Security Education for Smart Grid: Materials, Experiments, and Evaluation

Weichao Wang[1], Chuang Wang[1], Le Xie[2], Wen-zhan Song[3], and Yi Pan[3]

[1]: University of North Carolina at Charlotte

[2]: Texas A&M University

[3]: Georgia State University

Abstract: With the fast development of Cyber-Physical systems (CPS), security in these special application environments starts to attract more and more efforts. In this project, we form a team of researchers in information security, power systems, simulation, and education evaluation to jointly develop educational materials and experiments for security education in smart grid. Multiple course modules for infrastructure and data security in smart gird have been designed. We design a simulation/emulation based experiment platform and develop several student projects upon it. These materials have been adopted by both graduate and undergraduate level security courses. Formal evaluations are conducted by third party evaluators.

1. Introduction

With the proliferation of Cyber-Physical Systems (CPS) and Mission Critical Operations (MCO), both educational institutions and industrial corporations start to emphasize the connection between computer networks and real shop-floor environments. The ever-increasing penetration of Internet of Things (IoT) introduces new security threats to the application scenarios. Equipping the engineering students with basic knowledge and skills in information security and privacy becomes an essential task for these majors.

Compared to the demands in these areas, the availability of educational materials, especially hands-on experiment platforms, falls behind in many aspects. Several reasons lead to this

deficiency. First, the tight integration of Internet and control networks of physical systems happened very recently. Researchers are still focusing on resolving the technical problems caused by the merge. The development of training environments has just started to attract attention. Second, there exists a large gap in problem presentation and solving between the cyber systems and physical systems. Therefore, the development team of the educational materials must have expertise from both sides. Last but not least, educational experiment platforms for smart grid often involve physical equipment. They are different from simulation or emulation environments for computer networks. For example, an emulator for security education in smart grid needs to have the corresponding components in an electric system, which could raise serious safety issues. These issues restrict the development of such materials.

To bridge the gap, we have proposed and implemented a suite of educational materials for security in smart grid. The materials include a group of class modules and several student projects upon a simulation/emulation environment. The class modules cover the network and data security in smart grid. The experiment platform is built upon the open source Common Open Research Emulator (CORE) system designed by US Navy [1]. We implement components for power systems and data communication in the platform. The functionalities such as information exchange and event synchronization between the two types of networks are also implemented. We design multiple student projects on security of Advanced Metering Infrastructure (AMI) protocols, and impacts of contaminated data on system performance and stability. We have adopted the materials in both undergraduate and graduate courses. Our education evaluation expert conducted surveys and interviews with these students to assess the learning outcomes.

The contributions of our project are as follows. First, we have designed a suite of educational materials and corresponding experiments for security education in smart grid. The materials cover security and safety of both cyber and physical systems and their mutual impacts. Second, different from the co-simulation environments for smart grid [2][3][4] that are designed by power industry companies, our experiment platform is open source.

Therefore, other researchers and educators can adopt it or develop new educational materials upon it with relatively low learning curves. Last but not least, the learning outcomes of our approach are formally evaluated by education experts.

The remainder of the paper is organized as follows. In Second 2, we introduce the architecture of our simulation/emulation platform. In Section 3, we will describe the design of the course modules and experiments using the proposed approach. The evaluation activities and results will be discussed in Section 4. Finally, Section 5 concludes the paper.

2. The Experiment Platform

The proposed experiment environment consists of three components: smart grid emulator, computer network simulator, and a front-end user interaction component. These components are tightly coupled together. They regularly exchange information to maintain synchronization. Below we will describe each component in detail.

