MICRO-CI: A CRITICAL SYSTEMS TESTBED FOR CYBER-SECURITY RESEARCH

William Hurst, Nathan Shone, Qi Shi
Department of Computer Science
Liverpool John Moores University
Byrom Street
Liverpool, L3 3AF, UK
{W.Hurst, N.Shone, Q.Shi}@ljmu.ac.uk

Behnam Bazli
School of Computing
Staffordshire University
Beaconside
Stafford, ST18 0AD
Behnam.Bazli@staffs.ac.uk

Abstract— A significant challenge for governments around the globe is the need to improve the level of awareness for citizens and businesses about the threats that exist in cyberspace. The arrival of new information technologies has resulted in different types of criminal activities, which previously did not exist, with the potential to cause extensive damage. Given the fact that the Internet is boundary-less, it makes it difficult to identify where attacks originate from and how to counter them. The only solution is to improve the level of support for security systems and evolve the defences against cyber-attacks. This project supports the development of critical infrastructure security research, in the fight against a growing threat from the digital domain. However, the real-world evaluation of emerging security systems for Supervisory Control and Data Acquisition (SCADA) systems is impractical. The research project furthers the knowledge and understanding of Information Systems; specifically acting as a facilitator for cyber-security research. In this paper, the construction of a testbed and datasets for cyber-security and critical infrastructure research are presented.

Index Terms—Critical Infrastructure, SCADA, Testbed, Security

1. INTRODUCTION

Interconnected control systems such as SCADA (Supervisory Control and Data Acquisition) monitor and govern public infrastructures, such as power plants and water distribution networks [1]. A constant assessment of the security of working control systems is necessary to ensure critical infrastructures are secured against external cyber-threats. However, this assessment process can impact the availability and performance of the control system. These types of environment require constant service provision and any disruption can be costly, and impact upon the end-users. For that reason, alternative approaches to assessing the security of automated control systems are needed.

Non-virtualised physical testbeds are costly and inaccessible, and are often location constrained [1]. As such, modern education and research for control system security is becoming increasingly reliant on virtualised labs and tools [2]. Any learning or research undertaken using these tools, however, is based around the limitations and characteristics of such tools, as well as any assumptions made by their developers. Additionally, the accuracy of data resulting from emulations and models may be further decreased if used outside of their intended usage scenario. It is for that reason that projects such as SCADAVT propose testbed frameworks for cyber-security experimentation, based on a simulation approach [1].

A virtualised approach offers significant cost savings and a self-paced and active approach to learning. However, it also has several key limitations including: no hands-on experience, no real-world training with specific equipment and no experience in identifying and interpreting incorrect or uncharacteristic data. Simulation is effective at representing "*correct*" behaviour. However, critical infrastructure systems need to be protected against situations where they are exposed to extreme abnormal events. Unfortunately, in such circumstances, systems do not always behave in the way expected or respond in the same consistent manner. Similarly, it is therefore difficult to accurately model how a system's erratic behaviour might cascade and impact other parts of the infrastructure.

The research presented in this paper provides an ideal solution. The practical element involved in the Micro-CI project introduces a level of realism that is difficult to match through simulation alone. It allows for the advantages of both physical and virtual tools to be combined, and some of these are discussed below.

- Pedagogical benefits: The Micro-CI approach offers students and researchers hands-on experience and first-hand knowledge of the unpredictability of a system under attack or stress. It will also help them to refine their problem solving and practical skills.
- Cost effectiveness: The Micro-CI project has been designed to be as cost effective as possible. For example, at the time of writing, we estimate that at the time of writing, the design presented in this paper can be replicated at low cost.
- Portability: As the project components are on a miniaturised bench top scale, it enables them to be packed away, stored and transported with ease. Projects can still be moved and/or stored whilst partially assembled.

- Platform independency: The Micro-CI project does not require any specific requirements, dependencies or operating systems to interact with the testbeds developed. Additionally, it is not tied or restricted by any licencing model, so it can be used on an infinite number of different machines, without incurring additional costs.

In this paper, the architecture for the Micro-CI testbed, which replicates a water distribution plant, is outlined. Both the physical design and construction of the testbed are detailed. A case study and evaluation, in which cyber-attacks are launched against the water distribution plant, are also presented.

