TCIPG

GridStat Middleware Communication Framework: Management Security and Trust

Resilience of State Estimation using Hybrid SCADA and PMU Data

Yujue Wang, Thoshitha Gamage, and Carl Hauser

GOALS

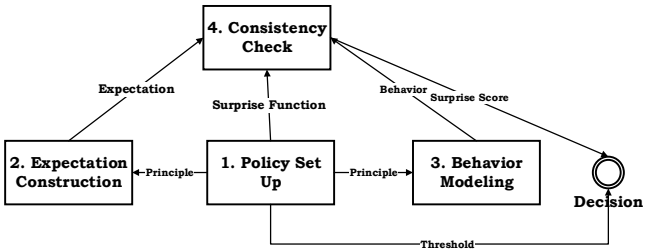
- Develop a trust modeling framework for large-scale critical infrastructures.
- Instantiate the trust framework in a wide-area data delivery system for the power grid.
- With the application of that trust modeling framework, power grid operators can make better use of hybrid SCADA and PMU data in state estimation.
  - PMU data are considered to be more accurate and to have a higher sample rate than SCADA data; PMUs are more "trustworthy."
  - Detect the bad SCADA sensors or identify PMUs that provide bad data.
- Enable sensitive information-sharing among stakeholders.

FUNDAMENTAL QUESTIONS/CHALLENGES

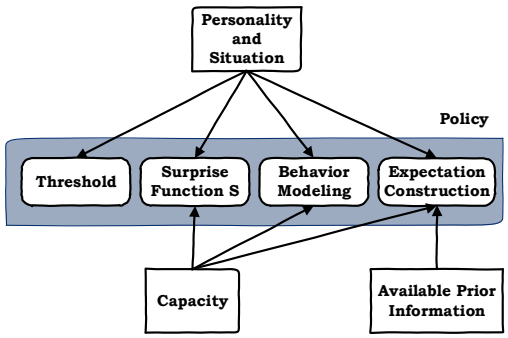
- Generally, trust is expressive enough to cover several types of uncertainties. But the expressiveness also brings difficulties:
  - Trust is inherently abstract, subjective, and uncertain.
  - Trust is multi-faceted and is *belief based on evidence.*
- Lack of mathematical framework for trust modeling that is suitable for the needs of modern critical infrastructure applications such as power grids.
- The bad-data detection problem can be reassembled as a trust-modeling problem, and bad-data sources can be identified either using other, more "trustworthy" data sources/streams, or using historical records. That approach is promising but challenging.

RESEARCH PLAN

- Trust is defined as the subjective belief in the **consistency** between a trustee's behavior and a trustor's expectation under a given context in an environment with uncertainty.
- The framework of trust modeling to identify bad-data sources:



- Policy set-up:



- Operators can shape their expectations pertaining to data streams or data sources according to some prior information, such as reference to more trustworthy data sources or to trustworthy players' endorsements.
- The consistency check function takes an expectation, *E,* formed in the expectation construction phase, and a behavior model, *P,* generated in the behavior modeling phase, as inputs, and outputs a score indicating the difference between the operator's expectation and the actual data streams.

RESEARCH RESULTS

- Two sample models: PE model and SE model.
  - PE model is the model with prior expectation.
    - Using the Kolmogorov-Smirnov (KS) test, measure the difference between the operator's expectation of the SCADA data stream and the actual data stream. The expectation can be obtained from more "trustworthy" data sources.

$$\mathcal{S} = \sup_{x_i} ||F(x_i) - K_N(x_i)||, \forall x_i \in X$$

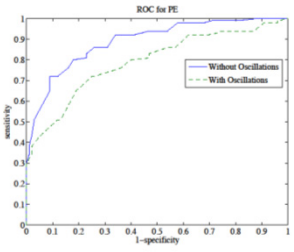
  - SE model is the model with self-built expectation.
    - Assume that each data source's historical data and future data are correlated.
    - Model the distribution of historical data in a time window with the Gaussian kernel function.

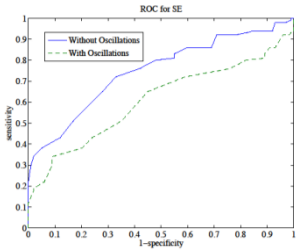
$$\hat{g}(x) = \frac{1}{n(2\pi)^{\frac{d}{2}}\sigma^d} \sum_{i=1}^{n} \exp\{-\frac{||x - x_i||^2}{2\sigma^2}\}$$

    - Measure the difference between distributions of two consecutive windows using relative entropy (Kullback-Leibler (KL) divergence).

$$\mathcal{S} = KL(\hat{g}(x), \hat{h}(\bar{x})) = \int_{-\infty}^{\infty} \ln(\frac{\hat{g}(x)}{\hat{h}(\bar{x})})\hat{g}(x)dx$$

- Simulation test on data streams generated by GridSim.





BROADER IMPACT

- The trust-modeling framework can be extended to other forms of bad-data source identification in addition to SCADA sensors and PMUs.
  - It could possibly be applied to detect malicious users in future transactive energy systems.
- This framework will be useful for entity authentication: The scheme can help agents involved in an authentication procedure determine whether other agents could act in predictable manners.

INTERACTION WITH OTHER PROJECTS

- In general, our research provides a scheme to take advantage of trust models to identify bad-data sources. So it is potentially relevant for any projects focused on trust modeling and bad-data detection.
- Specifically, research studying information-sharing between power utilities can benefit from our research in constructing a reliable and trustworthy channel for sharing operating information.

FUTURE EFFORTS

- Within the framework, refine algorithms and techniques to make the method more functional to data streams with oscillations.
- Expand the method to other applications of PMU data and design a mechanism to do PMU bad-data detection as a service at the system level.
- Examine hybrid SCADA and PMU data streams using a Byzantine data inference view and analyze the resilience of state estimation in various situations.

FUNDING SUPPORT PROVIDED BY DOE-OE AND DHS S&T