Analytics for Smart Grid Cybersecurity

Nazli Choucri
Professor
Political Science Department
Massachusetts Institute of Technology

Gaurav Agarwal
Alumnus
Sloan School of Management
Massachusetts Institute of Technology

April 25, 2017

Abstract

Guidelines, directives, and policy statements are usually presented in "linear" text form – word after word, page after page. However necessary, this practice impedes full understanding, obscures feedback dynamics, hides mutual dependencies and cascading effects and the like, -- even when augmented with tables and diagrams. The net result is often a checklist response as an end in itself. All this creates barriers to intended realization of guidelines and undermines potential effectiveness. We present a solution strategy using text as "data", transforming text into a structured model, and generate a network views of the text(s), that we then can use for vulnerability mapping, risk assessments and control point analysis. We apply this approach using two NIST reports on cybersecurity of smart grid, more than 600 pages of text. Here we provide a synopsis of approach, methods, and tools. (Elsewhere we consider (a) system-wide level, (b) aviation e-landscape, (c) electric vehicles, and (d) SCADA for smart grid).

Keywords

Cyber-physical systems, national security, smart grids, system of systems.

Citation: Choucri, N., & Agarwal, G. (2017). Analytics for smart grid cybersecurity. *Proceedings of the 2017 IEEE International Symposium on Technologies for Homeland Security (HST)*, 1–3.

Unique Resource Identifier: https://doi.org/10.1109/THS.2017.7943512

Analytics for Smart Grid Cybersecurity

1 The Challenge

As a general practice, guidelines, directives and policy documents are presented in text form, page-by-page and word-by-word – supported with figures, diagrams and tables as needed. Rooted in the legal tradition, this practice reinforces a linear logic, where sequence dominates, and the focus is on compliance, step by step. Invariably this situation supports a checklist approach to meeting requirements. By definition text undermines any attention to feedback, delays, interconnections, cascading effects, indirect impacts and the like – all embedded deep into the idiom or structure of the textual form. The text-form may be necessary, but it is not sufficient. In fact, it may create barriers to understanding, obscure the full nature of directives, and generate less than optimal results – all of which impede the pursuit of effective outcomes. Table 1 outlines the opportunity costs of cybersecurity guidelines.