The remainder of this paper is organised as follows. Section 2 presents a background discussion on testbed and critical infrastructure modelling. Cyber-security and cyber-threats are also highlighted. Section 3 presents the approach used to construct the Micro-CI testbed and a case study of the impact of an attack on the system. The resulting data is evaluated in Section 4. Finally, the paper is concluded and the future work is highlighted in Section 5.

2. BACKGROUND

As automation grows in all areas of critical infrastructures [4], increased pressure is put on control systems to oversee and monitor operations at all times. A central control unit has the job of governing the behaviour of a vast system, ensuring the infrastructure is run smoothly and automated efficiently [7].

A. Control Systems

Centralised control systems enable operators to control components from remote locations without physically needing to be there [5]. The requirement of operators having to travel to distant locations has been replaced by carefully designed user interfaces that allow the operator to interact with the system. The interfaces are often comprised of off-the-shelf components constructed to a specification suitable for the type of infrastructure being used [6]. The use of off-the-shelf components is a cause for concern, as the technology used in infrastructures is readily available for anyone to use, which can make them more vulnerable to attack [6].

A typical modern control system consists of a network of sensors acquiring data, which are used to control devices using programmable logic controllers (PLC). They are typically composed of three parts, the master terminal unit (MTU), remote terminal units (RTU) and the communication links. The MTU acquires data and sends instructions between various components such as a Human Machine Interface (HMI), databases for storing past information, workstations for engineers and business information systems for industrial applications. RTUs are essentially PLC devices that automate the actions ordered by the master terminal unit. The communication links are responsible for proficient communication and usually consist of fibre optics, microwave, telephone lines, pilot cables, radio or satellite.

Figure 1 provides a simplified illustrative overview of a control system, whereby the RTU provides the communication link between various components of the infrastructure and the MTU, which is linked to the graphical user interface.

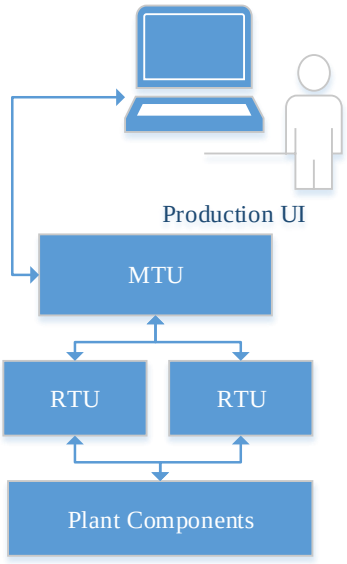

Figure 1 SCADA Overview

At this stage, the SCADA system is typically software that enables the operator to interact with the MTU and observe the on-going activities in the infrastructure [10]. Other approaches to control system construction can include a DCS (Distributed Control System) layout, which tends to have no central controller but is operated by various control components working together to decide the required action [8], [9]. However, the Micro-CI testbed detailed in this paper follows the traditional centralised control system structure for the purposes of simplicity and replicability.

B. The Cyber Threat

Control system data is coded in protocol format to exchange information with components and RTUs. The protocol formats provide automation and send information back to the control user interface to deliver a status of system operations. Communication protocols are designed for real-time operation [11]. Two examples of industrial control network protocols include Modbus (Modicon Communication Bus) and DNP3 (Distributed Network Protocol). They are commonly used in modern day critical infrastructures and able to match the specific requirements of the system. However, they are susceptible to disruption and security breaches [11]. One of the most common methods of attack is the Distributed Denial of Service (DDoS) attack, where systems are sent large volumes of traffic that is intended to make the system fail by overloading it. This attack is effective. It is a challenge to distinguish between good and bad requests, making attacks problematic to block [13].

Often cyber-attacks are specifically targeted at individual parts of infrastructures. For example, various attacks are

designed with the precise intention of disrupting or infiltrating SCADA systems [12]. One such attack is known as a Process Network Malware Infection (PNMI), which involves injecting a worm into the process network. The process network is often used for hosting the whole of the SCADA where communication is conducted through protocols like ModBus or DNP3 [12]. Another common technique is the Man in the Middle attack (MITM) [14] where false commands or system instructions and fake responses are inserted into the system. Not only can a MITM attack be used to cause disruption; it can also be used to provide a way of eavesdropping; making it important to use authentication protocols to ensure the confidentiality and integrity of the communications [15].

