SCyPES: Socio-Cyber-Physical-Energy Systems

Rajesh G. Kavasseri, Department of Electrical and Computer Engineering
North Dakota State University, Fargo, ND
rajesh.kavasseri@ndsu.edu
https://sites.google.com/site/rkavasseri/

Abstract: The landscape of electric power grids, especially distribution systems, is rapidly changing. The classic power grid, which can be viewed as a cyber-physical system, is rapidly transforming into a Socio-Cyber-Physical-Energy System (SCyPES) because of smart grid technologies that enable an unprecedented social component. Each layer: Social, Cyber and Physical, has its own attributes which when intertwined with the attributes of another layer, pose unique challenges to the system as a whole. This position paper examines the attributes of these layers, analyzes pair-wise dependencies, points out research challenges, identifies opportunities, along with potential solutions to address some of these challenges.

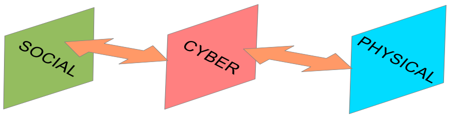

Figure 1: SCyPES: The Cyber layer is sandwiched between the Physical and Social Layers. The bidirectional arrows represent exchange of information in the form of messages/data/commands.

1 Challenges for the "Cyber-Physical" complex

The challenges for the "Cyber" layer mainly stem from Big-data (terra/peta scale) generated from: (i) millions of embedded sensing, monitoring, and recording devices in the physical layer and (ii) end-user interactions from the social layer. The specific questions that arise are:

• Can the present grid (that relies on legacy tools) factor in Big-data to generate real-time analytics, compute and execute (near)-real time control solutions? What changes must be made to existing automation algorithms/tools to ensure scalability?, to survive Big-data?

Achieving scalability while maintaining accuracy in near-real time computations is non-trivial. This cannot be overcome with current *centralized* tools that use a "process data first-compute next" approach, or by merely sizing-up computing resources or by enhanced hardware acceleration schemes. It requires foundational shifts **simultaneously** on two fronts: a) Big-data processing and b) the core computational tools for electric power systems analysis. Drawing clues from Google analytics which processes several thousand of peta (10^{15}) bytes every day and Electronic Design Automation (EDA) tools - which generate circuit solutions at very large scale ($\sim 750,000$ nodes), we recommend pursuing the following two research directions.

- **Merge Big-data processing with domain specific computing routines**: Examine methods to integrate data-handling with core processing such that electric distribution systems of the order of 100,000+ buses can be analyzed with streaming-results available in near-real time.
  One approach is to use *MapReduce* $-$ a programming model used by Google to process Big-data. An emulator can be built to mimic streaming data received from distributed smart meters. The collected data can be consolidated with MapReduce (using Apache/Hadoop) to recover individual data field attributes. Such attributes include real/reactive power, power-factor, peak-load, and other meter acquired electrical variables. This component allow near real-time processing of streaming electrical data for immediate analysis and archival for forensic analysis.

- **Examine alternate computing paradigms for core computational tasks**: The power-flow algorithm serves as the backbone for several other automation tools. Current state-of-the-art power-flow solvers (for distribution systems), still largely rely on a forward-backward sweep which requires complete system information (topology, parameters and data) and central computational resources for convergence. Since certain control/optimization actions in the grid can be executed only after a network solution is available, the latency involved in the data-acquisition, processing and computation stages limits the possibility for near-real time control action. Speeding up this computational task with hardware accelerators/GPUs is one, however unviable brute-force option, because Big-data can quickly outstrip computing resources. EDA tools however have been successful in using statistical approaches to generate circuit solutions at the scale nearly a million nodes. **Can these paradigms be exploited to analyze electric power systems?** As an alternative, one can explore a class of randomized algorithms to generate probabilistic network solutions. This may alleviate the burden of iteratively solving large systems of nonlinear-algebraic equations on a central basis. This may also allow selective and concurrent analysis of any desired portion of the electric grid to any desired level of accuracy. Since this is amenable inherently to parallelization, reasonably accurate solutions to disjoint portions of the grid may be obtained at the gain of several orders of computational time.

2 Challenges for the “Socio-Cyber” complex

The end-users are the final beneficiaries of the electric grid. The smart-grid allows two-way communication between users and service providers via AMIs (Automated Meter Interfaces). Given the “reach” (every home, business, enterprise, industry) of AMI technology, and the sensitivity and volume of control/command/information exchanges, security concerns are of paramount importance. The National Institute of Standards and Technology (NIST) lays out overarching guidelines and security architectures, [1] in terms of *Confidentiality, Integrity,and Availability [CIA].* **To address security concerns, the NIST recommends “mutual authentication” between AMIs and service providers - which is easier said than done!**. Achieving mutual authentication in smart-grids is challenging because of: (i) the large number -several hundreds or thousands within a service territory of AMI devices; and (ii) the large large message volumes and frequency of message exchanges. Traditional public-key infrastructure (PKI) schemes cannot be blindly extended because of:

- increased communication burden, i.e. large key sizes imply a larger communication bandwidth;
- increased time for decryption/verification which implies increased latency;
- resource limitations of AMIs and field devices.

