A Cyber Topology Model for the Texas 2000 Synthetic Electric Power Grid

Patrick Wlazlo[1], Kevin Price[1], Christian Veloz[1], Abhijeet Sahu[1],
Hao Huang[1], Ana Goulart[1], Katherine Davis[1], and Saman Zounouz[2]
[1]Texas A&M University
[2]Rutgers University

{[1]pjrwlazlo, kevincprice, diamond17, abhijeet_ntpc, hao_huang, goulart, katedavis}@tamu.edu,
[2]saman.zonouz@rutgers.edu

***Abstract*—Numerous works have modeled the synthetic power system network that represents the U.S. electricity transmission system. In this paper, the Texas 2000 synthetic bus model is adopted as a use case; however, the Texas 2000 synthetic bus model only contains the physical elements of the electric grid, such as power buses, substations, and transmission lines. Our research seeks to model the communication infrastructure for this synthetic grid to create a comprehensive cyber-physical model of the electric grid. Our communication model relies on the information gathered from the synthetic power system model as well as the network topologies used in substations, utility control centers, and balancing authorities like the Electricity Reliability Council of Texas (ERCOT). A Python program generates the communication model as a collection of JSON objects with lists of nodes and links, similar to the Cyber-Physical Topology Language (CPTL). The nodes represent the cyber components such as routers, firewalls, switches, relay controllers, relays, or remote terminal units (RTUs). The links represent the communication channel types such as microwave links, Ethernet, or MPLS/fiber links. Finally, the proposed communication model is applied to one region in the Texas 2000 synthetic model and Splunk is used to create the visualization for the cyber-physical model of this region.**

***Index Terms*—Smartgrid, Substation control network, Cyber-physical system, Network topology, Splunk**

I. Introduction

The electric grid consists of geographically separated components such as generators, power transformers, and distribution feeders [1], which can be modeled as a synthetic power system [2]. This synthetic system provides a research platform on power system analysis. For instance, the Texas 2000 bus synthetic grid model [3] contains publicly shared data of real power systems that are placed in artificial geographic locations. This synthetic bus model is used in power simulators, such as PowerWorld [4], to simulate the generation and transmission grid.

The communication network infrastructure also plays a vital role in the management, control, and security operations of the electric grid [5]. The inter-dependence between the *cyber* and the *physical* infrastructure makes it necessary to model a synthetic communication network to accompany the power system model. A realistic synthetic communication network model would enable industrial security analysts to test the resiliency of the electric grid to various cyber-attacks. For example, in the 2015 Ukraine power grid attack [6], hackers infiltrated multiple networks and remotely shut down substations to disrupt that country's energy distribution system.

Communication models can also be used as a platform to research and simulate algorithms for prevention of worm and malware propagation, for intrusion detection, in addition to training and educational purposes. With this in mind, the goal of our project is to combine the topologies of the physical and cyber systems as a cyber-physical model that can be applied in a next-generation Energy Management System (EMS). This next-generation EMS would help operators to visualize and assess cyber-threats and vulnerabilities in the power system.

However, realistic communication models are difficult to model as the actual network information deployed in the real world are considered to be highly sensitive information for the asset owners to release. Using public communication network models inside a substation [7][8], between substation and control center [9], and between control center and balancing authority or Independent System Operator (ISO) [10], we propose a communication network model that can be used with the Texas 2000 bus synthetic model. The communication model is represented in Java script object notation (JSON) format, and is based on the approach used in the Cyber Physical Topology Language (CPTL) [11], to describe cyber-physical architecture.

The remainder of this paper is organized as follows. Section II reviews previous works on cyber-physical modeling. Then, Section III presents our communication model, with details of its architecture and Python implementation based on an object-oriented approach in Section IV. A use case in a southern region of the Texas 2000 model is presented in Section V. Finally, future improvements to the model are discussed in Section VI with conclusions in Section VII.

978-1-7281-4201-2/19/$31.00 c 2019 IEEE

Accepted for publication in Principles, Systems and Applications of IP Telecommunications (IPTComm), October 2019

II. Background

A network topology model can be represented using a graph with vertices (nodes) and edges (links). However, to model a cyber-physical system, such as the smart grid, the inter-dependencies between the communication network and physical assets need to be fully captured in the model.

Numerous works have been dedicated to represent network topologies using Unified Modeling Language (UML) [12], eXtensible Markup Language (XML) [13], Java Script Object Notation (JSON) [14], Core Information Modeling (CIM) by Open Networking Foundation [15], GraphSpace [16], NetJSON [17], and models by Metro Ethernet Forum (MEF) [18]. Due to its widespread use, we adopted JSON to represent the proposed network topology, especially using the approach of representing nodes and links similar to the JSON format used by Grotto Networking [19].