C. *Related Projects*

The growing cyber-threat has led to a switch in research focus from physical protection to digital infrastructure security measures. However, this cyber-security research is hampered by a lack of realistic experimental data and opportunities to test new theories in a real-world environment [16]. For that reason, projects such as SCADAVT, have developed simulation-based testbed, which builds upon the CORE emulator, for building realistic SCADA models [1].

In their approach, Almalawi *et al*., develop a framework to construct a water distribution system [1]. The testbed consists of SCADA components, including the Modbus/TPC slave and master, and the Modbus/TPC HNI server. Functioning together, the testbed employs the use of the dynamic link library (DLL) of EPANET to simulate the water flow within the system. The testbed combines the use of existing techniques to produce a novel testbed application. The system tested through a case study involving a DDoS attack to demonstrate that convincing data-construction is possible. Software-based simulation data, such as this approach, is often used to test theoretical cyber-security systems; however, data constructed through emulators is inherently lacking in realism and a hands-on learning experience is missed.

In addition to the aforementioned water distribution testbed approach, there are several existing proposals for critical infrastructure testbed architectures, which focus on specific systems, such as electricity substations [17]. However, our long-term goal is not to constrain our testbed to a single role, but to adopt a modular approach; whereby new critical infrastructure roles can be integrated at a later stage. This would make it suitable and useful to a wider audience. Specifically, the proposed system focuses on a water distribution plant; however, the design is extendable and testbeds can be extended to incorporate other infrastructure types, such as an ecologically-aware power plant.

This project provides research opportunities for the testing and development of security enhancements in a real-life scenario. As such, the aim of the research is to have a practical output; a fully working critical infrastructure testbed. The goal is to demonstrate the suitability of the datasets generated by the Micro-CI testbed, which can also provide a benchmark for future comparison against those created by industry-standard software.

3. Approach

The Micro-CI project addresses the lack of both access to experimental data and the hands-on experience needed to properly understand the challenges involved in an era of growing digital threats. As such, the intended output of this project is to support the construction of a bespoke bench-top testbed for data generation; consisting of a model critical infrastructure and control system. The testbed will be used for cyber-security research purposes and testing new experimental methods for enhancing the level of security in cyber-critical systems, specifically those under current exploration by the investigators. In this section, an outline of the architecture of the Micro-CI project is presented.

A. *System Design*

The design displayed below in Figure 2, presents a rudimentary water distribution plant. The specification is modest, meaning there is scope for future expansion; yet is sufficient in size to produce realistic infrastructure behaviour datasets for research purposes. As illustrated in the diagram, there are two reservoir tanks, which are fed by two pumps moving water from external sources. The remote terminal unit (RTU) is used to monitor the outgoing flow rate and water level, to dynamically adjust the pump speed ensuring adequate replenishment of the reservoir tanks. However, vulnerabilities exist in the system, meaning that it is possible for an external source to cut off the water supply or flood the reservoir tanks. This can be achieved by switching off or speeding up either of the pumps used to control the water flow.

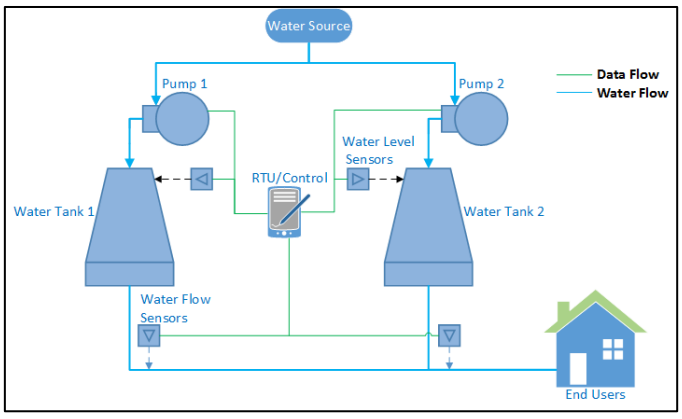

Figure 2. Water distribution plant testbed architecture

B. *Practical Micro-CI implementation*

The practical implementation of the testbed includes the following physical components: an Arduino Uno Rev. 3 as the RTU, two 12v peristaltic pumps as the water pumps, two liquid flow meters, two water level sensors, two amplification transistors, diodes, resistors and an LCD.

