Trust Quantification for Networked Cyber-Physical Systems

Yan Wang

***Abstract*—Cyber-physical systems (CPS) are highly integrated hardware-software devices that electro-mechanical components are tightly coupled with advanced computational algorithms for data collection, processing, communication, and control. Internet of Things is the emerging application of CPS. The main research challenge in designing CPS devices and systems is the quantification of complex system behaviors such as consciousness, adaptation, and evolution. Particularly trust becomes an important element that affects system behavior in the networked society. To capture the unique human societal and systems aspects of trustworthiness quantification for CPS systems, in this paper, trustworthiness is measured by the perceptions of ability, benevolence, and integrity quantitatively. Ability measures one's sensing and reasoning capability and influence to others. Benevolence captures the genuineness of intention and the extent of reciprocity in information exchange. Integrity provides the confidence about system dependability and predictability. A generic probabilistic graph model is developed to represent CPS system functionality at mesoscale and demonstrate the perception based quantification of ability and benevolence. Trust-based CPS network design and optimization are also demonstrated with the metrics of ability and benevolence.**

***Index Terms*—Cyber-Physical Systems, Information Exchange, Sensor Fusion, Graph Theory, Network Architecture, Security and Privacy, Trustworthiness, Trust, Ability, Benevolence, and Integrity.**

I. INTRODUCTION

CYBER-physical systems (CPS) are highly integrated hardware-software devices that electro-mechanical components are tightly coupled with advanced computational algorithms for data collection with sensor and actuator on-board, data processing and reasoning with embedded processors, and communication via Internet. They have the integrated capabilities of information collection, processing, and exchange. Compared to traditional mechatronics systems, CPS have much higher levels of complexities and functionalities for sensing and control, especially with advanced communication capability. They also possess more computational power than traditional sensor networks, where data collection and communication are the main functions. Examples of CPS are smart appliances for home and office, intelligent manufacturing systems with sensing and control units, personal devices for health monitoring, disruption-free energy supply and transportation infrastructure, and situation-aware automotive vehicles. With the advancement of novel materials for sensing, actuation, computation, and communication, future CPS may have different physical forms and properties, including those with self-adaptive, self-organizational, biomorphological, and soft structures [1]. CPS may also be created at different size scales and can exist at micro- and nano-scales. Therefore, future CPS can be highly integrated with vehicles, machines, and consumer products, as well as attached on or embedded in human bodies, and work collectively and collaboratively.

The unique value of CPS is their collaborative functions with information shared over the networks formed by themselves, which is being implemented and known as Internet of Things (IoT). Data are constantly collected and shared in a geographically distributed environment. Processing and reasoning for decision making are also done locally in a distributed fashion. This provides the infrastructure of crowdsourcing, where information is no longer processed in centralized locations.

Information crowdsourcing can improve the resilience of systems. In such a federated environment, fewer subsystems or nodes in the network play dominant roles in processing. The negative effects of the breakdown of some nodes thus can be minimized. The costs of maintaining such systems can also be reduced because of the less restrictive requirements on reliability of certain nodes. Nevertheless, there are new technical issues associated with crowdsourcing in IoT, such as interoperability, scalability, adaptability, usability, resilience, security, trust, and privacy [2]. Particularly, most of the CPS functions rely on information exchange. How to design a trustable network so that the data can be shared more freely inside the network than outside. Data security is critical for trust. However, security along cannot guarantee the trustworthiness. Although security protocols and policies can prevent data from being compromised during transmission, they provide no guarantee against the misuse by the receiving party or fraud by the transmitting party. In a secured network, partners could still avoid sharing necessary information because of competitions or conflicts of interests. In those cases, trustable relationship still cannot be established even in a secured environment.

Here, how to quantify trustworthiness in the context of systems of CPS and apply the concept to design IoT are studied. Although trust has been studied in the fields of computer science, psychology, marketing, and management, most of the studies remain qualitative. In order to be useful for system design, quantification of trust is necessary so that quantitative criteria can be used for design optimization.

In recent research of trust quantification for networked systems, two approaches are taken to quantify trustworthiness. In the top-down approach, trustworthiness is treated as an overall perception or belief about an individual's reputation or ability. It is quantified with probabilistic or non-probabilistic

Y. Wang is with Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA e-mail: yan.wang@me.gatech.edu

measures. In the bottom-up approach, trustworthiness is not treated as one single concept. Rather, it consists of multiple factors such as availability, dependability, and quality of services, each of which can be calculated from the statistics of physical systems, e.g. data transmission rates, executed routing protocols, and positive recommendations. For applications, the metrics of trustworthiness were used in assessing system availability or designing dynamic routing protocols.

Compared to traditional computational hardware and software systems, trustworthiness for CPS has some unique properties and challenges. The first one is the system-oriented quantification and its complexity. CPS networks formed by industrial or consumer products are exposed to severe threats of malicious attacks. Security breach at individual components levels cannot be completely prevented. Working along with disrupted services or compromised information is becoming a norm in networked CPS. Furthermore, CPS have highly integrated and autonomous functions of sensing, data processing, predictive modeling, decision making, and control. Therefore, to be meaningful for CPS systems design and engineering, the quantification of trustworthiness in CPS needs to target at the systems level with multiple functions and networked communities, instead of only at individual components or functions. The complexity of quantification at systems level is increased with more factors to be considered, such as availability, resilience, and adaptability of networks. These factors further increase the complexity of trust quantification and management.

