109/TCNS.2020.3029150

Stevens, C. (2020). Assembling cybersecurity: The politics and materiality of technical malware reports and the case of Stuxnet. *Contemporary Security Policy*, *41*(1). https://doi.org/10.1080/13523260.2019.1675258

Tellerreport. (2020). *A Renewed Scenario... The death of huge numbers of fish in southern Iraq*. https://www.tellerreport.com/news/2020-08-04-watch---a-renewed-scenario-----the-death-of-huge-numbers-of-fish-in-southern-iraq.SkECTj3UZv.html

Tesfahun, A., & Bhaskari, D. L. (2016). A SCADA Testbed for Investigating Cyber Security Vulnerabilities in Critical Infrastructures. *Automatic Control and Computer Sciences*, *50*(1), 54–62. https://doi.org/10.3103/S0146411616010090

Texas, C. E. (2012). *BLOOM FILTER BASED INTRUSION DETECTION FOR SMART GRID SCADA Saranya Parthasarathy and Deepa Kundur*. *May*.

Tuna, G., Nefzi, B., Arkoc, O., & Potirakis, S. M. (2014). Wireless Sensor Network-Based Water Quality Monitoring System. *Key Engineering Materials*, *605*(April), 47–50. https://doi.org/10.4028/www.scientific.net/KEM.605.47

Valli, C. (2009). Snort IDS for SCADA Networks. *Management*, *2009*, 618–621. http://ro.ecu.edu.au/ecuworks/529/

Vinchurkar, D. P., & Reshamwala, A. (2012). A Review of Intrusion Detection System Using Neural Network and Machine Learning Technique. *International Journal of Engineering Science and Innovative Technology (IJESIT)*, *1*(2), 54–63.

Waagsnes, H. (2017). *SCADA Intrusion Detection System Test Framework* (Issue May). University of Agder.

Walliman, N. (2018). *Research Methods - The Basics* (Second). Routledge.

Werling, J. R. (2014). *Behavioral Profiling of Scada Network Traffic Using Machine Learning Algorithms*. Air Force Institute of Technology.

WHO. (2018). *WHO EMRO | WHO and Ministry of Health and Environment investigate the mass death of fish in southern governorates of Iraq*. https://www.emro.who.int/irq/iraq-news/death-of-fish-investigation.html

Witte. (2021). *Modbus test and simulation*. https://www.modbustools.com/index.html

Yang, H., Cheng, L., & Chuah, M. C. (2019). Deep-Learning-Based Network Intrusion Detection for SCADA Systems. *2019 IEEE Conference on Communications and Network Security, CNS 2019*. https://doi.org/10.1109/CNS.2019.8802785

Zhang, J., Gan, S., Liu, X., & Zhu, P. (2016). Intrusion Detection in SCADA Systems by Traffic Periodicity and Telemetry Analysis. *Proceedings - IEEE Symposium on Computers and Communications*, *2016-Augus*, 318–325. https://doi.org/10.1109/ISCC.2016.7543760