In the schematics shown in Figure 3, potentiometer symbols have been used in place of sensors; this is due to the

limited symbols available in the blueprint software. As the maximum output of the Arduino is only 5v, transistors amplify this to the 12v required by the pumps. Lastly, the diodes are used to ensure the current can only travel in one direction, thus preventing damage to the Arduino. The hardware specification used is modest, meaning there is scope for future expansion; yet is sufficient in size to produce realistic infrastructure behaviour datasets for research purposes.

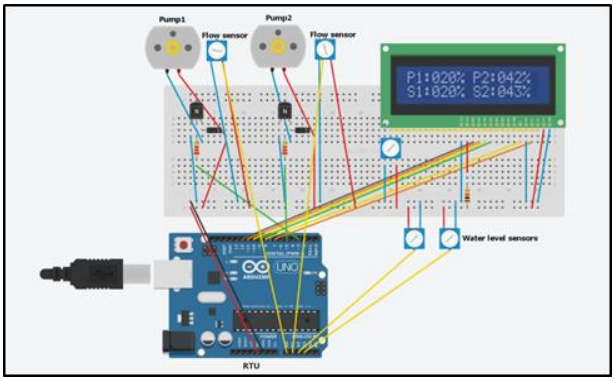

Figure 3. Physical wiring schematics

The construction is displayed in Figure 4. For the purpose of this experiment, the Arduino board remains connected to a PC via a USB cable (although this could be replaced with a network connection for similar experiments). The system is also inactive.

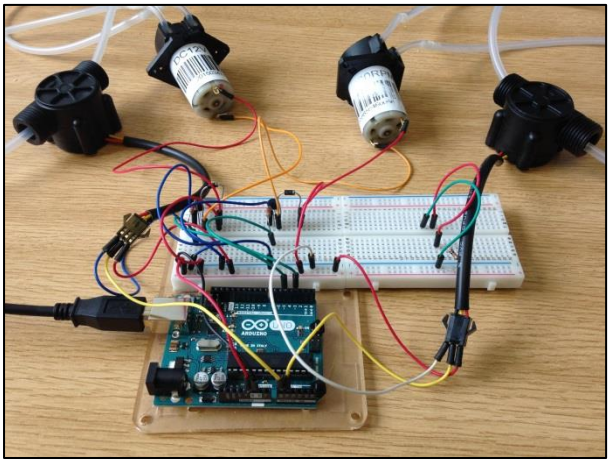

Figure 4. Testbed Construction

Through this USB connection, a serial connection is established to supply a real-time data feed, which is recorded and preserved by the PC (as illustrated in Figure 3). The metrics collected in this instance include: Water level sensor1/2 readings, Flow meter1/2 readings and Pump1/2 speeds. These readings are taken from each sensor every 0.25 seconds (4Hz) and written to the serial data stream.

To examine the quality of the data produced by the Micro-CI implementation, a dataset was recorded over the period of 1 hour. During this time, the testbed was operating under normal parameters (i.e. no cyber-attacks were present).

Essentially, this means that the pump speeds are configured to slowly continue filling the tanks at a controlled speed until full (even if no water is being used) and to cover the current rate of water consumption (if possible). The outflow (water being consumed) is a randomly applied value within a specific range (to make usage patterns more realistic). In this instance, the water source pipe is 60% smaller than the outflow pipes, which allows for a more accurate representation and to enable water tank starvation in case of overload. A sample of the data collection process is displayed in Figure 5, which shows the Arduino Serial Monitor.

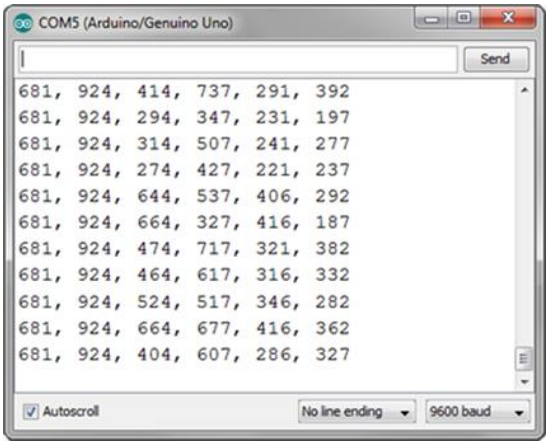

Figure 5. Example Serial Data Stream

The datasets produced by the testbed are evaluated in the following section, as a demonstration of their applicability in a critical infrastructure research setting.