The second uniqueness of CPS trustworthiness is the influence of perception in a human societal context. Trust is a state of mind, subjective and multi-faceted. Given the intensive interaction between humans and CPS, user-centric trust management is essential for CPS. Traditional security procedures (e.g. for authentication, nonrepudiation, confidentiality, data integrity, privacy protection, etc.) can enhance trust level based on the assumption that security control is always available. However, complete control of information access by content owners in CPS networks is impossible, because there is deep interdependency of transactions between heterogeneous subsystems, implementation of full security procedures for traditional computers on low-cost CPS units is infeasible, and the controllability can be quickly diminished as information is more likely to propagate through CPS networks than traditional computer networks. As a result, people's perception of trust is likely to change and become more tolerable for security and privacy related challenges, as new CPS technologies are adopted gradually in the society and new regulations or other social insurance for protection are available. Therefore, the dynamics of human perception and subjectivity needs to be emphasized in trustworthiness quantification for CPS. Since perception is influenced by social activities and interactions between people, the social behavior aspect of trust also needs to be captured.

To address the above challenges, especially the second one, in this paper a perception-oriented approach is taken to quantify trustworthiness. Trustworthiness is measured by perceptions of three major metrics instead of an abstract one. The three metrics, which include perceptions of *ability*, *benevolence*, and *integrity*, are carefully chosen based on the concepts studied in social sciences and to avoid redundancy. To model large-scale CPS networks, a probabilistic graph model is proposed to capture the functions of sensing, prediction, and communication. This mesoscale model provides a generic abstraction of CPS networks with scalability consideration. The proposed three trustworthiness metrics are calculated based on the probabilistic graph model. It is demonstrated that these three perception-level metrics can be calculated with the combination of Bayesian and statistical methods. Compared to other trustworthiness quantification approaches, the uniqueness of the proposed approach includes the considerations of different CPS functions including sensing, prediction, and communication. The perception based quantification method directly models subjectivity of beliefs and the influence of social behavior, with quantitative measures of ability, benevolence, and integrity, which have not been considered in other quantitative approaches.

In the remainder of the paper, a review of relevant work on trust quantification in the domains of computer and social sciences is first provided. Then the proposed probabilistic graph model is introduced in Section II. The metrics of ability, benevolence, and integrity are provided in Sections III, IV, and V respectively. The performance of the three metrics is evaluated with simulation studies in Section VI. In Section VII, the network design and optimization approach based on the metrics of ability and benevolence is demonstrated.

A. Trust quantification

In computer science, the study of trust had been traditionally centered around security policy for exchanging credentials, controlling access, and referring reputation [3], [4], [5], [6]. Recently, it was expanded to the context of social networks and semantic web [7], [8]. In the vast majority of the studies, trust was only treated qualitatively without providing specifications of how to calculate it. Limited work is on trust quantification.

In the context of social networks and multi-agent environments, most researchers ([9], [10], [11], [12], [13], [14], [15], [16], [17], [18]) model trust as reputation and rely on users' explicit ratings and recommendations to estimate the levels of trust. For instance, Beth et al. [19] quantified trust by the numbers of positive and negative experiences. Yu and Singh [20] calculated it from scaled reputation ratings in social networks. Lee et al. [21] calculated trust as the number of finished transactions. O'Doherty et al. [22] combined users' explicit ratings with the similarities of opinions and preferences for online recommendation.

In the context of computer networks, trust was mostly measured by quality of services [23] in a bottom-up fashion. Trustworthiness has been calculated from weighted metrics such as the numbers of forwarded data packets, executed routing protocols, modified packet addresses, etc. [24], [25]. For sensor networks, different approaches to calculate trust has been proposed. These include weighted average between local and global success rates of transactions [26], weighted average of consistency factors including consistency of individual nodes from their historical data [27] and consistency between nodes in local regions [28], as well as neighbors' data forwarding behaviors [29].

To capture the stochastic and subjective nature of trust, probabilistic approaches have been developed. Barber and Kim [30] modeled trust as belief of information reliability and the belief update is based on Bayesian networks. Yu and Singh [31] applied Dempster-Shafer evidence theory to trust modeling. Patel et al. [32] modeled trust as the expectation of fulfilled commitments that follows a Beta distribution. Wang et al. [33] extended this approach to Bayesian modeling based on the natural conjugate Beta distribution and generic Bayesian networks [34]. Kim et al. [35] calculated trust as the probability of resource availability. Li et al. [36] calculated trust as time averaged information entropy in data exchange.

Fuzzy logic was also applied to capture the linguistic imprecision of trust description, either as one concept [37], [38], or a combination of multiple factors such as ability, availability, motivation, usefulness, honesty, and others [39], [40], [41], [42].

To quantify trustworthiness in IoT, Chen et al. [43] proposed a fuzzy model to consider communication reputation factors including packet forwarding, package delivery, and energy efficiency. Saied et al. [44] quantified it based on user ratings and recommendations. Nitti et al. [45] used an weighted average between quality of service and opinions of credibility in transactions as a combined objective and subjective measure. Chen et al. [46], [47] treated trust as the combination of an overall probabilistic assessment from direct interaction and the social similarity in a recommendation system. Al-Hamadi and Chen [48] calculated trust from user ratings aggregated from different time periods and different locations. Jayasinghe et al. [49] used a composite trust metric that is consisted of eight different quantities such as probabilities of successful execution, cost of execution, completeness and accuracy of data records, etc. In the above approaches, the perception of trust in a social environment is not explicitly modeled. Only the communication function of IoT objects is focused on, whereas the functions of sensing and reasoning are not considered.