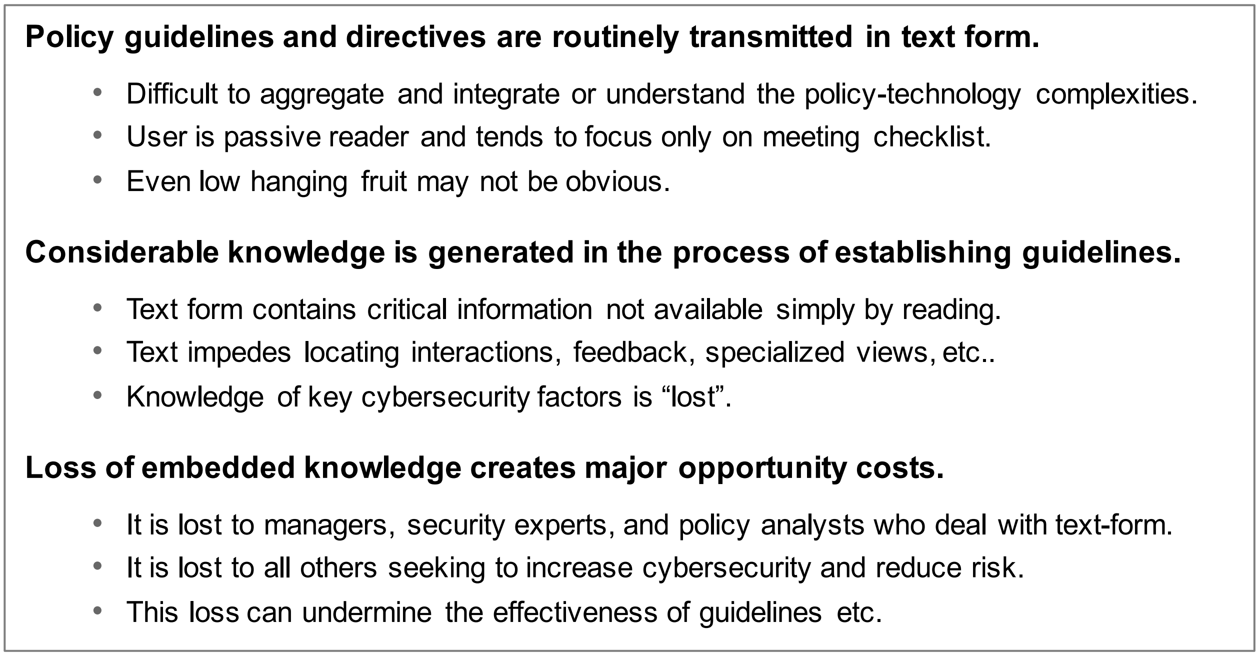
2 The Strategy

Focusing on the salience of cybersecurity in both private and public sectors, we draw on major reports presented by the National Institute for Standards and Technology (NIST) in its efforts to improve cybersecurity by providing analyses system state, risk assessments, probability metrics, all with detailed annotations to help guide the user community.

This material is rich in content, based on considerable collective knowledge, and subjected to a careful scrutiny and evaluation. While some efforts [1-4] have already been made to mine NIST materials, few exploit [5-8] the value of multi-analysis for knowledge mining and analytical tools to support user understanding, analysis, and eventually action. [5-6] visualizes the information on smart grid conceptual model provided in [9]. [7] analyses the dependencies within the same conceptual model and then restructures the later. [8] provides a filtered view of the conceptual model for electric vehicle.

This paper presents a comprehensive solution strategy to overcome the limitations of the text based view of guidelines and directives, as outlined in **Table 1**. The strategy consists of deploying analytical tools to formal text for the purpose of capturing as much of the features and intents of policies, guidelines, and directives as possible. **Table 2** provides a research summary with implementation details provided in the next section.

Table 1: ***Incidence-view*** **(binary) by cell in DSM for** ***Rules*** **in** ***Tallinn Manual 2.0***



Policy guidelines and directives are routinely transmitted in text form.

- Difficult to aggregate and integrate or understand the policy-technology complexities.
- User is passive reader and tends to focus only on meeting checklist.
- Even low hanging fruit may not be obvious.

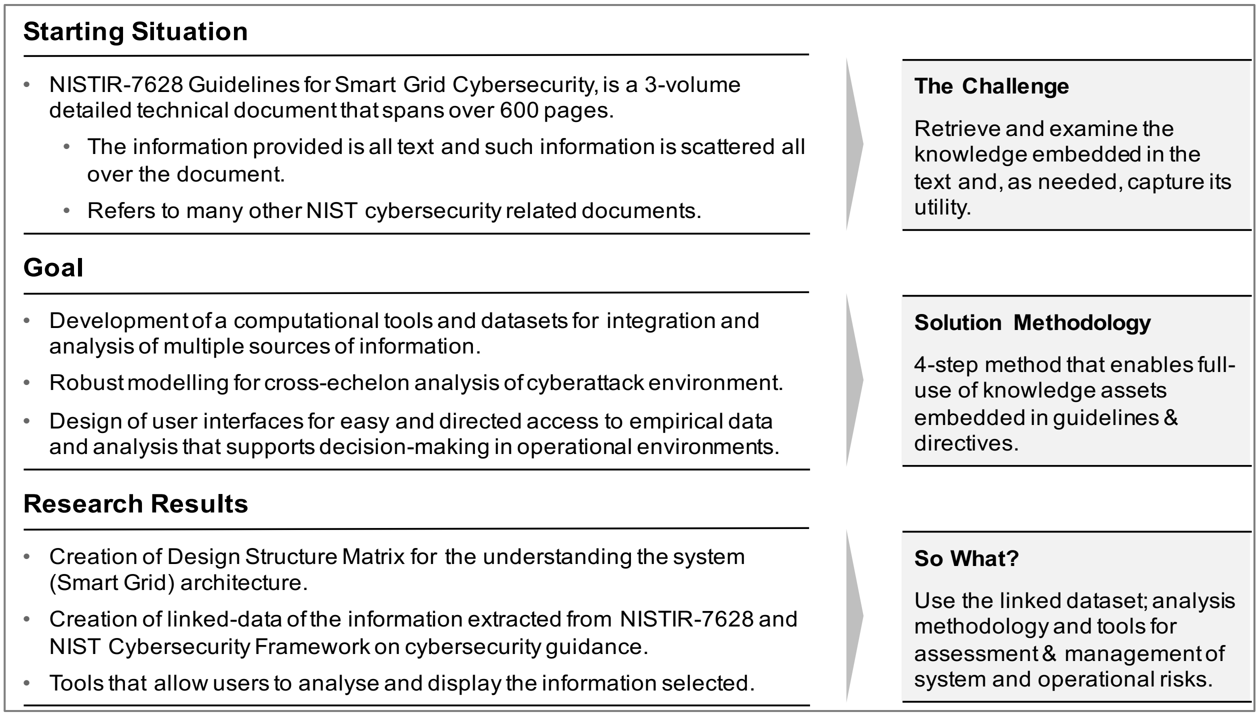
Considerable knowledge is generated in the process of establishing guidelines.