4. Evaluation

Similar to the SCADAVT project, this testbed is evaluated through the demonstration of a Distributed Denial of Service attack [1]. The effect of a DDoS attack in comparison with normal behaviour of the testbed is evaluated.

A. Data Construction

As such, for the first part of this case study, data for the water distribution plant is recorded whilst operating under normal conditions. This enables the building of a behavioural norm profile for the system. Within the testbed, during the DDoS attack, only intermittent readings from the sensors are received, forcing it to make drastic (and therefore uncharacteristic) changes to the pump speeds, rather than gradual as when operating normally.

A small sample of the data obtained at 00:10.5 in run time is shown in Table 1. There is no significant variation present in the data. All the metrics maintain consistent trends in operation.

Within Table 1, C1 to C6 denote the system components used for data collection. As such, C1 and C2 denote the water level in tank 1 and 2 correspondingly; C3 and C4 signify the water levels in tank 2 and 3; C4 represents the water flow from tank 2; C5 denotes the speed of pump 1 and C6 indicates the speed of pump 2.

Table 1. Normal Physical testbed Data Sample

Sample(*t*)	C1	C2	C3	C4	C5	C6
00:10.5	65.0	69.9	47.3	55.4	81.9	85.1
00:10.7	65.0	69.9	39.4	48.5	74.1	78.8
00:11.0	65.0	69.9	39.4	53.4	74.1	83.1

Table 2 represents the distribution of values for each of the components over the 1 hour simulation. The unique value, max, min, median, mean and standard deviation of the values are demonstrated.

Table 2. Distribution Values for Normal Data

Assessment	C1	C2	C3	C4	C5	C6
unique (est.)	23.00	4.00	55.00	52.00	51.00	53.00
min:	65.00	69.99	34.82	44.86	68.91	74.83
max:	66.14	70.19	38.19	47.96	72.91	77.92
median:	65.59	70.09	36.86	46.69	71.30	76.64
mean:	65.58	70.07	36.72	46.54	71.14	76.50
std:	0.328	0.063	0.691	0.694	0.772	0.695

B. Attack Data Construction

The DDoS attack on the system, which is launched against the RTU's communications channel, results in intermittent sensor readings. Whilst no new values are readily available, the RTU continues to maintain the previous pump speed. As before, Table 3 represents the distribution of values for each of the components over the 1 hour simulation during the cyber-attack scenario. The unique value, max, min, median, mean and standard deviation of the values are again demonstrated.

Table 3. Distribution Values for Attack Data

Assessment	C1	C2	C3	C4	C5	C6
unique (est.)	25.00	7.00	52.00	51.00	58.00	59.00
min:	65.00	69.89	34.91	44.69	56.42	56.84
max:	66.18	70.02	38.16	47.98	85.05	91.45
median:	65.69	69.99	36.86	46.69	71.64	75.51
mean:	65.62	69.99	36.71	46.54	72.07	74.88
std:	0.36	0.015	0.689	0.705	6.832	7.327

Whilst under attack, the testbed continues to function, service is disrupted and the output is visible in the dataset constructed. For example, the min and max values from C5 are considerably different. This change is data is identifiable in a visual comparison of the pump speeds.

Figure 6 presents a two feature scatter plot of the normal and abnormal operation of the two testbed pumps. Pump speed 2 is displayed along the y-axis with pump speed 1 detailed on the x-axis.