In particular, the cyber-physical topology language (CPTL) [11] defines primitives, such as vertices and edges, that represent a cyber-physical network. For instance, the vertices are network nodes or physical devices, and the edges relate a pair of nodes, such as a communication link. Vertices and edges have attributes. A vertex that is a router could have attributes such as the IP address of its interfaces and its configuration information. Attributes of an edge could be for instance the protocol used in that link, such as DNP3, Modbus, or Inter-Control Center Communications Protocol (ICCP). One advantage of using CPTL is that it defines operations in the vertices and edges, such as contraction and expansion. We are interested in the contraction operation, which allows auditors or operators to assess and visualize a group of substations (e.g., by contracting several substations into a super-node). Vertex attributes can also be contracted, for instance, to assess a specific property of a group of relays.

Previous research works that focused on capturing the inter-dependencies among the cyber and physical side include the design of a synthetic communication model considering the Power Line Communication (PLC) in one utility [20]. A Common Information Model (CIM) is used in [21] to represent cyber information along with electrical information.

As a previous tool we expand upon, the Cyber-Physical Security Assessment Project (CyPSA) [5] provides a common format for future planning and risk assessment of the electric grid. CyPSA models the cyber-physical dependencies within the electric power grid, while adopting a common format using CPTL. It modeled the electric and communication systems of an 8-bus substation and a central control center. The 8-substation model included communication links such as Ethernet, serial, multimode fiber, etc and each node was assigned a type (e.g., a distance relay) and an IP addresses.

Our contribution is to leverage the idea of CPTL and the 8-substation model of CyPSA [5] to design a cyber model for a realistic Texas Synthetic Grid that mimics the communication infrastructure of substations, multiple utility control centers, and an actual balancing authority or Independent System Operator (ISO). The Python classes and JSON file can be used by other researchers in different fields of study, such as cyber-physical simulations of cyber attacks, improvements to human machine interface (HMI) at control centers, or developing intrusion detection tools for electric utilities.

III. Communication Model Design

The traditional power system model uses bus-branch topology to represent power grids. This simplifies the model for power system studies, at the expense of the detailed information of each substation. To build the communication model for power systems, it is necessary to consider the detailed substation topology. In this paper, we expand the bus-branch topology into node-breaker topology [5] to better represent the substation network.

With the substation topology, this paper presents the communication network from the device level to the utility control center (where the EMS is located), and from control center to balancing authority, as shown in Fig. 1. At the substations, in addition to equipment specific to electric utilities, we include generic network nodes to represent a corporate local area network (LAN). At the utility control center, a demilitarized zone (DMZ) with a public-facing web server has also been added.

A. Substation Network Topology

The substation information in the PowerWorld simulator contains zone name, bus number, substation number, and a substation's latitude and longitude. Using the bus numbers that are associated with a substation, we obtain information regarding the number of relays as well as type of relays (line and load relay considered in the PowerWorld case). Once the type of relay is determined, we add a link between the relay and the associated substation's relay controller. Substations are comprised of these nodes and links with the objective of transmitting data to the utility control center, usually using the Distributed Network Protocol (DNP3).

In conjunction with relays and relay controllers, miscellaneous network nodes are added to the cyber model of the substation to represent a LAN, which includes desktop computers for employers to access email, IP phones, security cameras, and card readers for physical access to the substation devices. These nodes are important to assess security vulnerability, since they are prone to cyber attacks (e.g. phishing emails).

B. Representing Multiple Utilities

In the Texas 2000 model, the power network topology is represented as a connected graph with n substations as nodes and m transmission lines as links. Using the substation's geographic coordinates, we created a k-mean area clustering [22] where k represents the number of utilities we use to distribute these substations. This k-mean area clustering of the substations also produces a set of central location coordinates. A geographical map showing the utilities connections to their substations is shown in Fig. 2, with seven utilities ($k = 7$). The central location set of coordinates is where we chose to place the utility control center.