Intuitively, trust is a willingness to be vulnerable to another. Different qualitative definitions of trust exist in the domains of psychology, marketing, human behavior, and organization. Mayer et al. [50] carefully studied dozens of characteristics of trust in literature, identified commonality, and defined trustworthiness as a set of three categories of perception: *ability*, *benevolence*, and *integrity*. Ability is about perception of skills, expertise, and competency associated with trustee. Benevolence is the extent to which the trustor believes that the trustee acts for the welfare of the trustor, rather than just maximizes its own profit. Benevolence is a summary of related characteristics such as loyalty, openness, receptivity, and availability. Integrity is the trustor's perception that the trustee will be honest and adhere to an acceptable set of principles. Integrity is associated with consistency, discreetness, fairness, promise fulfillment, reliability, and value congruence.

The trust model of ability, benevolence, and integrity has been widely adopted in different fields. The three factors have been applied in designing psychological and behavioral studies of trust [51], [52]. The model was applied to measure the trustworthiness of online shopping merchants [53], [54], [55], electronic banking service providers [56], [57]. The model has also been adopted in designing information systems with better privacy policies [58], better understanding of users intention [59], [60], user participation [61], security [62], and technology integration [63]. The concerns of security, privacy, and trust in CPS networks are similar to those in traditional information systems. The trust model of ability, benevolence, and integrity thus can be applied in CPS. Nevertheless, in all of the above studies ability, benevolence, and integrity were defined qualitatively without providing quantitative measures.

Each of the three trust factors (ability, benevolence, and integrity) captures some unique aspects of trust, and they are independent to each other. That is, the degree of perceived ability of trustee does not indicate its level of benevolence or integrity, and vice versa. To compare the multi-criteria trustworthiness of two parties, if two factors are at the same level, whoever has a higher degree of the third factor dominates and is more trustable. For non-dominant cases, different methods have been used, such as multi-objective optimization where Pareto front is identified without combining the criteria, multi-attribute utility theory to evaluate alternatives strictly based on preferences, and weighted sum of attributes to combine the criteria. Quantitative methods such as data envelopment analysis, analytic hierarchy process, and ELECTRE (elimination and choice expressing the reality) have been developed to determine weights and rankings with consistency. The effects of the three perceptions also dynamically change over time. Mayer et al. [50] also postulated that among the three factors the effect of integrity on trust is the most salient one at the early stage of trust relationship, and the effect of perceived benevolence will increase over time.

In summary, although trust has been considered as a critical component in electronic business, computer networks, and sensor networks, there is still a lack of formal and quantitative methods to study how trustworthiness can affect system design. In the design of CPS networked systems, trustworthiness directly affects the formation of information sharing policies adopted by the products, which in turn influences the design of networks. The proposed quantification method is targeted at system and network design.

B. *Proposed quantitative trustworthiness metrics*

Trust has a few unique characteristics. It is subjective (trust is personal and based on individual's perception), asymmetric (trust relationship is mutual but not necessarily symmetric), and dynamic (the level of trust is not static and changes along time). Trustworthiness metrics need to incorporate these properties.

In this paper, ability, benevolence, and integrity are quantitatively measured for the first time, even though the three concepts have been widely used in qualitative assessments and user studies. The quantification process is based on a probabilistic graph model. The graph model provides a mesoscale abstraction to represent the major functions of information gathering and exchange between nodes in CPS networks.

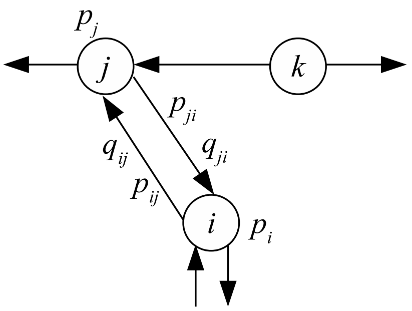


Fig. 1: The probabilistic graph model

Particularly, the probability of accurate sensing and prediction by each node, as well as the probabilities of positive and negative correlations as mutual influence between nodes are explicitly modeled, which allows for quantitative measurement of ability and benevolence directly. Notice that ability, benevolence, and integrity aspects of trust co-exist in trustworthiness quantification. These three independent dimensions need to be considered separately and not simply combined into one metric. Also different from other approaches, a perception based belief modeling approach is taken to calculate ability, benevolence, and integrity, given that the level of trust is based on perception.

II. Probabilistic Graph Model

Recently a probabilistic graph model was proposed to provide an abstraction of information collection, processing, and exchange between CPS in an IoT environment [64], [65]. Here, the model is generalized. A probabilistic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{R}, \mathcal{P}, \mathcal{Q})$, where $\mathcal{V} = \{v_k\}$ is a set of nodes, $\mathcal{E} = \{(v_i, v_j)\}$ is a set of directed edges, as shown in Fig. 1. Each node v_k is associated with a *prediction probability* $p_k \in \mathcal{R}$, whereas each edge (v_i, v_j) is associated with a *P-reliance probability* $p_{ij} \in \mathcal{P}$ and a *Q-reliance probability* $q_{ij} \in \mathcal{Q}$. The probabilities are defined as follows.