- Text form contains critical information not available simply by reading.
- Text impedes locating interactions, feedback, specialized views, etc..
- Knowledge of key cybersecurity factors is "lost".

Loss of embedded knowledge creates major opportunity costs.

- It is lost to managers, security experts, and policy analysts who deal with text-form.
- It is lost to all others seeking to increase cybersecurity and reduce risk.
- This loss can undermine the effectiveness of guidelines etc.

Table 1 Research Summary



Starting Situation

- NISTIR-7628 Guidelines for Smart Grid Cybersecurity, is a 3-volume detailed technical document that spans over 600 pages.
  - The information provided is all text and such information is scattered all over the document.
  - Refers to many other NIST cybersecurity related documents.

The Challenge

Retrieve and examine the knowledge embedded in the text and, as needed, capture its utility.

Goal

- Development of a computational tools and datasets for integration and analysis of multiple sources of information.
- Robust modelling for cross-echelon analysis of cyberattack environment.
- Design of user interfaces for easy and directed access to empirical data and analysis that supports decision-making in operational environments.

Solution Methodology

4-step method that enables full-use of knowledge assets embedded in guidelines & directives.

Research Results

- Creation of Design Structure Matrix for the understanding the system (Smart Grid) architecture.
- Creation of linked-data of the information extracted from NISTIR-7628 and NIST Cybersecurity Framework on cybersecurity guidance.
- Tools that allow users to analyse and display the information selected.

So What?

Use the linked dataset; analysis methodology and tools for assessment & management of system and operational risks.

3 The Implementation

Almost everyone recognizes salience of cybersecurity as a fundamental requisite for socio-economic and polity stability and wellbeing. Reports of serious breaches of established practice in terms of unauthorized access, damages to data and systems, deployment of malware, outright theft, invasion of privacy and a host of rapidly growing disruptions -- to note some of the most recurrent themes – all of which have created a vocabulary that expands day by day. This study focuses on key NIST reports on cybersecurity for smart grid, a ubiquitous feature of power systems.

There is little need for introduction of NIST, the premier standard setting entity in the nation and often for the international community as a whole. In this study, we go beyond appreciating the contributions of NIST to viewing reports as a source of new knowledge, a basis for identifying risk, valuating alternative courses of action, and facilitating prioritization in the deployment of corrective measures.

Here we illustrate our approach with the use of two key documents from the overall NIST ecosystem–NISTIR-7628: Guidelines for Smart Grid Cybersecurity [9], and (ii) NIST Cybersecurity Framework [10] —all totaling more than 600 pages. We use the NISTIR-7628 Guidelines as the basis and augment our investigations with the Framework. We consider these as distinctive generic meta-representations of system-state and risk assessments. Rather than evoking the "one size fits all" idiom, NIST highlights the necessary as well as the sufficient.

Table 2 Pragmatics for Cybersecurity and Risk Management.

A	**Create Linked Data for Structured Model of Smart Gird**	Identify essential system elements designed to fulfil intended functions of a Smart Grid and create a linked database.	***Sources:*** NISTIR 7628 - *Guidelines for Smart Grid Cybersecurity;* and *NIST Cybersecurity Framework.* ***Tools:*** *Relational Database.*
B	**Construct Design Structure Matrix & Exploratory Tools**	Use linked data base to build Design Structure Matrix (DSM) of essential elements of Smart Grid (actors, domains, etc. per NIST frame).	***Source:*** Relational Database created in step (A) above. ***Tools:*** *Excel for DSMs; Tableau for exploratory tool; Protégé for Hypertext views.*
C	**Generate system-wide Network View**	Create network view from DSM model to examine dependencies among system elements, control points, salience of edges, & implications system-wide implications of guidelines.	***Source:*** Relational Database created in steps (A) and (B) above. ***Tools:*** *Gephi for creating network visualizations.*
D	**Identify Enterprise Risks & Assessments**	Utilize exploratory tools, databases and network views to situate vulnerabilities of system elements and analyse system-wide impacts on the smart grid using network views.	***Additional Sources:*** *NIST National Vulnerability Database; NIST Common Vulnerability Scoring System.* *Tools: DSMs (Excel); Exploratory Tools (Tableau) ; Hypertext (Protégé) and Network visualizations (Gephi).*

Simply put, the first step is to transform the basic text (NISTIR 7628: Guidelines for Smart Grid Cybersecurity) into a structured model, a design structure matrix, of the entire system in question. Overall approach is presented in **Table 3** for enabling full-use of knowledge assets embedded in guidelines & directives.