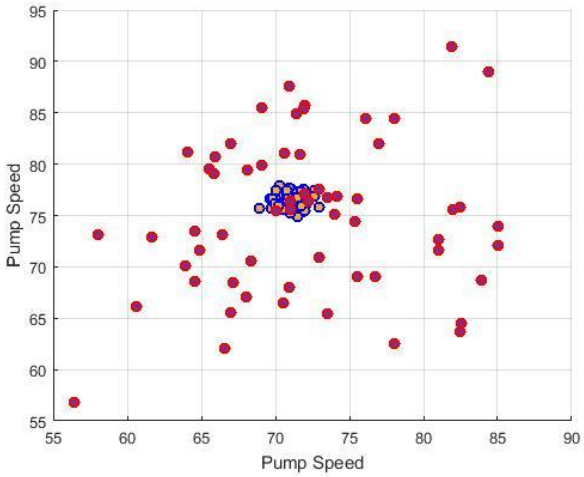

Figure 6. Scatter Plot for Normal and Abnormal Operation of Pumps

Variance can be seen in the clustering of the data. The colour indicates the grouping. The normal behaviour of pump shows a deviation in the operational speed to adjust for water flow changes and maintaining the water level in the tanks. This is displayed in the red circles, distributed throughout Figure 6. Within the attack data, the RTU's communications channel is unable to maintain active communication the water levels meaning that pump 2 is unable to adjust efficiently to match the tank water levels. This is reflected in the small data cluster in the centre of Figure 6.

C. Discussion

Current anomaly detection systems function by identifying a deviation from established patterns within given datasets. For example, in the case of network security, algorithms compare network flow with historical flows and outliers in the datasets are subsequently identified. Threats are marked as an anomaly. Supervised learning algorithms are generally employed to learn about the threats and establish patterns of attack behaviour, which are labelled as signatures. This allows for the detection of novel attacks. Based on the above evaluation, we envision that this testbed would be ideal for training and research which focuses on anomaly and signature-based detection. Specifically, this testbed offers the following benefits:

- Normal/Abnormal dataset construction: Normal datasets can be constructed to act at the established pattern of system behaviour. Abnormal dataset can be constructed to act as deviation from normal patterns of behaviour. This can allow for experimentation with novel detection algorithms. Both types of data are needed for the core functionality for how anomaly detection practice is done.
- Diverse application: The testbed is applicable to a range of CI scenarios and cyber-attack types. The above evaluation focuses on DDoS on a water plant; however, Denial of Service, Signature Injection, are further examples of attack scenarios which can be implemented for dataset construction. The main components of the

system, for example the pumps, can also be altered to change the CI infrastructure type.

5. CONCLUSION

As previously discussed, one of the aims of this project is to devise a testbed, which is suitable for cyber-security training and research. It is our belief that the use of real-life data is more suitable for cyber-security research, than that of simulation only. One of the most effective aspects of the Micro-CI testbed is its expandability; meaning that in future work the scale of the testbed can be expanded to incorporate additional components and sensors.

However, as with all solutions, there are some drawbacks to our approach. The first is that the use of low cost hardware reduces the level of accuracy that can be achieved. For example, the Arduino Uno uses an ATMega microcontroller, which is only capable of recording 4-byte precision in double values. This can present problems if precision is a crucial part of the research being undertaken. However, this can be mitigated by purchasing more expensive hardware. Another limitation is that in comparison to simulation software, the practical approach may require a greater level of improvement to students' skillsets (which is not a detrimental attribute), and a longer initial construction time, to accomplish a working implementation.

One of the main challenges for governments around the globe is the need to improve the level of awareness for citizens and businesses about the threats that exist in cyberspace. The arrival of new information technologies has resulted in different types of criminal activities, which previously did not exist, with the potential to cause extensive damage to internal markets.

Society is becoming increasingly reliant upon critical infrastructure systems, which is forcing them to become more accessible and interconnected, in a short space of time. When this is combined with the growing sophistication of cyber-attacks, this poses a considerable physical and digital security threat. Hence, critical infrastructure security is a key area of much-needed research that is under-supported. We hope that Micro-CI will provide a cost-effective, yet realistically accurate tool for future cyber-security research and learning. In our future work, we will compare the results from Micro-CI against existing industry-specific simulation software. We will also make the datasets available for cyber-security and critical infrastructure research. In addition, the construction design and instructions will be made available to other researchers.

6. ACKNOWLEDGEMENTS

The authors would like to thank the UK Academy for Information Systems (UKAIS) as the funding body for this research project (http://www.ukais.org.uk/).

REFERENCES

[1] A. Abdulmohsen., Z. Tari., I. Khalil., and A. Fahad., SCADAVT-A framework for SCADA security testbed based on virtualization technology, Proceedings of the 38th IEEE Conference on Local Computer Networks (LCN), pp639-646, 2013

[2] L. Topham, K. Kifayat, Y. A. Younis, Q. Shi and B. Askwith, Cyber Security Teaching and Learning Laboratories: A Survey, Information & Security: An International Journal, vol. 35, 2016.