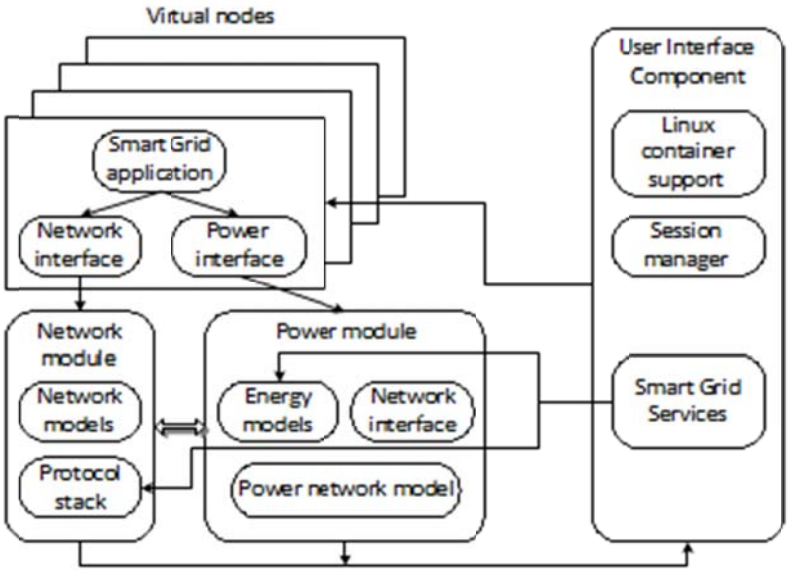

Figure 1: Architecture of the proposed experiment platform.

The power grid emulator is built upon the Common Open Research Emulator (CORE) system. It provides a platform for emulating

smart grid environments, and running real smart grid applications by applying virtualization techniques in Linux. Here each emulated node can have its own instance of virtual or real network devices, network protocol stack, and process space while sharing the file system of the emulation server with other nodes. Since the virtual nodes are communicating with each other through sockets, they can be distributed over different physical machines. This light-weight virtualization approach guarantees the scalability of our approach. At the same time, researchers can develop smart grid applications and educational materials under Linux. They can be easily ported into our environment.

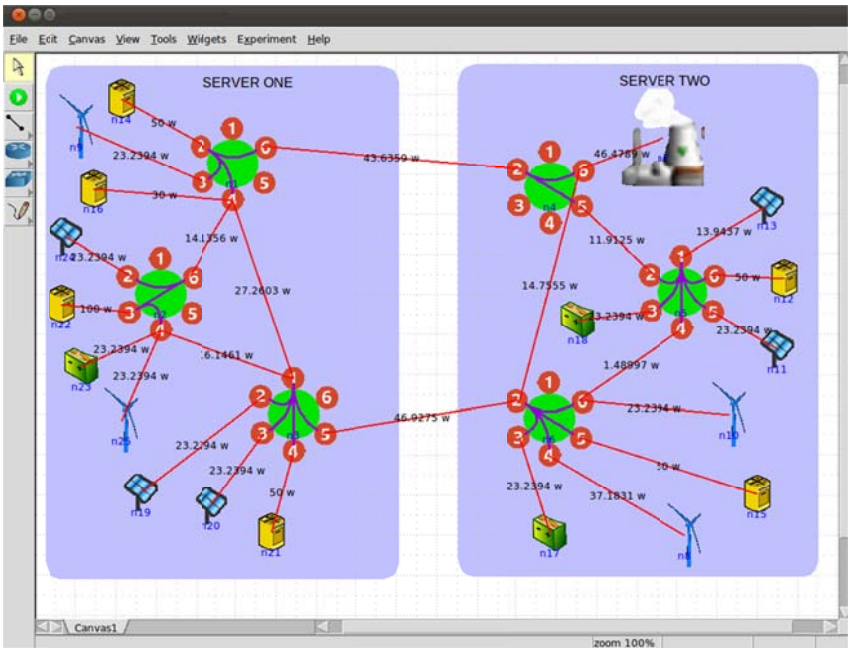

Figure 2: User interface of the proposed platform.

While communication among the virtual nodes can be achieved through the APIs of Linux, the emulation system faces the problem of synchronization between the cyber and physical sub-systems: while a communication network is often modeled as a discrete, event-driven simulator, smart grid emulator is a continuous-time program. Here several existing methods in parallel and distributed simulation [5] cannot be directly used. To solve this problem, we adopt the mechanism designed in [6] and build a scheduler in the

federation component to sort the events in both components. Here we will run both components till the time that the first event will happen. Any new events that will be triggered by this event will then be inserted into the scheduler. In this way, synchronization between the two components will be maintained and both pre-scheduled and dynamic events can be properly handled.