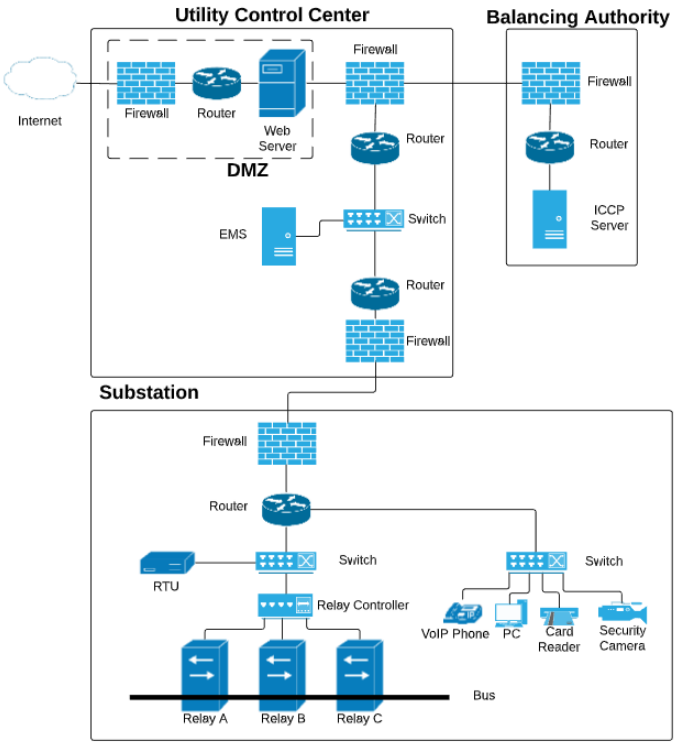

Fig. 1: Comprehensive network topology in the substation, utility control center, and the balancing authority.

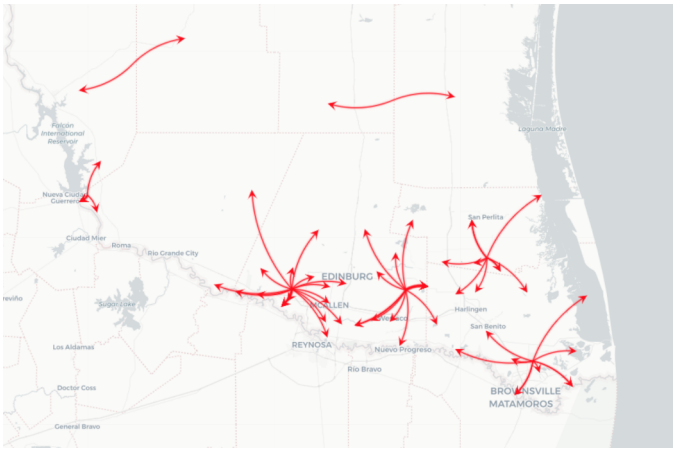

Fig. 2: Splunk map showing k-mean clustering of substations in the McAllen zone.

C. Substation to Utility Control Center

The utility control center is tasked with the job of monitoring and controlling their respective substations. Utilities contain supervisory systems such as supervisory control and data acquisition (SCADA) servers to manage their substation remotely and semi-autonomously. The modeling of how these systems relay information to each other will help us assess their vulnerabilities and risk of cyber attacks.

D. Communication between Utility and Balancing Authority

As an Independent System Operator (ISO) (or balancing authority), Electric Reliability Council of Texas (ERCOT) manages the flow of electric power on the Texas Interconnection providing electricity to 25 million customers. Its market participants are composed of Qualified Scheduling Entity (QSE), and Transmission and Distribution Service Provider (TSP/DSP). There are many utilities that act as QSEs and TSP/DSPs. Most of these utilities are located in a specific area in Texas, such as a municipality with substations nearby. Hence, we cluster the substations in the Texas 2000 Synthetic electric grid into multiple utilities. This allows us to represent multiple businesses in our model, where each utility has its own network, as described in Section III-B.

In particular, ERCOT has two control centers in Taylor, Texas (Primary) and Austin, Texas (Secondary). Each communicates with its market participant using Inter-Control Center Communications Protocol (ICCP) (i.e., Layer 7 protocol in the OSI architecture), Multiprotocol Label Switching (MPLS), Digital Access and Cross-connect System (DACS) network (as Layer 2 in the OSI architecture) [10]. DACS is usually used for voice communication and also as the backup communication in case of MPLS failure.

ERCOT provides the router, switch, and firewall that monitors the network traffic for each market participant. Therefore, each QSE, TSP, or DSP will have firewalls in the premises of each market participant as shown in Fig. 1.

E. Use of Different Types of Communication Links

The type of link employed for substation communication depends on various factors such as latency, geographic location of the substation (e.g., urban or remote areas), security, and cost.

Cellular technologies are used in inter-substation as well as substation to utility WAN communication [23]. Authors in [24] present the use of cellular-based sensor communication for overhead transmission line monitoring in power delivery systems. In addition, microwave point-to-point wireless links are common in substations located in rural areas.

Synchronous optical networking (SONET) is widely used in long distance communication based Wide Area Measurement (WAM) applications [25]. In [26], availability analysis on a SONET ring networks is performed for a power grid communications confirming the abundant use of fiber. Recently, the use of MPLS networks wherever available is also common for many fiber links between substation and utility control center. Additionally, legacy serial point-to-point links (over fiber or radio) are also common for wide area communication between utility and substations.