The prediction probability p_k is the probability that node k detects the true state of world θ and is defined as

$$P(x_k = \theta) = p_k, \tag{1}$$

where x_k is the state variable of node k. The information dependency between node j and node i is described by P-reliance probability

$$P(x_j = \theta | x_i = \theta) = p_{ij}, \tag{2}$$

and Q-reliance probability

$$P(x_j = \theta | x_i \neq \theta) = q_{ij}. \tag{3}$$

P-reliance probability indicates the positive effect of information exchange between nodes, whereas Q-reliance probability captures the negative influence. It is also possible to have $P(x_k = \theta | x_k = \theta)$ and $P(x_k = \theta | x_k \neq \theta)$ indicating how much a node's prediction relies on its own observation.

The probabilistic graph proposed here can be regarded as a generalization of classical graph where weights are associated with both nodes and directed edges. The weights are probabilities of prediction and reliance. Also note that the probabilistic graph model here is different from Bayesian belief network.

CPS nodes collect information by itself or from their neighboring collaborators. With the new information, the prediction probabilities are updated. Different *information fusion rules* can be adopted by nodes to update their prediction probabilities. Example rules include *best-case*, *worst-case*, and *Bayesian* rules. They are listed as follows. To simplify the notation, we use $P(x_k)$ to denote $P(x_k = \theta)$, $P(x_k^c)$ to denote $P(x_k \neq \theta)$, and $P(x_j|x_i)$ to denote $P(x_j = \theta | x_i = \theta)$.

The *best-case* or *optimistic* fusion rule is

$$P'(x_k) = 1 - \Pi_{i=1}^{M}(1 - P(x_k|x_i)), \tag{4}$$

where node k has a positive prediction with updated probability P' if any of the M nodes as its information sources provides a positive cue. Some variations of this rule can also be used, such as

$$P'(x_k) = 1 - \Pi_{i=1}^{M_1}(1 - P(x_k|x_i))\Pi_{j=1}^{M_2}(1 - P(x_k|x_j^c)), \tag{5}$$

where both positive cues from M_1 nodes and negative cues from M_2 nodes ($M_1 + M_2 = M$) are considered. Another version could be

$$P'(x_k) = 1 - \Pi_{i=1, i\neq k}^{M}(1 - P(x_k|x_i)), \tag{6}$$

where the node's own prior observation is not included in the update.

The *worst-case* or *pessimistic* fusion rule is

$$P'(x_k) = \Pi_{i=1}^{M} P(x_k|x_i), \tag{7}$$

where the prediction of a node is positive only if all cues it receives from other nodes are positive. Similarly, some variations of the rule exist, such as

$$P'(x_k) = \Pi_{i=1}^{M_1} P(x_k|x_i)\Pi_{j=1}^{M_2} P(x_k|x_j^c), \tag{8}$$

and

$$P'(x_k) = \Pi_{i=1, i\neq k}^{M} P(x_k|x_i). \tag{9}$$

The *Bayesian* fusion rule is

$$P'(x_k) \propto P(x_k)\left[(P(x_k))^r(1 - P(x_k))^{M-r}\right], \tag{10}$$

where the prediction of node k is updated from prior prediction probability $P(x_k)$ given that r out of a total of M cues provided by others are positive.

For simplicity, only binary value of state variables is considered here. Obviously, further generalization to multi-valued discrete state variables is straightforward. Suppose there are a finite set of discrete values $\{\theta_1, \ldots, \theta_N\}$ that the state variable x_k can take. The multi-valued prediction probability $P(x_k = \theta_n)$ ($n \in \{1, \ldots, N\}$) can be obtained similar to binary values. Similarly, reliance probabilities $P(x_j = \theta_n | x_i = \theta_m)$ ($m, n \in \{1, \ldots, N\}$) can be obtained enumeratively.

The above information fusion rules can be similarly extended to multi-valued state variables. For instance, the optimistic fusion rule in Eq.(5) becomes

$$\begin{aligned} P'(x_k) = 1 - &\Pi_{i=1}^{M_1}(1 - P(x_k|x_i = \theta_1)) \cdots \\ &\Pi_{j=1}^{M_N}(1 - P(x_k|x_j = \theta_N)), \end{aligned} \tag{11}$$

whereas the pessimistic fusion rule in Eq.(8) becomes

$$P'(x_k) = \Pi_{i=1}^{M_1} P(x_k|x_i = \theta_1) \cdots \Pi_{j=1}^{M_N} P(x_k|x_j = \theta_N), \tag{12}$$

where $M_1 + \cdots + M_N = M$. Obviously, if only one of the N values is true or of concerned, the problem setting can be simplified and converted back to the binary case. For continuous state variables, discretization of values is necessary, which is typical in the context of digital world.

Another closely related question is how the prediction and reliance probabilities can be obtained in a physical system. The most straightforward way is to estimate the probabilities from the collected historical data [66]. The prediction probability of a CPS node can be based on data collected by its sensing and reasoning units. It can be estimated as the frequency of correct prediction. The reliance probabilities can be estimated similarly from the frequencies of positive and negative predictions by the neighboring nodes given the node's own prediction. If no data are available, subjective estimations from domain experts can be elicited.

Probability elicitation is well known in both practice and literature. Standard procedures are taken to elicit probabilities associated with some events from domain experts subjectively. The major issues to be avoided during elicitation are possible personal bias, inconsistency, and incoherent assessment. Quantitative methods [67] such as quantile intervals and scoring rules have also been developed. For instance, in the scoring rule approach, a scoring function defines rewards such that the domain expert's true belief about how likely an event occurs is reported, as it is his or her best interest to maximize the belief-weighted expected payoff, which is to ensure consistency.