To help instructors and students to ease the learning curve, we design a graphic user interface (GUI) to facilitate access to our smart grid emulator. As shown in Figure 2, our interface provides drag-and-drop functionality. A user can easily set up her/his simulation environment. A user can create links among the nodes and configure the properties of the links. For example, for network communication the user can choose the communication protocol and bandwidth. For a power line, the capacity of the line can be adjusted. This GUI provides an easy-to-use interface for users with little experience in smart grid and hides the system complexity from them. It is convenient for students to set up a smart grid simulation environment to run specified power grid models and obtain related power and communication security knowledge under this model.

2.1 User Scheduling in the Experiment Platform

Our system provides a reservation component through which a user can schedule her/his emulation experiment. The interface is shown in Figure 3. By clicking on an available time slot in the table, a user can reserve experiment time and identify configuration files. The configuration file specifies the grid topology information including nodes, power line, and communication link information. During the experiment, a log file will be generated that records the dynamic power flow and network information. After the experiment, the user can download the log file and analyze the experiment results.

Figure 3: Experiment platform reservation component.

3. Educational Materials and Example Experiments

3.1 Course Modules

We have designed several course modules on security of smart grid. Since they emphasize different aspects of this topic, they can be adopted by different courses. Below we introduce several modules in detail.

3.1.1 Overview of Smart Grid and Its Cyber Security

This is the introduction to the overall architecture of smart grid including both the physical power system and the communication system. For power system, we will introduce the functions of the energy generation components, the distribution and transmission mechanisms, and the load control and demand response algorithms. For the communication system, we will cover various communication networks in smart grid, the Advanced Metering Infrastructure (AMI), and major smart grid industry standards. The mutual impacts between power system and the communication system will be emphasized in the materials.

The cyber security of smart grid will provide high level discussion on the problems such as the attack surface in both sub-systems and threats from both inside and outside attackers. Since many traditional security measures such as authentication, authorization, and accounting demonstrate unique properties in smart grid, we will re-introduce these concepts based on the new application

environments. As a special emphasis, we will demonstrate how cyber attacks can lead to catastrophic results in physical power systems. Since this module provides a high level introduction to smart grid and its security, it can be integrated in most network security, electric systems, or infrastructure protection courses.

3.1.2 Network Security and Infrastructure Stability in Smart Gird

This module provides an in-depth coverage of the information network infrastructure in smart grid and its security. We will first describe the roles of wide area (WAN), local area (LAN), and home area networks (HAN) in smart grid. The advantages and disadvantages of different techniques to provide last-mile access connection to end users through power line communication, wireless networks, or cellular systems will be presented. Since many network attacks in Internet have their companions in smart grid, we will describe these kinds of attacks in detail. The power system needs to maintain a balance between the demands and supply in real-time. We will demonstrate how network attacks in the cyber system can impact the infrastructure stability of the whole power grids. This module will help students to connect information network security with power grid stability. It could be integrated into network security, power grid stability, or critical infrastructure protection courses.

3.1.3 Data Security and Privacy in Smart Grid

This module provides an in-depth coverage of the threats to confidentiality and privacy of data in smart grid. We will first describe the data collection, aggregation, processing, transmission, and storage procedures in the AMI and control systems of smart grid. Since different network protocols may be used at different stages of data processing, we will also discuss the vulnerabilities during the data format transformation procedures. In this module, we will introduce several concrete examples of data manipulation attacks [7][8] such as data injection [9][10]. Since data transmitted in smart grid without sanitization may lead to disclosure of sensitive information of end users, we will also introduce the countermeasures to preserve user privacy. This component can be

adopted by data security or critical infrastructure protection courses.

3.2 Example Experiments

We have designed a group of projects upon the proposed experiment platform. These experiments focus on the mutual impacts between the cyber and physical systems. In addition to the topics such as power system stability, we also consider attacks that may lead to financial misconduct. Below we describe a few examples in detail.

