Evidence-based Trust for Critical-Infrastructure Decision-Making

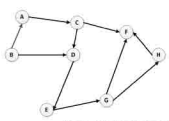
Yujue Wang and Carl Hauser

Goals

- Large-scale critical infrastructures such as the North American electric power grid, which contains $n \times 10^5$ devices and is owned by $m \times 10^3$ organizations, have great diversity among the entities involved. Providing security for such systems is made difficult by uncertainty of trust between the entities.
- The typical security mechanisms need to make explicit "trust assumptions" about other parties participating in the protocol. Trust is then used as a substitute for knowledge in order to demonstrate that the protocol has the security properties that the principal desires.
- We want to make trust explicit in order to:
  a) Ensure trusted entity-key binding in Public Key infrastructure;
  b) Provide guarantees that the information producer and the network meet the QoS properties as contracted;
  c) Share valuable information safely within the power grid.

Fundamental Questions/Challenges

- As our previous research indicates, trustworthiness can be computed directly if the trustor has obtained enough observations y to estimate the probability of trustworthiness, θ, of a trustee using a Bayesian estimator $\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d\theta}$. But how much information is enough? What can the trustors do if there is not enough information?
- On the other hand, the binary trust relation formed by thresholding directly computed trustworthiness values forms a digraph G. An edge from Node A to Node B indicates that A considers B trustworthy. How can trustors make use of the information revealed in this graph for further trust-related inference?



	A	B	C	D	E	F	G	H
A	0	0	1	0	0	0	0	0
B	1	0	0	1	0	0	0	0
C	0	0	0	1	0	1	0	0
D	0	0	0	0	1	0	0	0
E	0	0	0	0	0	0	1	0
F	0	0	0	0	0	0	0	0
G	0	0	0	0	0	1	0	1
H	0	0	0	0	0	1	0	0

An example digraph and adjacent matrix of trust relationships

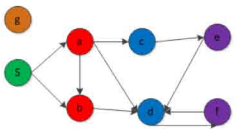
Research Plan

- In some situations, it is not necessary to obtain the exact trustworthiness value. An agent may just need a ranking of the trustees. For example, a data center may rank all of the possible data sources and choose the best ones. On the other hand, the trustor may not have enough information to compute the trustworthiness value directly.
- There should be a mechanism to justify whether an agent has enough information to compute the probability of a trustee's trustworthiness directly.

  At time t, the agent can make use of the past observations $Y_{1:t}$ to estimate the observation at time $t+1$, which is $\tilde{Y}_{t+1}$. At time $t+1$, the agent obtains the real observation Y_{t+1}.

  For example, the agent can then compute the cosine similarity $\cos(\theta_{t+1})$ between $\tilde{Y}_{t+1}$ and Y_{t+1}. If $\cos(\theta_{t+1}) < \epsilon$, where ϵ is a pre-set threshold, it means there was not enough information to compute the trustworthiness probability directly. The trustor then has to compute trust using the indirect method.
- A personalized ranking system can be applied for indirect trust computing. A trustor can make use of the established trust relationships it has with some agents to make a subjective ranking for other agents.
- A Personalized Ranking System should obey four axioms: self-confidence, transitivity, independence of irrelevant alternatives, and incentive compatibility [1].

Research Results



- A distance rule can mostly satisfy these four axioms.
- We can add a discriminative function $r(i)$ to break the ties.
- We use the *PageRank* algorithm to get the value of $r(i)$.

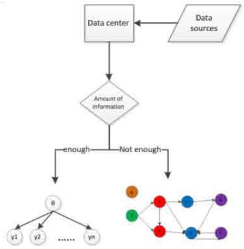
s	a	b	c	d	e	f	g
0.2	0.08	0.1013	0.0213	0.3324	0.0171	0.2579	0

The personalized trust ranking of s:

$$b \succ a \succ d \succ c \succ f \succ e \succ g$$

Broader Impact

- These techniques will be useful for entity authentication. The scheme can help agents involved in an authentication procedure determine whether they locally have enough information to make an authentication decision or if they should refer to others. This reduces the risk of trusting third parties (e.g., CAs) or peers blindly or when making an authentication decision based on partial information.
- With the proposed method, the controllers can improve their ability to assess the trustworthiness of data sources.



Interaction with Other Projects

- In general, our research provides a scheme to solve a trust assessment problem. So it is potentially relevant for any projects focusing on authentication.
- 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

- We will relate the comparison of predictions with the concept of entropy to more formally measure the lack of information.
- We will improve the algorithm to compute the direct trust and make it more practical and easier to implement.
- The current scheme is still a general approach. Applying it to the power grid requires data to determine trustworthiness of power system devices and entities.
- It is necessary to make the ranking algorithms handle a digraph with edge weights representing trustors' trustworthiness value to trustees.
- Various prediction techniques should be assessed for their ability to perform observation prediction.

Reference:
[1] ALTMAN, A. AND TENNENHOLTZ, M. 2006. An axiomatic approach to personalized ranking systems. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*.