III. Ability

The ability of CPS consists of *capability* and *influence*. In the context of probabilistic graph model, the capability of a node is generally quantified as the perceived probability that it can provide accurate prediction about the true state of the world based on its available information. Within a networked society, the influence of a node to others, which can be interpreted as leadership, is also regarded as part of its ability. The leadership that a CPS node has is characterized as the extent of its positive or negative influence to its neighbors.

A. Capability of prediction

The perception of ability is subjective and varies among different people. Suppose that the perceived capability for node j is the perceived prediction probability $A_j(\theta) = \mathbb{P}(P(x_j = \theta))$ with respect to the true state of world θ, which follows a Gaussian distribution with mean p_j and precision τ_0. Here $\mathbb{P}(\cdot)$ denotes perception. In other words, the perceived capability of node j is randomly distributed, with expectation

$$\mathbb{E}(A_j(\theta)) = p_j, \tag{13}$$

and variance

$$\mathbb{V}(A_j(\theta)) = \tau_j^{-1}. \tag{14}$$

In a society with extensive information exchange, the perception of capability can be updated with newly obtained information. For instance, if reliance probabilities with respect to node j are made available to the public, then the perceived capability of the node can be updated. Suppose that the perceived reliance probabilities $L_{ij} = \mathbb{P}(P(x_j|x_i))$ and $L_{ij}^c = \mathbb{P}(P(x_j|x_i^c))$ for all $i, j \in \mathcal{V}$ are Gaussian random variables, with expectations

$$\mathbb{E}(L_{ij}|A_j) = p_{ij} (\forall i, j \in \mathcal{V}) \tag{15}$$

and

$$\mathbb{E}(L_{ij}^c|A_j) = q_{ij} (\forall i, j \in \mathcal{V}) \tag{16}$$

under the condition of the perceived capability A_j.

The variances of the perceptions may depend on the nature of information sources. For the perceptions related to the information shared with node j from others, the variances are

$$\mathbb{V}(L_{ij}|A_j) = \tau_{ij,p}^{-1} (\forall i \in \mathcal{S}_j), \tag{17}$$

and

$$\mathbb{V}(L_{ij}^c|A_j) = \tau_{ij,q}^{-1} (\forall i \in \mathcal{S}_j), \tag{18}$$

where $\mathcal{S}_j = \{v_i|(v_i, v_j) \in \mathcal{E}\}$ is the collection of *source nodes* with respect to node j and each of the source nodes sends information to node j. Without the loss of generality, we can assume that the variances of the perceived reliance probabilities are the same, i.e. $\tau_{ij,p} = \tau_{s,p}$ and $\tau_{ij,q} = \tau_{s,q}$ $(\forall i \in \mathcal{S}_j)$. The complete set of perceived P- and Q-reliance probabilities for the source nodes with respect to node j is denoted as $\mathcal{L}^{(+j)} = \{L_{ij}|i \in \mathcal{S}_j\} \cup \{L_{ij}^c|i \in \mathcal{S}_j\}$.

With the reliance probability information, the perception of capability is updated based on the Bayesian belief update scheme or Bayes' rule. Because the perceptions follow Gaussian distributions, the expectation of posterior perception for capability of node j is

$$\mathbb{E}(A_j(\theta|\mathcal{L}^{(+j)})) = \frac{\tau_j p_j + \tau_{s,p} \sum_{i \in \mathcal{S}_j} p_{ij} + \tau_{s,q} \sum_{i \in \mathcal{S}_j} q_{ij}}{\tau_j + \tau_{s,p} s_j + \tau_{s,q} s_j}, \tag{19}$$

where $s_j = |\mathcal{S}_j|$ is the number of source nodes with respect to node j.

The consideration of Q-reliance probabilities in capability in Eq.(19) is necessary. When a node gives correct prediction even when its information sources provide negative or wrong predictions, the node exhibits good capability. Also note that if the assumption of equal variances is not made, the posterior perception of capability in Eq.(19) still can be calculated with $\tau_{s,p} \sum_{i \in \mathcal{S}_j} p_{ij}$ replaced by $\sum_{i \in \mathcal{S}_j} \tau_{ij,p} p_{ij}$, $\tau_{s,q} \sum_{i \in \mathcal{S}_j} q_{ij}$ by $\sum_{i \in \mathcal{S}_j} \tau_{ij,q} q_{ij}$, $\tau_{s,p} s_j$ by $\sum_{i \in \mathcal{S}_j} \tau_{ij,p}$, and $\tau_{s,q} s_j$ by $\sum_{i \in \mathcal{S}_j} \tau_{ij,q}$ respectively.

The posterior perception of capability in Eq.(19) can be regarded as the weighted average of prediction and reliance probabilities, denoted as

$$\mathbb{E}(A_j(\theta|\mathcal{L}^{(+j)})) = \alpha_j p_j + \alpha_{s,p} \sum_{i \in \mathcal{S}_j} p_{ij} + \alpha_{s,q} \sum_{i \in \mathcal{S}_j} q_{ij}, \tag{20}$$

where $\alpha_j = \tau_j / (\tau_j + \tau_{s,p} s_j + \tau_{s,q} s_j)$, $\alpha_{s,p} = \tau_{s,p} / (\tau_j + \tau_{s,p} s_j + \tau_{s,q} s_j)$, and $\alpha_{s,q} = \tau_{s,q} / (\tau_j + \tau_{s,p} s_j + \tau_{s,q} s_j)$.