3.2.1 Experiment 1: Worm attacks on smart meters

This experiment will help students to understand how software diversity and network topology of AMI can improve the system's robustness against worm attacks. Students can assign different software versions to smart meters. They can also configure the connections among meters to form different network topologies such as an ad hoc network or a star-shaped hierarchy connected by data concentrators. A malicious node will send out packets to infect other smart meters. The infected nodes will then become new sources of worm attacks. Two rules will be applied to restrict the propagation of malware: (1) a malicious meter can infect only the nodes using the same version of software; and (2) since data signals cannot penetrate the voltage converters, the malware can infect only those meters connected by the same data concentrator. Students can examine worm propagation under different network topologies and their impacts on power grid stability. For example, if a large number of meters are infected, attackers can turn on/off power provision to these devices simultaneously. Fluctuations in power usage of this size are beyond the adjustment capabilities of the system and may cause cascading failures in the grid.

3.2.2 Experiment 2: DoS attacks on transaction control sublayer

Similar to the TCP/IP protocol stack, many control protocols in smart grid infrastructure also use 'transactions' to enable communication among meters, concentrators, and control nodes. Since most control networks need to handle only a limited number of event types, these communication protocols do not expect many

concurrent transactions. As an example, the LonTalk protocol [11] is widely used by many intelligent control devices including smart meters. In LonTalk, the maximum number of active transactions that a server needs to support is 16. This small number allows malicious nodes to conduct DoS attacks. In this experiment, students are required to initiate more than 16 transactions to saturate the service provider. Students will evaluate the impacts of the DoS attacks by measuring the success rate and average response delay of other requests.

3.2.3 Experiment 3: Impacts of power system topology errors on locational marginal price (LMP)

This experiment tries to expose the relationship between the power transmission network topology and real time electricity market prices. We consider the scenario in which the undetected false status of circuit breakers from topology error processing may lead to wrong modeling of real-time network topology, which, in turn, misleads the results of state estimation and real-time economic dispatch. To demonstrate the impacts of the errors, we focus on the changes in locational marginal price (LMP). Here an attacker can generate false data on status of circuit breakers and manipulate the sensed electricity flows along specific links. In this way, some links may be excluded from the state calculation procedures while the calculation results still roughly match with the sensed data. This error, however, could lead to changes in LMP through which the malicious party can make a profit.

The project uses the IEEE 14-bus system with the bus-breaker model, as shown in Figure 4.(a). Here the solid squares represent closed breakers, and hollow squares represent open breakers. The red line between 5 and 6 represents a congested line. In this figure, the misconfigured status of the circuit breaker at bus 5 leads to the (dotted) line 4-5 exclusion error as long as the line 5-6 is congested. The corrupted network topology information is fed into economic

dispatch module without being detected by topology error processing [12].