Power Line Communication (PLC) was an earlier method used for intra-substation communication that has been replaced by Ethernet and fiber technology [27]. In some scenarios, satellite communication is used as a backup communication infrastructure [28], although cellular links are more common nowadays as backup links to fiber optics links.

To represent the variety of types of communication links employed by utilities, the proposed cyber model includes link attributes such as the type of link (Ethernet, MPLS, cellular, serial), distance in meters, and bandwidth in bits per second. It is also useful to indicate which type of protocol (DNP3, ICCP, HTTP) is used in each communication link.

IV. JSON Object Generation

To build the model, initially the Python program extracts the power system information from an Excel file exported from a Texas 2000 model's PowerWorld simulation. Then, the application generates the communication model as a collection of JSON objects with lists of nodes, links, substations, utility control center, and a balancing authority. A hierarchical class model (Fig. 3) is chosen to represent the nodes within the Python code. Each class has unique attributes that allow the utility operator to distinguish it from the others. The main parts of this class include device attributes, such as node identifications (ID), utility, or IP address.

The *Node* and *Link* classes are used within a substation, utility, or balancing authority classes. These three classes are designed to be unique and not sub-classes of one another. This is the quickest way of indexing through the JSON data generated and classify the data into three domains (i.e., substation, utility control center, or balancing authority). However, by using this structure in the program it is important to make sure to have a common link between the substation and utility, and the utility and balancing authority. For this common link, we chose the link between two connected *Firewall* nodes, based on the network topology of Fig. 1. The node and link attributes are explained next.

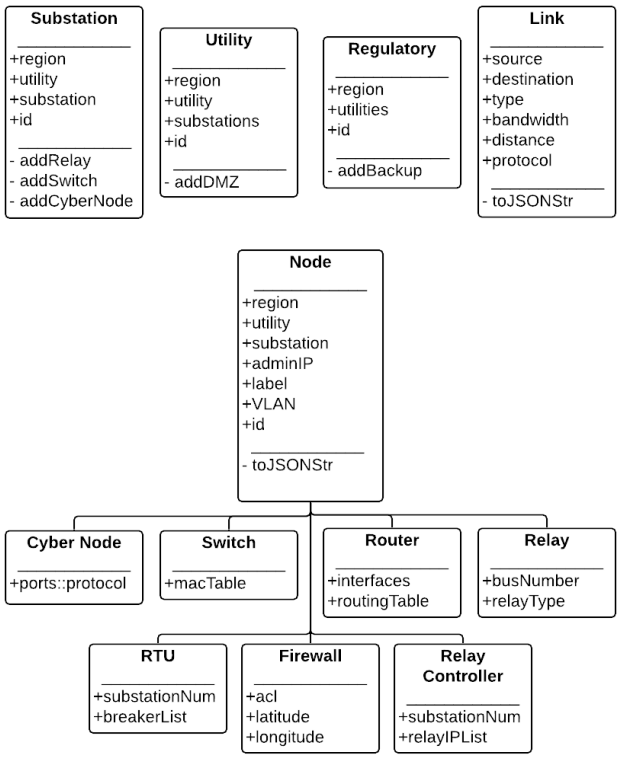

Fig. 3: Cyber model's Python classes and sub-classes with attributes.

A. Node Attributes

The *Node* parent class is created in Python to generate model for networking devices such as a firewall, router, switches, remote terminal units (RTUs), relay controllers, and relays (Fig. 3). The *Node* class stores attributes that every subclass of *Node* would contain, such as:

- **region** - ISO or balancing authority.
- **utility** - the utility provider that owns the substations.
- **substation** - the substation the device is operating in.
- **adminIP** - administrator IP address used for remotely configuring a device.
- **label** - naming convention for the device.
- **vlan** - the Virtual Local Area Network (VLAN) which the devices operate from within the utility.

Since nodes in the utility control center do not necessarily have a substation number, and nodes in the balancing authority do not have a utility number, the substation and utility attributes may be left empty by default.

Each sub-class of the *Node* class has further attributes that are specific for the device. For instance, the sub-class router have all the node attributes with the addition of the interface attribute (used to keep track of which IP address that are

assigned to a physical interface of a router), and the routing table.

B. Link Attributes

Similarly the interconnections of each node are stored in a Python *Link* class. The *Link* is not a sub-class of the *Node* class but its own parent class. The *Link* class attributes are as follows:

- **source** - source node ID.
- **destination** - destination node ID.
- **type** - serial, Ethernet, cellular, microwave, or MPLS/fiber.
- **bandwidth** - bandwidth in Mbps of the connection.
- **distance** - relative distance in meters of connection.
- **protocol** - main application layer protocol used in the link (DNP3, Modbus, or ICCP).

In this network model, the *links* are not directional since the *nodes* have full-duplex capability. However, the convention used for the links is to assign the parent node as the **source** and the child node as the **destination**.