[3] D. Lewis, The pedagogical benefits and pitfalls of virtual tools for teaching and learning laboratory practices in the Biological Sciences, HE Academy, 2014

[4] L. H. de Melo Leite, L. de Errico, and W. do Couto Boaventura, Criteria for the selection of communication infrastructure applied to power distribution automation, Proceedings of the IEEE PES Conference on Innovative Smart Grid Technologies (ISGT Latin America), pp. 1–8, 2013.

[5] O. Gerstel, AControl Architectures for Multi-Layer Networking: Distributed, centralized, or something in between? Optical Fiber Communications Conference and Exhibition (OFC), pp 1-16, 2015.

[6] C. Esposito, D. Cotroneo, R. Barbosa, and N. Silva, Qualification and Selection of Off-the-Shelf Components for Safety Critical Systems: A Systematic Approach, Proceedings of the Fifth Latin-American Symposium on Dependable Computing Workshops, pp. 52–57, 2011.

[7] V. Urias, B. Van Leeuwen, and B. Richardson, Supervisory Command and Data Acquisition (SCADA) system cyber security analysis using a live, virtual, and constructive (LVC) testbed, Proceedings of the IEEE Military Communications Conference, (MILCOM), pp. 1–8, 2012.

[8] Z. Liu., D. Li., L. Yun., and S. Xu., An assessment method for reliability of distributed control system, Proceedings of the IEEE International Conference on Information and Automation, pp. 1300-1304, 2015.

[9] H. Fayyaz Abbasi., N. Iqbal., M. Rehan, Distributed Robust Adaptive Observer-Based Controller for Distributed Control Systems with Lipschitz Nonlinearities and Time Delays, Proceedings of the 13th International Conference on Frontiers of Information Technology (FIT), pp. 185–192, 2015.

[10] J. Adrian Ruiz Carmona., J. César Muñoz Benítez and J. L. García-Gervacio., SCADA system design: A proposal for optimizing a production line, Proceedings of the International Conference on Electronics, Communications and Computers (CONIELECOMP), pp. 192-197, 2016.

[11] R. Gao and C. Hwa Chang, A scalable and flexible communication protocol in a heterogeneous network, Proceedings of the 13th International Conference on Computer and Information Science (ICIS), pp 49-52, 2014.

[12] Y. Zhang., L. Wang., Y. Xiang and C. Ten, Inclusion of SCADA Cyber Vulnerability in Power System Reliability Assessment Considering Optimal Resources Allocation. IEEE Transactions on Power Systems, Vol:PP, No 99, pp 1-16, 2016.

[13] Q. Yan., F. R. Yu., Q. Gong., and J. Li., Software-Defined Networking (SDN) and Distributed Denial of Service (DDoS) Attacks in Cloud Computing Environments: A Survey, Some Research Issues, and Challenges, IEEE Communications Surveys & Tutorials, Vol. 18 No. 1, pp. 602–622, 2015.

[14] A. Sahi Khader., and D. Lai., Preventing man-in-the-middle attack in Diffie-Hellman key exchange protocol, Proceedings of the 22nd International Conference on Telecommunications (ICT), pp. 204–208, 2015.

[15] R. Divya., and S. Muthukumarasamy., An impervious QR-based visual authentication protocols to prevent black-bag cryptanalysis, Proceedings of 9th IEEE International Conference on Intelligent Systems and Control (ISCO), pp. 1–6, 2015.

[16] T. Benzel, R. Braden, D. Kim and C. Neuman, Experience with DETER: a testbed for security research, in Proceedings of the 2nd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, 2014.

[17] Z. L. H. Wei, G. Yajuan, and C. Hao, Research on information security testing technology for smart Substations, in Proceedings of the International Conference on Power System Technology (POWERCON), pp. 2492–2497, 2014.

[18] M. Ficco, G. Avolio, L. Battaglia, and V. Manetti, Hybrid Simulation of Distributed Large-Scale Critical Infrastructures, Intell. Netw. Collab. Syst., pp. 616–621, 2014.