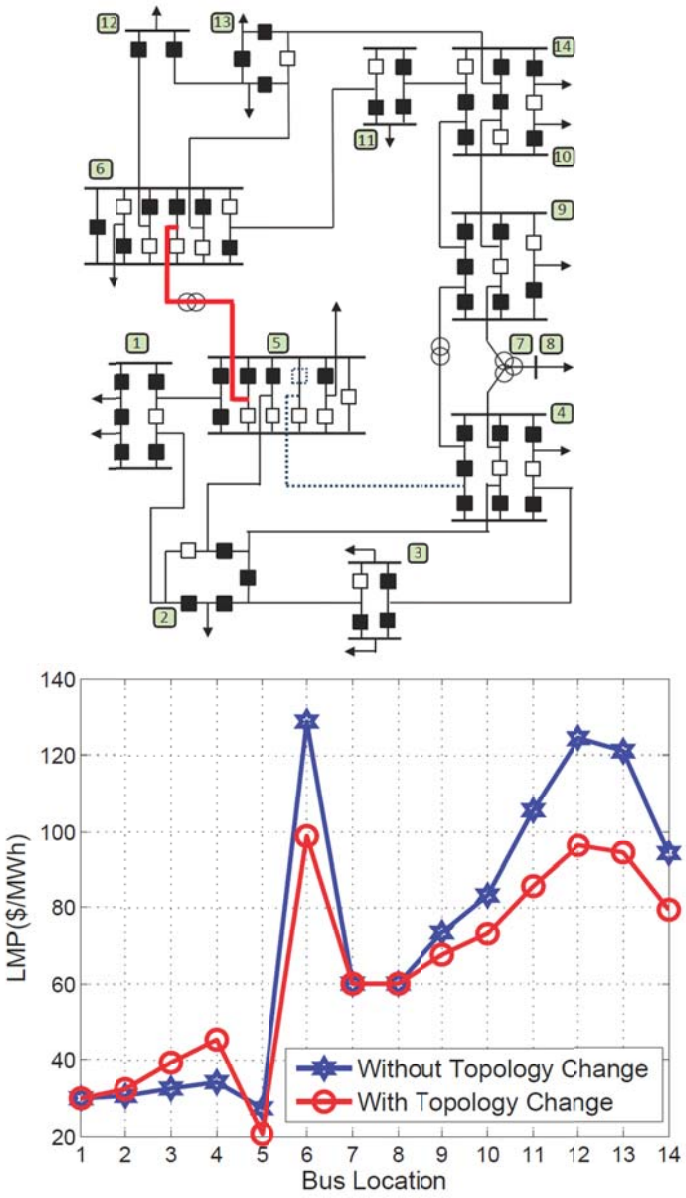

Figure 4: (a) IEEE 14-bus system with bus-broker model; and (b) the LMP results with and without line exclusion error.

In this project, students need to generate the network configuration. They will also generate the false information to mislead the calculation model. Our platform will then use the SCED (Security Constrained Economic Dispatch) model and the sensor data to calculate the LMP [12]. Figure 4.(b) shows the marginal price with and without the topology errors, respectively. Please note that under both cases the marginal units (at buses 1 and 8) and congestion pattern (the line 5-6 congestion) stay unchanged. In this way, it is very hard for the detection component to sense the anomaly. From this figure, students can learn that attackers can bypass the current anomaly detection components in SCADA system to manipulate the real-time power price and make a profit in a stealthy way.

4. Evaluation and Results

4.1 Material Adoption and Design of Evaluation

The proposed materials were adopted in two courses in the Fall semester of 2014. During the evaluation, repeated measured pre-post design was employed. Convenience sampling method was chosen because this serves as a pilot study of a large Smart Grid Project implemented at multiple institutions. The project was piloted in two classrooms at our university, one at the undergraduate level and the other at the graduate level. All students in the two classrooms participated in this project after the debriefing of the objectives of the project. The project was approved by the Institutional Review Board at our university and all student consent forms were signed and collected. All students completed pre- and post-surveys. Five randomly selected students from each classroom were interviewed for their experiences with this project after the instruction. Each interview lasted about 15-20 minutes and was recoded on a voice-recorder. The interviews were transcribed verbatim.

4.2 Sample

In the undergraduate classroom, 24 students participated in the project. Of these 24 students, 23 (96%) were male and one student (4%) was female. They are either junior (33%) or senior (67%)

students. As for ethnicity, 12 (50%) were European American, 4 (17%) were African American, 2 (8%) were Hispanic, 1 (4%) was Asian, 5 (21%) were multiple racial, and one student (4%) did not report this information. The graduate classroom was consisted of 39 students, of whom 17 (44%) were females and 22 (56%) were males. The ethnic background of these graduate students is as follows: 6 (15%) European American, 1 (3%) African American, 13 (33%) Asian, 10 (26%) multiracial, and 9 (23%) did not report this information. The average age of the undergraduate students was 23.68 with a standard deviation of 3.06 and that of the graduate students was 25.13 with a standard deviation of 3.74.

4.3 Survey Instruments

Self-Efficacy to Use Smart Grids and Security Knowledge was developed by the project director and the program evaluator to measure how well students feel they can reach the course objectives. Theories and practices in self-efficacy were adopted while designing the items [13][14][15]. It is consisted of 12 items on a 5-point Likert scale ranging from 1 (cannot do at all) to 5 (certainly can do). This survey was administered to the students before and after the instruction on the topics.

Self-Regulation for Smart Grids Security was also developed by the project director and the program evaluator to measure student self-regulated learning behaviors while learning the topics in the course. It is consisted of 13 items on a 5-point Likert scale ranging from 1 (not at all) to 5 (all the time). This survey was administered to the students before the instruction. It was not used as a repeated measure because we do not expect these behaviors to change in a short period of time. Student behavior is also not part of the goals of the project. It was used as a control variable in data analyses.