Both the *Node* and *Link* classes store the attributes in a dictionary data type. The dictionary data type is used in conjunction with the Python JSON library [29] in order to generate the JSON data structure. The JSON data structure is generated and stored in a local file. It will then be utilized by Splunk [30] to index through the data and generate network topology maps. Further details are given in Section V.

C. Method of Generating Node ID

Each node has a unique ID attribute. This is required to map the node in any network topology application or plugin. If a node does not have a unique ID then the network topology application cannot index the node properly. In case of duplicate IDs the network topology application would likely ignore the second node with the same ID. We developed a regular expression to create the unique ID for each node. The regular expression used is as follows:

[region].[utility].[substation].[VLAN].[unique node no.]

where:

- the *region* is the ISO or balancing authority, for this instance "ERCOT";
- the *utility* is a k-mean clustering number from 0 to 6 (or a utility name);
- The *substation* is a substation number from the Texas 2000 bus model;
- The *VLAN* is used to distinguish between the different VLAN types, such as the corporate VLAN with cyber nodes (e.g., office computers, card reader, etc) and the operational VLAN with relays, controllers and RTUs;
- The *unique node no.* is the instance count of the node class.

While the unique node number can be in place of the unique ID attribute for a node, but it does not give any information about where the node is physically located within the network. The decision is made to include additional hierarchical information in order to make the debugging process simpler.

TABLE I: Table for IP Allocation

Location	Subnet	Network ID
Balancing Authority	/24	172.30.0.0
Utility WAN	/24	10.0.0.0-10.51.255.0
Utility DMZ	/27	10.0.0.0-10.51.255.0
Substation	/24	10.52.0.0-10.254.255.0

Furthermore, our unique node ID regular expression can be parsed in order to index and group nodes within the same level in the hierarchy.

D. Methods of IP allocation

The balancing authority is given a class B private IP address space with a /24 subnet mask (172.30.0.0/24). The /24 is used in order to allow for the addition of other regulatory subnets for future development of the model. Currently only a single /29 subnet is used to allow for six adminIP addresses in the nodes of the balancing authority, as in Fig. **??**.

The utility is also assigned a class A private address with a /24 subnet mask. Addresses from 10.0.0.0/24 to 10.51.255.0/24 are used. The mask is required since the utility will have multiple subnets. The main subnet will contain five nodes (a balancing authority side firewall and router, an EMS server, and a substation side firewall and a router), with a /27 subnet mask. The additional subnet will come from a demilitarized zone (DMZ) that will have two nodes. The DMZ subnet will also have a /27 subnet mask to allow for more servers to be placed within the DMZ subnet without the need for reallocating IP addresses.

Similarly, the substation will be given a class A private address space with a /24 subnet mask. Addresses from 10.52.0.0/24 to 10.254.255.0/24 are used. The substation will have multiple VLANs. The main VLAN is between the substation's firewall and router. The next VLAN is the operational technology (OT), which includes the relay controller, RTU, and relays. The other substation is the corporate VLAN which has a PC and VoIP Phone. Each of the VLANs listed prior will have a /27 subnet. Using only /27 subnets will allow for eight VLANs per substation, which leaves five extra /27 potential VLANs for more complex utility networks. The /27 subnets will allocate 30 host IP's for a single VLAN. Currently the largest VLAN has used only 10 out of the 30 host IP addresses for the OT VLAN with nine relays. Table I summarizes the broad overview of network ID and subnet allocated to different parts of the communication model.

E. Pseudo Code for Model Generation

The algorithm for cyber network model generation is given by Algorithm 1. It shows how the JSON files are used to store our model. The bottom-up approach outlined by the algorithm is based on the substation information imported from the Texas 2000 model built in PowerWorld. The utilities are generated by passing the substation's firewall node to ensure that both the utilities and substations files have a unifying link. Likewise, all the utilities firewalls are passed to generate regulatory domain

Algorithm 1 Pseudo code for network model generation

```
function Generate_Network_Model (Substation_list)
    . Substation_list is a list of substation numbers imported from PowerWorld
    for substation sub in Substation_list do
        Create instance of substation sub
        Populate instance sub with nodes and links
        Write nodes and links of sub into a JSON file
            . One JSON file is generated per substation
    end for
    Group substation instances into N areas (or utilities) using k-mean clustering based on latitude and longitude of substations . Consider N=7 for our model
    utility_list = Two dimensional list of substations separated by the k-mean area
        . Columns are the utility area number, rows are substation numbers
    for utility utl in utility_list do
        Create instance of utility utl
        Populate instance utl with nodes and links
        Create links to each substation in utl's area
        Write nodes and links of utl into a JSON file
            . One JSON file is generated per utility
    end for
    Create instance of regulatory, reg
    Populate reg with nodes and links
    Create links to each utility in utility_list
    Write nodes and links of reg into a JSON file
    . One JSON file is generated per balancing authority
return json_files
end function
```

(i.e., balancing authority). This construction of each domain allows for the option of merging all the nodes and links in our model into a single network topology map or just choosing to display a single domain's LAN.