Figure 1 shows the NIST logical reference model of Smart Grid. Figure shows the actors, the domains and interfaces between them, albeit as a "spaghetti plate". (Color represents the Smart Grid sub-domain). The key elements are:

- Actor "...is a device, computer system, software program, or the individual or organization that participates in the smart grid"[9].
- Domains encompass smart grid conceptual roles and services
- Logical Interfaces connect any two actors.

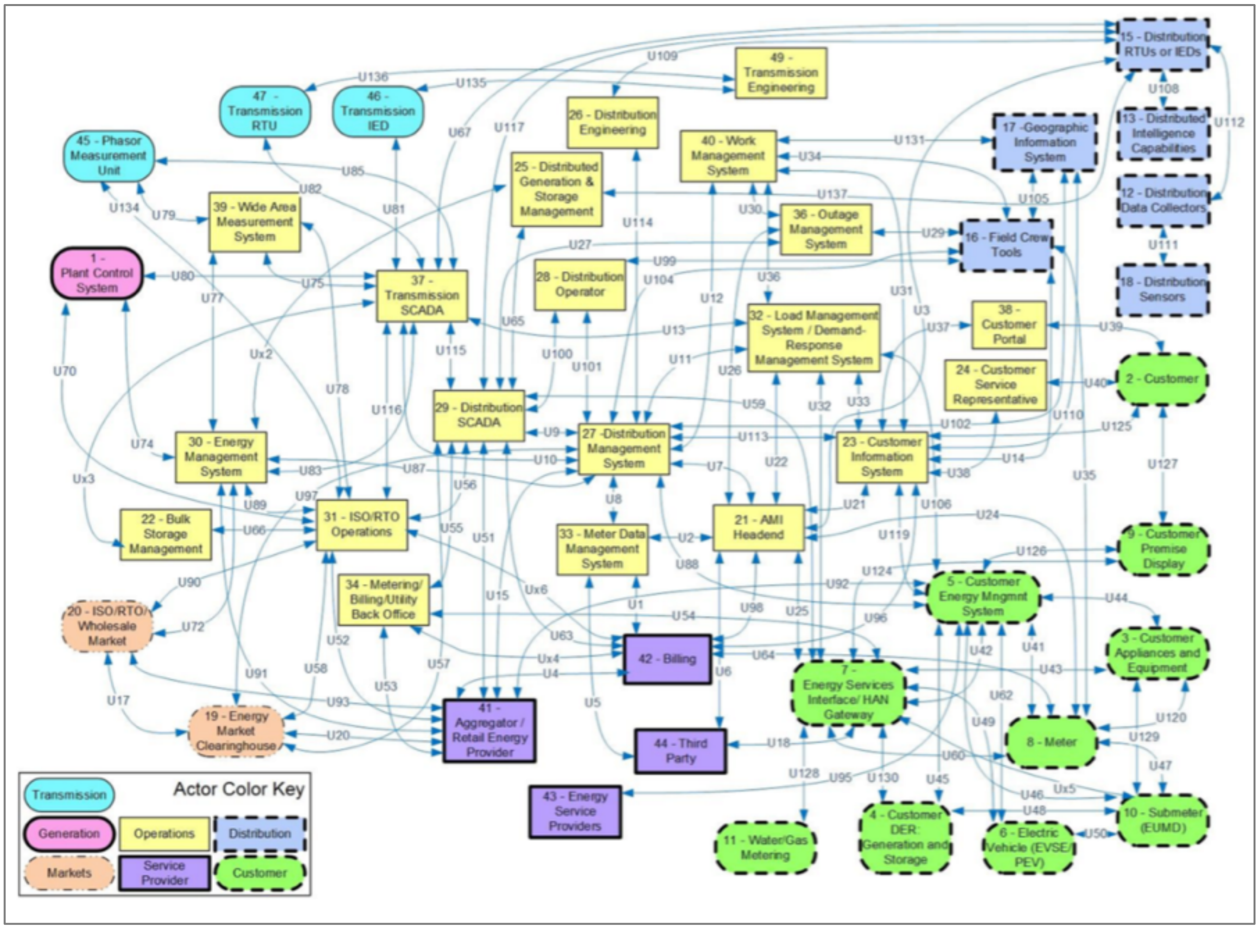
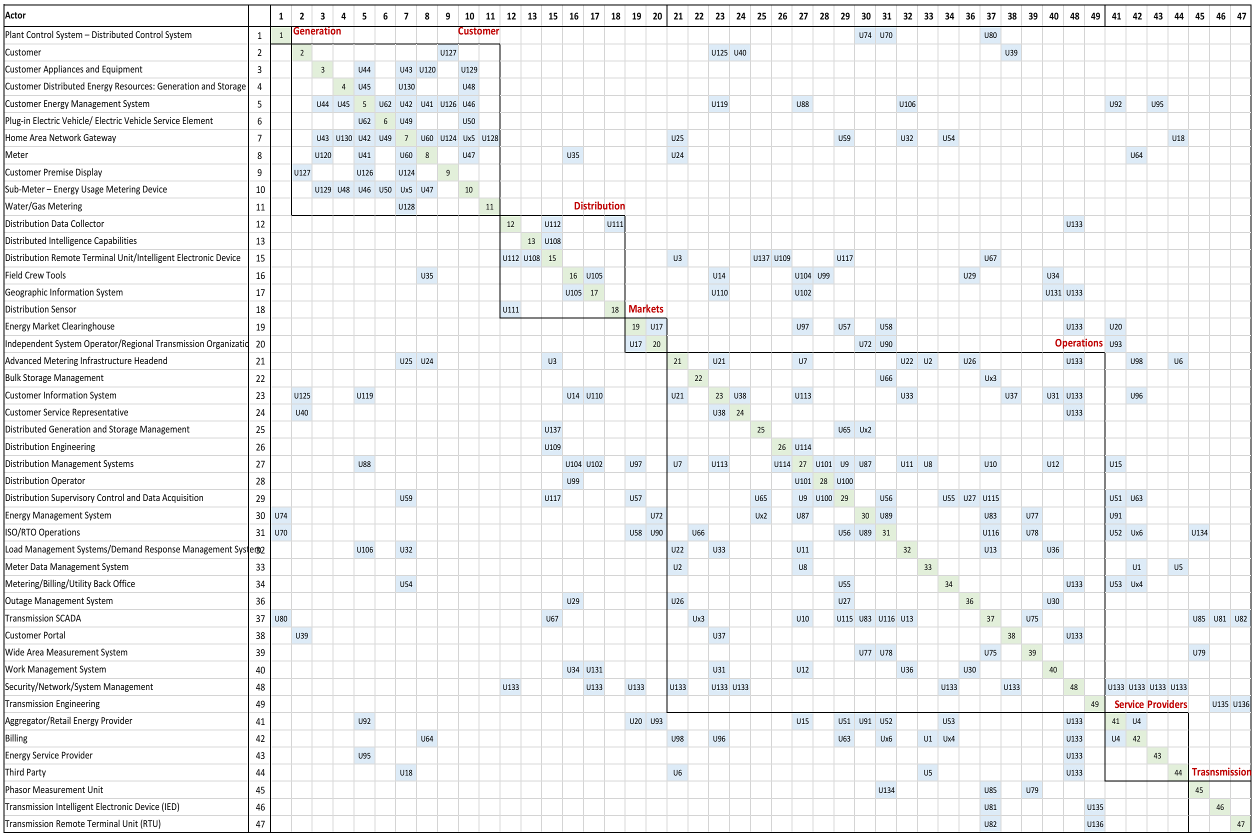