In this context, the main challenges are:

• How can we balance the “communication-encryption” trade-off, without sacrificing the functionalities achievable in the smart-grid? For example, how much? and what type of authentication is required for services such as residential load control programs or demand side management? To address this, we recommend pursuing the following two research directions.

- **A one-size-fits-all approach via encryption is not viable. Instead, examine provably secure “lightweight” authentication schemes [2] for two party communication in the smart-grid**, i.e., those that have low communication and computational burden. One alternative to the prevailing hash based message authentication codes (MAC) is variable length (i.e. message dependent) MAC schemes, [3]. These require very low verification times, and supports a higher upstream communication rate from smart-meters without extra buffering - which translates to reduced memory usage.
- **Multicast settings: Functionality versus Security**: Several applications in the smart-grid require the transmission of a message that is shared (or common) with multiple users. Such messages are transmitted as multicast because unicast may be too expensive (in terms of computational resources) and therefore inefficient. Given the sensitivity of control/command messages exchanged in multicast, authentication is crucial, given the potentially catastrophic consequences that my occur if malicious parties gained unauthorized access, forged messages or mounted replay attacks. A promising solution to multicast authentication is the use of one-time signatures (OTS). A OTS scheme (Lamport and Rabin) generates one digital signature based on a cryptographically secure one-way function without trapdoors

for several messages that are multicast. OTS schemes are attractive for smart-grid applications because signature generation and verification can be done very efficiently, i.e. with low computation complexity. We proposed in [5], for the first time, the generation of OTS from sigma protocols for multicast authentication in the smart grid. A sigma protocol is an interactive three move protocol between a prover P and verifier V to establish the veracity of a statement without explicitly revealing the contents. While sigma protocols, as an emerging technique is finding applications in e-cash, e-voting and e-credentials, its potential for communication in smart-grids has not been explored so far. Our preliminary/ongoing work suggests that σ protocols yields a dramatic reduction in: signing cost (three orders), pre-computation cost (two orders), and storage overhead (two orders) at a very modest increase (four fold) in signature size. This research direction may be explored to several multicast applications from the distribution through sub-transmission and bulk power systems with resource constrained devices.

3 Challenges for the Social layer

The social layer with end-users is probably the hardest to understand because of the complexity of human behavior. For example, the needs and behaviors of users vary with stark contrast. What motivates one user may turn off another and one's necessity may be another's luxury. When these traits are factored with electricity, or energy as an end product, several interesting questions arise.

- **How can one decode human behavior (with respect to energy usage) from a quantum of information concealed in data-traces?, Can the behavior of an end-user be modeled simply from a data-trace? If so, can this information be used for better utilization, or conservation of energy resources?** Despite Google's dominance in the information business, the Google powermeters failed. Why? Even as reports point to inconsistencies with market/regulatory policies, the answer probably lies in complex rules that govern data usage - about what types of data can be shared and how, and with whom? and for what purpose. With emerging paradigms such as "**Transactive Control**" [6] and the "Green button initiative", questions still remain about **how, precisely, can these models "optimize" grid operations**. One line of research is to apply information retrieval and data mining techniques to reduce high dimensional data into a low dimensional space to understand its essential characteristics. What methods or modes of information can best induce change in user behavior? What incentives or interventions are optimal to realize the objectives from (both the user and the grid) - from a control, regulatory, operational or even financial perspective? These are complex questions that require research into a new class of models that factor human behavior into currently established cyber-physical hybrid models.

4 Summary

Addressing these challenges requires closer interaction and joint work from three distinct/diverse disciplines: electric power systems, computer science, cryptography and even behaviorial psychology to create a new class of models that can study the complexities posed by Socio-Cyber-Physical systems.

References

[1] NISTIR 7628: Guidelines for Smart Grid Cyber Security: `http://www.nist.gov/smartgrid/upload/nistir-7628_total.pdf`.

[2] M. M. Fouda, Z. M. Fadlullah, N. Kato, R. Lu and X. Shen, "A lightweight message authentication scheme for smart grid communications", *IEEE Trans. on Smart Grid*, 2(4), pp: 675-685, 2011.

[3] R. Sule, R. S. Katti and R. G. Kavasseri, "A variable length fast Message Authentication Code for secure communication in smart grids", *Proc. IEEE PES General Meeting*, San Diego, July 2012.

[4] C. Hazay and Yehuda Lindell, "Efficient and secure Two Party protocols:Techniques and Constructions", Springer 2010.

[5] R. S. Katti, R. Sule and R. G. Kavasseri, "Multicast Authentication in the Smart Grid with One-Time Signatures from Sigma-Protocols", *Proc. 4th ICCPPS(ACM/IEEE), CPS Week, Philadelphia, PA, April 2013*

[6] `http://www.qualitylogic.com/community/index.php/what-is-transactive-control/`