V. Application to the Texas 2000 Bus Synthetic Model

The synthetic grid in the Texas 2000 model is divided into eight areas: south, south central, coastal, east, north, north central, west, and far west. Each area is further divided into multiple zones. We selected the McAllen zone from the south area for modeling purposes.

A. Splunk Map of McAllen Zone

Splunk is the primary framework used to generate maps for our project. More specifically, the Network Diagram Viz application in Splunk is utilized to display the network topology for JSON nodes and links. The Network Diagram Viz application is chosen since it is already integrated into the Splunk environment and Splunk is preferred due to its broad use in the cyber-physical security industry.

A generic substation cyber-physical model is shown in Fig. 4. The text under the node is the node ID. The node ID is unique to that node network location, as described in Section III. The figure serves as an example of what a substation from our model looks like when displayed using the Network Diagram Viz application.

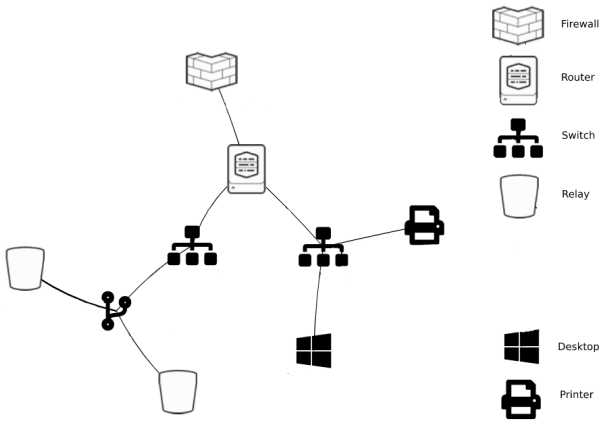

Fig. 4: Using Splunk map to display the network topology for a generic substation in the McAllen zone.

In the same way, Fig. 5 is showing the network connections from the utility control center to each of its substations. The connection from the substation routers to the utility router is based on a star configuration. The central node is the utility control center's firewall. There are five substation firewalls linked to the utilities firewall node. Behind each substation firewall there is a substation router, and behind the central firewall there is a router and an EMS server used to remotely monitor and control the substations. The application-layer protocol used in the communication link between utility and the five substations is DNP3. Attached above the EMS server is the balancing authority-side router, along with a firewall. This router and firewall are used to communicate with the balancing authority. The protocol used in that link is ICCP.

B. Evaluation of Computation Time for Cyber Model Generation Algorithm

For our model, the total processing time to output all 71 substations in McAllen zone to a file in the JSON format took $7.0338\ sec$. The average output time per file is $0.0991\ sec$. To output all seven of the utilities took $0.0274\ sec$. The average output time per file for a utility is $0.0037\ sec$. The total output time for all seven utilities does include the time to process the k-map clustering, which took $0.0012\ sec$ of computation time. To output the only balancing authority took $0.0005\ sec$. Run in unison, the whole process took $7.0617\ sec$.

The computation time is expected to grow exponentially when generating the entire Texas grid, because of the linear search algorithms used to index through the nodes in order to create links between them. Also, the linear search algorithm is used to import the PowerWorld data. More efficient non-linear search algorithms such as binary or interpolation search will be required to counteract the long computation time.

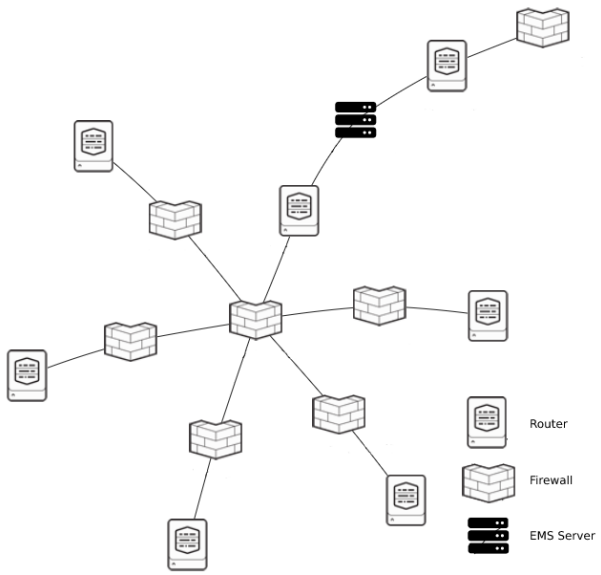

Fig. 5: Using Splunk to display a utility star topology.