15 - Distribution RTUs or IEDs
49 - Transmission Engineering
47 - Transmission RTU
46 - Transmission IED
45 - Phasor Measurement Unit
26 - Distribution Engineering
40 - Work Management System
17 - Geographic Information System
13 - Distributed Intelligence Capabilities
12 - Distribution Data Collectors
25 - Distributed Generation & Storage Management
39 - Wide Area Measurement System
36 - Outage Management System
16 - Field Crew Tools
1 - Plant Control System
37 - Transmission SCADA
28 - Distribution Operator
18 - Distribution Sensors
32 - Load Management System / Demand-Response Management System
38 - Customer Portal
24 - Customer Service Representative
2 - Customer
29 - Distribution SCADA
27 -Distribution Management System
30 - Energy Management System
23 - Customer Information System
22 - Bulk Storage Management
31 - ISO/RTO Operations
33 - Meter Data Management System
21 - AMI Headend
9 - Customer Premise Display
34 - Metering/ Billing/Utility Back Office
5 - Customer Energy Mngmnt System
20 - ISO/RTO Wholesale Market
42 - Billing
3 - Customer Appliances and Equipment
7 - Energy Services Interface/ HAN Gateway
19 - Energy Market Clearinghouse
41 - Aggregator / Retail Energy Provider
44 - Third Party
8 - Meter
43 - Energy Service Providers
11 - Water/Gas Metering
4 - Customer DER Generation and Storage
6 - Electric Vehicle (EVSE/ PEV)
10 - Submeter (EUMD)
Actor Color Key
Transmission
Generation
Operations
Distribution
Markets
Service Provider
Customer

Figure 1. Smart Grid Logical Reference Model.
Source: NISTIR-7628 [6]

Using the NIST "reference" model from NISTIR 7628 [9], as an entry point, the critical data are extracted from the NIST tables that define the elements of the reference model (Figure 3). The actors and logical interfaces identified in **Figure 2** are converted into a Dependency Structure Matrix (DSM). First proposed by Steward [11], DSM is a matrix-based information exchange tool for representation of interactions between the elements or activities of a decomposed system [12].

The full DSM matrix is shown in **Figure 2** at a high level of abstraction. The DSM depicts actors listed on the rows and columns in the order numbered in the "reference" model. Entries in the cell correspond to logical interfaces signaled in **Figure 2**. We can "dig in deeper" or to focus on some segment(s). Traditional exploratory tools allow us to examine parts and pieces of the DSM as done in [8].



Actor		1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	36	37	38	39	40	48	49	41	42	43	44	45	46	47
Plant Control System – Distributed Control System	1	1	Generation								Customer																			U74	U70					U80												
Customer	2		2							U127													U125	U40													U39											
Customer Appliances and Equipment	3			3		U44		U43	U120		U129																																					
Customer Distributed Energy Resources: Generation and Storage	4				4	U45		U130			U48																																					
Customer Energy Management System	5			U44	U45	5	U62	U42	U41	U126	U46												U119				U88					U106										U92		U95				
Plug-in Electric Vehicle/ Electric Vehicle Service Element	6					U62	6	U49			U50																																					
Home Area Network Gateway	7			U43	U130	U42	U49	7	U60	U124	Ux5	U128									U25								U59			U32		U54											U18			
Meter	8			U120		U41		U60	8		U47					U35					U24																						U64					
Customer Premise Display	9		U127			U126		U124		9																																						
Sub-Meter – Energy Usage Metering Device	10			U129	U48	U46	U50	Ux5	U47		10																																					
Water/Gas Metering	11							U128				11				Distribution																																
Distribution Data Collector	12												12		U112			U111																						U133								
Distributed Intelligence Capabilities	13													13	U108																																	
Distribution Remote Terminal Unit/Intelligent Electronic Device	15												U112	U108	15						U3				U137	U109			U117							U67												
Field Crew Tools	16								U35							16	U105						U14				U104	U99							U29				U34									
Geographic Information System	17															U105	17						U110				U102												U131	U133								
Distribution Sensor	18												U111					18	Markets																													
Energy Market Clearinghouse	19																		19	U17							U97		U57		U58									U133		U20						
Independent System Operator/Regional Transmission Organizatic	20																		U17	20										U72	U90								Operations			U93						
Advanced Metering Infrastructure Headend	21							U25	U24						U3						21		U21				U7					U22	U2		U26					U133			U98		U6			
Bulk Storage Management	22																					22									U66					Ux3												
Customer Information System	23		U125			U119										U14	U110				U21		23	U38			U113					U33					U37		U31	U133			U96					
Customer Service Representative	24		U40																				U38	24																U133								
Distributed Generation and Storage Management	25														U137										25				U65	Ux2																		
Distribution Engineering	26														U109											26	U114																					
Distribution Management Systems	27					U88										U104	U102		U97		U7		U113			U114	27	U101	U9	U87		U11	U8			U10			U12			U15						
Distribution Operator	28															U99											U101	28	U100																			
Distribution Supervisory Control and Data Acquisition	29							U59							U117				U57						U65		U9	U100	29		U56			U55	U27	U115						U51	U63					
Energy Management System	30	U74																		U72					Ux2		U87			30	U89					U83		U77				U91						
ISO/RTO Operations	31	U70																	U58	U90		U66							U56	U89	31					U116		U78				U52	Ux6			U134		
Load Management Systems/Demand Response Management System	32					U106		U32													U22		U33				U11					32				U13			U36									
Meter Data Management System	33																				U2						U8						33										U1		U5			
Metering/Billing/Utility Back Office	34							U54																					U55					34						U133		U53	Ux4					
Outage Management System	36															U29					U26								U27						36				U30									
Transmission SCADA	37	U80													U67							Ux3					U10		U115	U83	U116	U13				37		U75								U85	U81	U82
Customer Portal	38		U39																				U37														38			U133								
Wide Area Measurement System	39																													U77	U78					U75		39								U79		
Work Management System	40															U34	U131						U31				U12					U36			U30				40									
Security/Network/System Management	48												U133				U133		U133		U133		U133	U133										U133			U133			48		U133	U133	U133	U133			
Transmission Engineering	49																																								49	Service Providers					U135	U136
Aggregator/Retail Energy Provider	41					U92													U20	U93							U15		U51	U91	U52			U53						U133		41	U4					
Billing	42								U64												U98		U96						U63		Ux6		U1	Ux4						U133		U4	42					
Energy Service Provider	43					U95																																		U133				43				
Third Party	44							U18													U6												U5							U133					44	Trasnsmission		
Phasor Measurement Unit	45																														U134					U85		U79								45		
Transmission Intelligent Electronic Device (IED)	46																																			U81					U135						46	
Transmission Remote Terminal Unit (RTU)	47																																			U82					U136							47