4.4 Data Analytical Procedure

Independent and dependent samples t-tests were used, respectively, to examine statistically significant differences of student perceptions of self-efficacy and their self-regulated learning behaviors between the undergraduate and graduate classrooms as well as before and after the project. An alpha-level of 0.05 was chosen for these hypotheses tests. Student responses to interview

questions were analyzed with grounded theory [16]. Constant comparison method was employed during the iterative process of comparing and contrasting themes and concepts. The circumstances under which these themes occurred were examined closely by at least two researchers to avoid researcher's bias.

4.5 Evaluation Results

Descriptive statistics of the measured variables (self-efficacy to use smart grid) before and after the instruction as well as student self-report of self-regulated learning behaviors are reported in Table 1 as follows:

Table 1

Means and Standard Deviations of Self-Efficacy and Self-Regulated Learning behaviors

	Efficacy (pre)	fficacy	egulation
graduate (n = 24)	).87)	).94)	).66)
ate (n = 39)	1.04)	).64)	).93)

Note. Numbers in parentheses are standard deviations.

Independent samples t-test failed to detect statistically significant differences between undergraduate and graduate students' self-report of the use of self-regulated learning behaviors, $t(61) = 1.13$, $p = .27$, Cohen's $d = 0.31$. This effect size is small [17]. Similarly, no statistically differences were found between undergraduate and graduate students with respect to their self-efficacy to work on Smart Grid projects either before, $t(61) < 0.01$, $p = .99$, Cohen's $d < 0.01$ (small effect size), or after the instruction, $t(61) = -0.55$, $p = .59$, Cohen's $d = 0.16$ (small effect size).

Dependent samples t-test revealed that student self-efficacy to use Smart Grid knowledge and skills increased significantly after the

instruction. Both the undergraduate and graduate students reported significant increase in self-efficacy beliefs after the instruction: $t(18) = 7.70$, $p < .001$, Cohen's $d = 1.77$ (large effect size) for the undergraduate students; and $t(27) = 9.03$, $p < .001$, Cohen's $d = 1.71$ (large effect size) for the graduate students.

Data from the interviews suggest that most students did not have any exposure to the topic of smart grid security before taking this course. Based upon the comparison of student responses with regard to what the students knew before and after the instruction, we can see that students did learn a lot from the instruction. A common theme that came out of the student responses is that the topic was covered too fast. One student said, "I would prefer he spends more time. Smart grid security was one topic which not too many people knew in the class out of 30 students. It was one or two students who knew about it. So not too many people knew, so maybe we could've gone a little slower." Another student said, "I think we got a really good overview of what it is but it's very difficult I think to understand different types of smart grids and what really could be done to prevent them because of the complexity of our grid system itself, taking the smart out of the power grid itself is a different field of computer science." Still another student said, "I think it was covered too fast because we did it on a single day and when he mentioned smart grid he was also mentioning other technology like big data so it wasn't like a separate topic to learn about 'oh what is this, what is this; what is the difference between this and this' and so everything was jammed packed towards the end of the semester."

Another theme that emerged from the qualitative data is more discussion and interactions. One student commented, "There's more of the lecture so maybe reduce the lecture, give breaks to ask questions, talk about similar topics and have more of a discussion, that would engage students if they are not listening. Because just listening you cannot fully concentrate, so breaks and having discussions we would be able to put in our mind." Another student shared the same point of view, "If we had the time I think we should have more practical labs or something for some of the things he talks about in class if we had more hands on things to do so we can really understand what he is teaching you in class."

However, not all students agreed to have more discussion or interactions, such as hands-on activities. One student argued, “An activity could be interesting like playing with a fake smart grid, it could be interesting but it would be taking away time to be covering material. So I would be hesitant to recommend that since that time could be to cover much more material.”

Finally, all the students interviewed were very positive towards the instructor. They think that the instructor delivered the knowledge clearly and effectively. According to one student, “He did a very good job of explaining it, I thought it was taught well, and he used very good drawings on the board. He used some actual pictures that were very high quality and he made it a little bit humorous and interesting, and so it was easy to follow what he was saying and what he was teaching.” Another student added that the instructor was humorous and was able to maintain his interest throughout the class: “Yes he was able to maintain my interest, it seemed a little off topic from the course but it was very interesting. He showed a video regarding a lab for smart grids that had like a fake windmill, a fake solar panel and stuff like that and I thought it looked fascinating and I thought he did a very good job of explaining the content and catching my attention.”