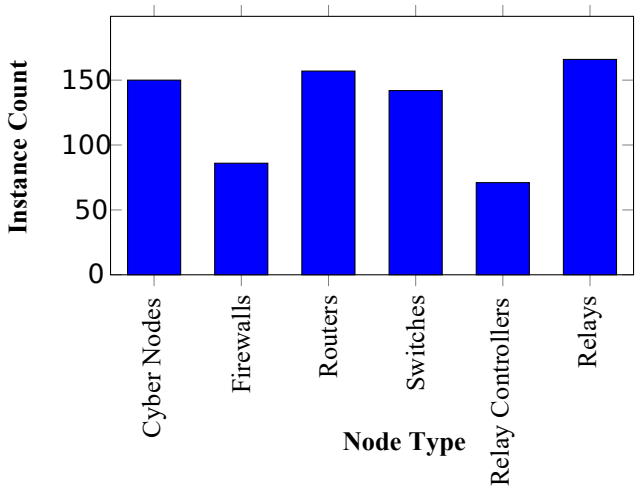

Fig. 6: Histogram showing the different amount of nodes in the McAllen zone.

C. Node and Links Size Evaluation for Cyber Model

For the communication model of the McAllen zone in the Texas 2000 model, links and nodes for 71 substations, 7 utilities, and 1 balancing authority are created. Between these three class domains, 150 generic cyber nodes (e.g., computers, printers, EMS servers) are generated. It created 142 switches and relay controllers nodes, as well as 86 firewall nodes, and 157 routers. There are also 166 load and line relays nodes (Fig. 6). This gives a total of 772 nodes and 456 links that are generated to create our cyber topology map of the McAllen zone. The cyber topology is stored within 79 JSON files.

Scaling our model to account for all of the Texas 2000 synthetic electric grid would roughly require a total of 14,000 nodes and 8,000 links, which would be stored in 1,400 JSON files. This is based on the fact that the Texas 2000 model has 1,250 substations and McAllen having only 71 substations. Note that this is a conservative estimate since the number of cyber nodes on the corporate WAN is going to increase as the model becomes more complex.

VI. Scope of Future Work

The communication model presented in this paper is developed for one zone in the south area of the Texas 2000 case (i.e. the McAllen region). This is a work in progress that will be further improved to accommodate the entire Texas grid. The model will have fine-grained attributes on every node and link, such as firewall configurations.

These are the key areas of improvements for the communication model:

- Modeling routing paths in the WAN, testing access control lists for firewalls, and simulating services running on the cyber hosts.
- Scaling our model to incorporate all eight areas of the 2000 Texas model.
- Improving visualization by merging power topology with cyber topology.
- Expanding our model to include some cases of inter-substation communication.
- Exploring the ability to expand and contract attributes of nodes and links, as proposed in CPTL, to give a concise assessment the cyber-physical system.
- Optimizing PowerWorld search algorithms for generating relay nodes and creating links inside substations.
- Parallel processing utilizing the parallelization package in Python to build and output the JSON for the full Texas grid. Parallel processing each area will reduce the total computation time.

VII. Conclusions

A synthetic communication network embedded on the physical grid makes the Texas 2000 model more comprehensive. Unlike the power network, the communication network is based on the location of the substations and the utility control centers. One balancing authority is also added to the cyber model. The current model stores data for the utilities' local and wide area networks in JSON format, as a network graph with nodes and links. Thus, the cyber topology model presented in this paper creates synthetic cyber assets on top of the existing physical buses of the Texas 2000 model. Up to this point the model is programmed in such a way that it can be implemented on any synthetic grid model, with the goal of creating a model that can simulate complex cyber attack scenarios on utilities.

Acknowledgement

This research is supported by the US Department of Energy Cybersecurity for Energy Delivery Systems program under award DE-OE0000895.

REFERENCES

[1] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "A survey on smart grid potential applications and communication requirements," *IEEE Transactions on industrial informatics*, vol. 9, no. 1, pp. 28–42, 2012.

[2] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on power systems*, vol. 32, no. 4, pp. 3258–3265, 2016.

[3] ACTIVSg2000: 2000-bus synthetic grid on footprint of Texas. [Online]. Available: https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg2000/

[4] "PowerWorld Simulator," 2019. [Online]. Available: https://www.powerworld.com/

[5] G. A. Weaver, K. Davis, C. M. Davis, E. J. Rogers, R. B. Bobba, S. Zonouz, R. Berthier, P. W. Sauer, and D. M. Nicol, "Cyber-physical models for power grid security analysis: 8-substation case," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2016, pp. 140–146.