Figure 2 Design structure matrix of NIST Smart Grid "reference" model.

4 Network Views

With the foundations in place the next step is to create the network view of the smart grid. A combination of graph theory and its underlying tools [13-14] help transform the DSM into a very informative and valuable network view. **Figure 3** shows the network view of NIST information flows based on the DSM. The network view and the DSM, based on exactly the same data differ considerably in explanatory value.

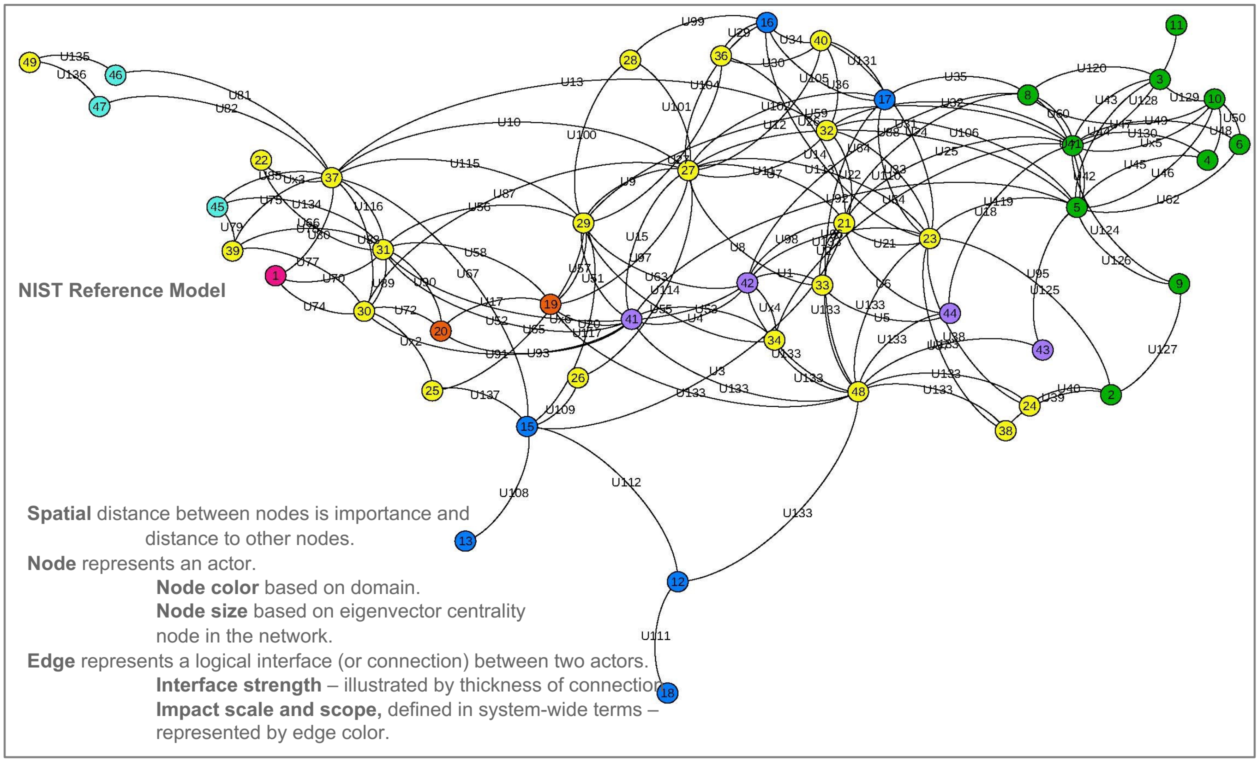

NIST Reference Model
Spatial distance between nodes is importance and distance to other nodes.
Node represents an actor.
Node color based on domain.
Node size based on eigenvector centrality node in the network.
Edge represents a logical interface (or connection) between two actors.
Interface strength – illustrated by thickness of connection
Impact scale and scope, defined in system-wide terms – represented by edge color.

Figure 3 Alternate view of NIST Smart Grid "reference" model.

5 What Have We Gained?

The network view (Figure 3), is built on the DSM (**Figure 2**), which is derived from information in a text on guidelines and directives for cybersecurity– in this case focusing for Smart Grid – and highlighted in Figure 1 Based on centrality criteria [14-15], **Figure 3** can then be used to locate nodes central to the Smart Grid, identify vulnerability points or pathways, and explore a range of "what …if…." contingencies.

Here we presented a snapshot of our purpose, as well as methods and tools to capture "full value" embedded in guidelines and directives for cybersecurity and sustainability. Our approach is generic, and can be customized to specific enterprise, operation, mission or other – text permitting.

References

[1] Cyber Security Procurement Requirements Traceability for the Electric Sector, Electric Power Research Institute, Palo Alto, CA, Tech. Rep. 3002003331, 2014.

[2] Security Posture using the Electricity Subsector Cybersecurity Capability Maturity Model (ES-C2M2), Electric Power Research Institute, Palo Alto, CA, Tech. Rep. 3002003332, 2014.