The variance of the updated perceptions for the capability of node j is

$$\mathbb{V}(A_j(\theta|\mathcal{L}^{(+j)})) = (\tau_j + \tau_{s,p} s_j + \tau_{s,q} s_j)^{-1}. \tag{21}$$

B. Influence

The influence or leadership of node j is associated with the effectiveness of information sharing from node j to others. When the information sharing from node j to *destination nodes* in $\mathcal{D}_j = \{v_k | (v_j, v_k) \in \mathcal{E}\}$ is considered, where each of the destination nodes receives information from node j, the perception of the capability of node j can be further updated. When the precision of the perceptions related to the information shared *from* node j to others are characterized by

$$\mathbb{V}(L_{jk}|A_j) = \tau_{jk,p}^{-1} (\forall k \in \mathcal{D}_j) \tag{22}$$

and

$$\mathbb{V}(L_{jk}^c|A_j) = \tau_{jk,q}^{-1} (\forall k \in \mathcal{D}_j), \tag{23}$$

the complete set of perceived P- and Q-reliance probabilities for the destination nodes with respect to node j is denoted as $\mathcal{L}^{(-j)} = \{L_{jk} | k \in \mathcal{D}_j\} \cup \{L_{jk}^c | k \in \mathcal{D}_j\}$. Similarly, to simplify the notation, it is assumed that the variances of the perceived reliance probabilities are the same, i.e. $\tau_{jk,p} = \tau_{d,p}$ and $\tau_{jk,q} = \tau_{d,q}$ $(\forall k \in \mathcal{D}_j)$. The expectation of the updated perception of ability based on the influence to others is

$$\begin{aligned}&\mathbb{E}(A_j(\theta|\mathcal{L}^{(-j)})) = \\ &\qquad \frac{\tau_j p_j + \tau_{d,p} \sum_{k \in \mathcal{D}_j} p_{jk} + \tau_{d,q} \sum_{k \in \mathcal{D}_j} (1 - q_{jk})}{\tau_j + \tau_{d,p} d_j + \tau_{d,q} d_j},\end{aligned} \tag{24}$$

where $d_j = |\mathcal{D}_j|$ is the number of destination nodes with respect to node j, or simply

$$\begin{aligned}&\mathbb{E}(A_j(\theta|\mathcal{L}^{(-j)})) = \\ &\qquad \alpha_j p_j + \alpha_{d,p} \sum_{k \in \mathcal{D}_j} p_{jk} + \alpha_{d,q} \sum_{k \in \mathcal{D}_j} (1 - q_{jk}),\end{aligned} \tag{25}$$

where $\alpha_{d,p} = \tau_{d,p}/(\tau_j + \tau_{d,p} d_j + \tau_{d,q} d_j)$, and $\alpha_{d,q} = \tau_{dq}/(\tau_j + \tau_{d,p} d_j + \tau_{d,q} d_j)$. Notice that $(1 - q_{jk})$ is used here to quantify the influence of node j to others, which captures how likely others end up with negative predictions given that node j provides a negative cue.

The variance of the updated perceptions for node j's ability after obtained information from destination nodes is

$$\mathbb{V}(A_j(\mathcal{L}^{(-j)})) = (\tau_j + \tau_{d,p} d_j + \tau_{d,q} d_j)^{-1}. \tag{26}$$

C. Overall ability

The expectation of the further updated perception of ability that includes both capability of prediction and influence to others is

$$\begin{aligned}&\mathbb{E}(A_j(\theta|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)})) = \\ &\qquad \frac{\begin{bmatrix} \tau_j p_j + \tau_{s,p} \sum_{i \in \mathcal{S}_j} p_{ij} + \tau_{s,q} \sum_{i \in \mathcal{S}_j} q_{ij} \\ + \tau_{d,p} \sum_{k \in \mathcal{D}_j} p_{jk} + \tau_{d,q} \sum_{k \in \mathcal{D}_j} (1 - q_{jk}) \end{bmatrix},}{\tau_j + \tau_{s,p} s_j + \tau_{s,q} s_j + \tau_{d,p} d_j + \tau_{d,q} d_j}\end{aligned} \tag{27}$$

or simply

$$\begin{aligned}&\mathbb{E}(A_j(\theta|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)})) = \\ &\qquad \alpha_j p_j + \alpha_{s,p} \sum_{i \in \mathcal{S}_j} p_{ij} + \alpha_{s,q} \sum_{i \in \mathcal{S}_j} q_{ij} \\ &\qquad + \alpha_{d,p} \sum_{k \in \mathcal{D}_j} p_{jk} + \alpha_{d,q} \sum_{k \in \mathcal{D}_j} (1 - q_{jk})\end{aligned}, \tag{28}$$

where weights $\alpha_j, \alpha_{s,p}, \alpha_{s,q}, \alpha_{d,p}, \alpha_{d,q}$ are defined accordingly.

The variance of the updated perceptions for node j's ability after both information from source and destination nodes is

$$\begin{aligned}&\mathbb{V}(A_j(\theta|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)})) = \\ &\qquad (\tau_j + \tau_{s,p} s_j + \tau_{s,q} s_j + \tau_{d,p} d_j + \tau_{d,q} d_j)^{-1}.\end{aligned} \tag{29}$$

D. Higher-order perception

In a society, one's perception can be influenced by others' perceptions. In the context of trust, one's perceived trustworthiness can be a function of others' perceived trust levels because of mutual influence in judgment and decision making. Therefore, the previous ability perception model can be further extended to a higher-order one with the consideration of mutual influence. The expected ability in Eq.(27) or (28) and variance in Eq.(29) are considered as the *first-order* model, where the perception of a node's ability is only affected by its interaction with the immediate neighbors. For the second-order model, the ability of a node is also affected by the perceived abilities of its intermediate neighbors, particularly the destination nodes which it directly shares information with. That is, the ability of a node is also related to the abilities of the nodes that it has direct influence on.