[6] G. Liang, S. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 ukraine blackout: Implications for false data injection attacks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3317–3318, 2017.

[7] S. Coppel, T. Tibbals, and A. Silgardo, "Practical considerations for ethernet networking within substations," *Schweitzer Engineering Laboratories, Inc.(SEL), Energy Central*, 2008.

[8] T. Skeie, S. Johannessen, and C. Brunner, "Ethernet in substation automation," *IEEE control systems magazine*, vol. 22, no. 3, pp. 43–51, 2002.

[9] P. Robertson, C. Gordon, and S. Loo, "Implementing security for critical infrastructure wide-area networks," in *Proceedings of the Power and Energy Automation Conference, Spokane, WA, USA*, 2013, pp. 26–28.

[10] Ercot nodal iccp communication handbook. [Online]. Available: http://www.ercot.com/services/mdt/userguides/

[11] G. A. Weaver, C. Cheh, E. J. Rogers, W. H. Sanders, and D. Gammel, "Toward a cyber-physical topology language: Applications to nerc cip audit," in *Proceedings of the first ACM workshop on Smart energy grid security*. ACM, 2013, pp. 93–104.

[12] V. Saxena and D. Arora, "Uml modeling of network topologies for distributed computer system," *Journal of computing and information technology*, vol. 17, no. 4, pp. 327–334, 2009.

[13] A. Rahman, A. Pakstas, and F. Z. Wang, "An approach to integration of network design and simulation tools," in *Proceedings of the 8th International Conference on Telecommunications, 2005. ConTEL 2005.*, vol. 1. IEEE, 2005, pp. 173–180.

[14] L. Lhotka, "Json encoding of data modeled with yang," Tech. Rep., 2016.

[15] Core information modeling (coremodel). [Online]. Available: https://www.opennetworking.org/wp.../ONF-CIM_Core_Model_base_document_1.1.pdf

[16] Graphspace 2.0 user manual. [Online]. Available: http://manual.graphspace.org/en/latest/index.html

[17] Netjson. [Online]. Available: http://netjson.org

[18] M. S. M. 59. Network resource management information model: Connectivity. [Online]. Available: https://www.mef.net/Assets/Technical_Specifications/PDF/MEF_59.pdf

[19] Network graph visualization. [Online]. Available: https://www.grotto-networking.com/NetworkGraphVisualization.html

[20] T. Hartmann, F. Fouquet, J. Klein, Y. Le Traon, A. Pelov, L. Toutain, and T. Ropitault, "Generating realistic smart grid communication topologies based on real-data," in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2014, pp. 428–433.

[21] P. Skare, H. Falk, M. Rice, and J. Winkel, "In the face of cybersecurity: How the common information model can be used," *IEEE Power and Energy Magazine*, vol. 14, no. 1, pp. 94–104, 2015.

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[23] A. Hajjawi, M. Ismail, and N. F. Abdullah, "A scheduling scheme for smart grid and mobile users over lte networks," in *2016 International Conference on Advances in Electrical, Electronic and Systems Engineering (ICAEES)*. IEEE, 2016, pp. 421–426.

[24] K. Hung, W. Lee, V. Li, K. Lui, P. Pong, K. Wong, G. Yang, and J. Zhong, "On wireless sensors communication for overhead transmission line monitoring in power delivery systems," in *2010 First IEEE International Conference on Smart Grid Communications*. IEEE, 2010, pp. 309–314.

[25] Y. Deng, H. Lin, A. G. Phadke, S. Shukla, J. S. Thorp, and L. Mili, "Communication network modeling and simulation for wide area measurement applications," in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*. IEEE, 2012, pp. 1–6.

[26] S. Jih and M.-L. Yin, "An availability analysis on sonet ring networks in power grid communications," in *2012 Proceedings Annual Reliability and Maintainability Symposium*. IEEE, 2012, pp. 1–6.

[27] R. Jenkins, D. Dolezilek, D. Darms, and B. Gurney, "Case study: Efficiently replace plc automation systems by integrating ieds with fiber-optic and ethernet communications in the substation," in *2006 Power Systems Conference: Advanced Metering, Protection, Control, Communication, and Distributed Resources*. IEEE, 2006, pp. 408–422.

[28] X. Xiao, Z. Fu, G. Liu, and C. Deng, "A backup data network for power system automations based on satellite communication," in *2010 International Conference on Power System Technology*. IEEE, 2010, pp. 1–5.

[29] Json encoder and decoder. [Online]. Available: https://docs.python.org/3/library/json.html

[30] P. R. J. D. M. J. D. Erickson Delgado, Ashish Kumar Tulsiram Yadav and B. page Sigman, *Splunk: Enterprise Operational Intelligence Delievered*. Packt Publishing, 2017.