[3] Risk Management in Practice, Electric Power Research Institute, Palo Alto, CA, Tech. Rep. 3002003333, 2014.

[4] Cyber Security Risk Management in Practice, Electric Power Research Institute, Palo Alto, CA, Tech. Rep. 3002004712, 2014.

[5] M. Harvey, D. Long and K. Reinhard, "Visualizing NISTIR 7628, Guidelines for Smart Grid Cyber Security," 2014 Power and Energy Conference at Illinois (PECI), Champaign, IL, 2014, pp. 1-8.

[6] D. Long, B. Drennan, and K. Reinhard, "NISTIR 7628 Visualization", Cyber Resilient Energy Delivery Consortium (cred-c.org).

[7] B. Rogers and E. Gilbert, "Identifying architectural modularity in the smart grid: an application of design structure matrix methodology", Grid-Interop Forum, Phoenix AZ, 2011.

[8] A. C. F. Chan and J. Zhou, "On smart grid cybersecurity standardization: Issues of designing with NISTIR 7628," in IEEE Communications Magazine, vol. 51, no. 1, pp. 58-65, January 2013. doi: 10.1109/MCOM.2013.6400439

[9] Guidelines for Smart Grid Cybersecurity, National Institute of Standards and Technology, Interagency Report 7628 Rev. 1, September 2014.

[10] Framework for improving critical infrastructure cybersecurity, National Institute of Standards and Technology (U.S.), 2014.

[11] D. V. Steward, "On an Approach to Techniques for the Analysis of the Structure of Large Systems of Equations", SIAM Review 4 (4):321-342, 1962.

[12] T. R. Browning, "Design Structure Matrix Extensions and Innovations: A Survey and New Opportunities," in IEEE Transactions on Engineering Management, vol. 63, no. 1, pp. 27-52, Feb. 2016.

[13] Bastian M., Heymann S., Jacomy M., Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media, 2009.

[14] W. de Nooy , A. Mrvar, and V. Batagelj. *Exploratory social network analysis with Pajek*. 2nd ed., Cambridge University Press, 2011.

[15] M. Jacomy, T. Venturini, S. Heymann, M. Bastian, "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software", *PLOS ONE* 9(6): e98679, 2014.