If the notations of $\mathbb{E}(A_j(\theta|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)}))$ and $\mathbb{V}(A_j(\theta|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)}))$ are simplified to

$$\mathbb{E}(A_j(\theta|+,-)) = E_j \tag{30}$$

and

$$\mathbb{V}(A_j(\theta|+,-)) = V_j \tag{31}$$

respectively, then in the *second-order* model, the expected ability is

$$\begin{aligned}&\mathbb{E}^{(2)}(A_j(\theta|+,-)) \\ &= \frac{\begin{bmatrix} V_j^{-1} \cdot E_j + \tau_{d,p} \sum_{k \in \mathcal{D}_j} p_{jk} (V_k^{-1} \cdot E_k) \\ + \tau_{d,q} \sum_{k \in \mathcal{D}_j} (1 - q_{jk}) (V_k^{-1} \cdot E_k) \end{bmatrix}}{V_j^{-1} + \tau_{d,p} \sum_{k \in \mathcal{D}_j} p_{jk} V_k^{-1} + \tau_{d,q} \sum_{k \in \mathcal{D}_j} (1 - q_{jk}) V_k^{-1}},\end{aligned} \tag{32}$$

which is the same as the first-order expectation, and the variance is

$$\begin{aligned}&\mathbb{V}^{(2)}(A_j(\theta|+,-)) \\ &= \left[V_j^{-1} + \tau_{d,p} \sum_{k \in \mathcal{D}_j} p_{jk} V_k^{-1} + \tau_{d,q} \sum_{k \in \mathcal{D}_j} (1 - q_{jk}) V_k^{-1} \right]^{-1}.\end{aligned} \tag{33}$$

Similarly, the third-order model can be constructed by incorporating the perceived abilities of the neighbors' neighbors, which the reference node indirectly shares information with. Therefore, the higher-order perception model incorporates the lower-order perceptions, as an extension of weighted averages where weights are the associated precisions. Recursively the n^{th} order model is defined based on the $(n-1)^{th}$ order ones.

From Eqs.(33) and (31), it is seen that the variance of the second-order perception is lower than that of the first-order, because the additional terms are added in the precision. By incorporating a higher-order perception, the estimation of trustworthiness in terms of ability can become more precise.

E. Perception update

One's perception will be updated as new information is available. For instance, when prediction capability of a node is increased with newly upgraded censors or reduced because of malicious attacks, the perception of prediction capability needs to be updated. Similarly, the modifications and fluctuations of reliance probabilities, because of the behavior dynamics between nodes or network topology changes, will also affect the perception of capability and influence.

The update of perception is based on Bayes' rule. Suppose that the expected value and variance of the originally prediction probability for node j at time t_k are $p_j^{(k)}$ and $\tau_j^{(k)}$. When new estimate of prediction probability for node j is available as $p_j^{(new)}$ with precision $\tau_j^{(new)}$ at time t_{k+1}, the prediction probability needs to be updated to

$$p_j^{(k+1)} = \frac{\tau_j^{(k)} p_j^{(k)} + \tau_j^{(new)} p_j^{(new)}}{\tau_j^{(k)} + \tau_j^{(new)}}. \tag{34}$$

An important step of perception update is determining the precision value $\tau_j^{(new)}$. To avoid the dominance of a very small variance that causes overly-emphasized new prediction in Bayes' rule, a pre-determined lower bound of variance V_{min} can be set. This procedure can be applied to prevent perception attack where a biased perception is used to swing other's perceptions. The precision is updated to

$$\tau_j^{(k+1)} = \tau_j^{(k)} + \tau_j^{(new)}. \tag{35}$$

The perception update because of reliance probability changes, either expected value or variance, as well as high-order perceptions can be done similarly.

F. Illustrative examples

Here, two examples are used to illustrate ability based on perceived capability and influence. The first example is a randomly generated graph which is to illustrate the relationships among capability, influence, and ability. The second example provides more details of how the ability of a node is affected by its neighbors and prediction and reliance probabilities.

In the first example, a directed graph with 50 nodes as shown in Fig. 2 is constructed, where the edge connections between nodes are randomly generated. The heavy tail at the end of an edge in the figure denotes an arrow, indicating an incoming vertex (e.g. the information flow direction from node 37 to node 47 is shown). The probability that there is an edge between two nodes is set to be 0.08. The prediction, P-, and Q-reliance probabilities are randomly generated from an uniform distribution between 0 and 1. Similarly, the variances associated with the prediction and reliance probabilities are randomly generated from a uniform distribution between 0 and 0.5.

Notice that the variance of 0.5 already over-estimates the possible variation ranges. A variance of 0.5 corresponds to a standard deviation of over 0.7. From a normal distribution, the probability of a random value falls within three times of standard deviation is 99.73%. For a probability value, this already goes beyond the range of 0 and 1. Choosing the potential large variance of 0.5 is for illustration purpose.