5 Conclusion

As an important component in Cyber-Physical systems, smart grid starts to attract more and more attention from both researchers and educators. The lack of security education materials for smart grid restricts the training of qualified workforce. In this project, we design multiple course modules to introduce infrastructure and data security in the systems. We have also designed an experiment platform upon which students can conduct simulation/emulation experiments. We describe a few experiments of cyber attacks on smart grid and their impacts on physical systems. We implemented the materials in both undergraduate and graduate courses. Evaluation and student interviews were conducted by education experts.

Immediate extensions to our project can be conducted in the following ways. First, we will implement these materials in collaborating institutions so that more students can benefit from

the project achievements. We will also evaluate the effectiveness of our approaches in a larger scale for further improvements. Second, we will refine the interface design of our experiment platform and make it public. In this way, other educators can start to use it and design new materials upon it. Last but not least, we will present our materials to our industrial partners and collect feedback from them. In this way, we can better train future workforce for them.

6 Acknowledgement

This research is supported in part by NSF Award #1303356, #1303359, and #1303378.

References

[1] J. Ahrenholz, “Comparison of CORE network emulation platforms”, in IEEE Military Communications Conference (MILCOM), pp. 166—171, October, 2010.

[2] T. Godfrey, M. Sara, R. C. Dugan, C. Rodine, D. W. Griffith, and N. T. Golmie, “Modeling smart grid applications with co-simulation”, in IEEE International Conference on Smart Grid Communications (SmartGridComm), pages 291–296, 2010.

[3] Hua Lin, S. Sambamoorthy, S. Shukla, J. Thorp, and L. Mili, “Power system and communication network co-simulation for smart grid applications”, in IEEE PES Innovative Smart Grid Technologies (ISGT), pages 1–6, 2011.

[4] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, “Matpower steady-state operations, planning and analysis tools for power systems research and education”, IEEE Transactions on Power Systems, 26(1):12–19, 2011.

[5] R. M. Fujimoto, “Parallel and Distributed Simulation Systems”, Wiley-Interscience, 2000.

[6] J. Nutaro, “Designing power system simulators for the smart grid: Combining controls, communications, and electro-mechanical dynamics”, in IEEE Power and Energy Society General Meeting, pages 1–5, 2011.

[7] S. Cui, Z. Han, S. Kar, T.T. Kim, H. V. Poor, A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions", IEEE Signal Processing Magazine, volume 29(5), pages 106-115, 2012.

[8] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao, "On false data injection attacks against distributed energy routing in smart grid", IEEE/ACM Third International Conference on Cyber-Physical Systems (ICCPS), pages 183-192, 2012.

[9] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets", in IEEE International Conference on Smart Grid Communications (SmartGridComm), pages 226-231, 2010.

[10] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations", IEEE Transactions on Smart Grid, volume 2(4), pages 659-666, 2011.

[11] Christine Simeone, "Echelon's control operating system (cos) software now enables fortum to reach beyond the smart meter," MSN Money Business Wire, 2011.

[12] Dae-Hyun Choi and Le Xie, "Impact analysis of locational marginal price subject to power system topology errors," Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on, pp. 55—60, 21-24 Oct. 2013.

[13] A, Bandura, "Self-efficacy: The exercise of control," New York: W. H. Freeman and Company, 1997.

[14] F. Pajares, "Motivational role of self-efficacy beliefs in self-regulated learning," in D.H. Schunk & B. J. Zimmerman (Eds.), Motivation and self-regulated learning: Theory, research, and applications. New York, NY: Routledge, 2009.

[15] C. Wang, D-H. Kim, R. Bai, and J. Hu, "Psychometric properties of a self-efficacy scale for English language learners in China," System, 44, 24-33, 2014.

[16] B. G. Glaser, "More grounded theory methodology: A reader," Mill Valley, CA: Sociology Press, 1994.

[17] J. Cohen, “Statistical power analysis for the behavioral sciences,” Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.