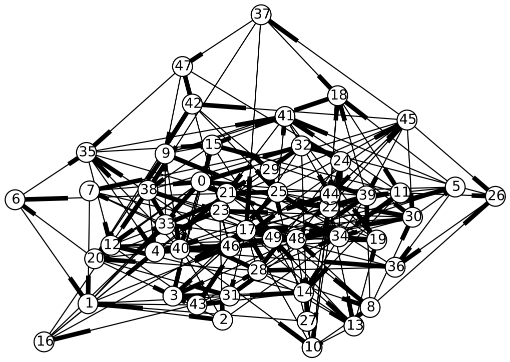


Fig. 2: A random graph with 50 nodes

Based on equations similar to Eqs.(19), (24), and (27), the expectations of the perceived capabilities, influences, and ability for the 50 nodes are shown in Fig. 3 respectively. Here the variance or precision of each probability is used during the calculation of the weighted averages, instead of the average values of $\tau_{s,p}$, $\tau_{s,q}$, $\tau_{d,p}$, and $\tau_{d,q}$. The expected values for each node are denoted as dots in Fig.3, where the error bar indicates the standard deviation, or the square-root of the variance. It is seen that the expected values of the perceived capability, influence, and ability could vary based on the available information. The average expected values of the 50 capabilities, influences, and abilities are 0.4722,0.5057, and 0.4965 respectively. The average standard deviations for them are 0.1192, 0.1309, and 0.0774 respectively. The variance of perceived ability is smaller than those of perceived capability or influence only. As more information is included, the precision of the perception can be improved. This can also be confirmed by the variances of these three standard deviations, which are 0.0027,0.0081, and 0.0010. That is, the precision of perception tends to be more consistent as new information is included.

From Eq.(29), it is also seen that the variance of perceived ability reduces as more information source or destination nodes are added. That is, the larger the number of nodes that a node directly exchanges information with, the more precise that the perception about this node's ability becomes. However, this does not necessarily lead to a higher ability. As seen in Eq.(27), different precision levels of prediction and reliable probabilities give different weights to the overall expected ability. For example, if the precision of prediction probability τ_0 is much larger than the precisions of reliance probabilities, then the ability will be dominantly determined by the level of prediction probability.

In the second example, a simple graph with 11 nodes is shown in Fig. 4. Different from the first example, the prediction and reliance probabilities are deterministic values instead of random ones, which helps understand their effects on ability. Two scenarios are studied as follows.

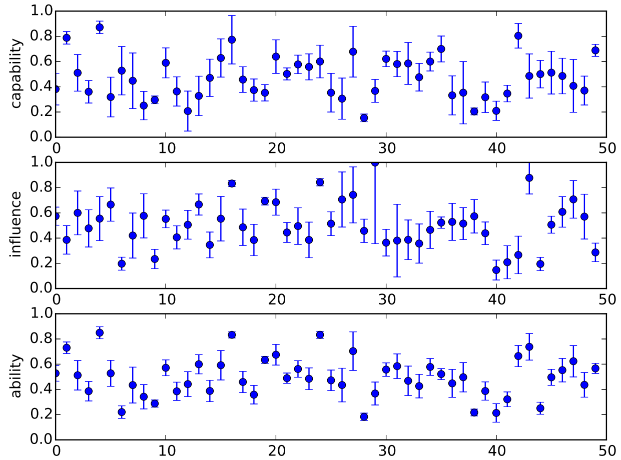


Fig. 3: The perceived capabilities, influences, and overall abilities of the 50 nodes in Fig. 2.

In the first scenario, shown in Fig. 5a, the mean values of prediction probabilities for all nodes are set to be 0.5 and variances are 0.3. The means of P-reliance probabilities for all edges are 0.9 and Q-reliance probabilities are 0.1. The variances of all reliance probabilities are 0.1. This is the scenario that the individual node's sensing and prediction capability is limited, the nodes work collaboratively, and reasoning and decision making rely very much on the communication between nodes. It is seen that nodes who send more information to others tend to have higher rankings of ability. Node 4 has the highest number of information outflows, thus is more influential and has the highest ability level. It is the most trustable node in terms of ability. In contrast, node 5 does not send information to others and is the least trustable. It is also seen that the variance of second-order ability is less than that of first-order ability. The extent of variance reduction is also related to how much information exchange is done with others. The variance associated with node 5 is not reduced in the second-order ability, since it does not have any influence to others.

A different scenario is illustrated in Fig. 5b. In this case, the means of prediction probabilities are 0.9 and variances are 0.1. The means of P-reliance probabilities are 0.5 and Q-reliance probabilities are 0.5. The variances of reliance probabilities are 0.3. It is a scenario that nodes are highly independent. Their decision makings mostly rely on own sensing and prediction capability. The other nodes' inputs are not as influential as in the first scenario. In this case, the nodes that try to be influential such as node 4 are not deemed to be trustable. Node 5 on the other hand has a higher level of ability, even though its variance is large. The ability of a node is mainly determined by its own prediction probability. Information exchange more or less only affect the variance of ability. In general, the variances of the nodes in the second scenario are higher than the ones in the first scenario.

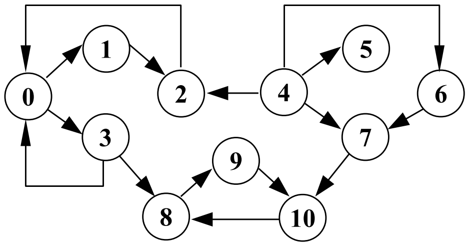
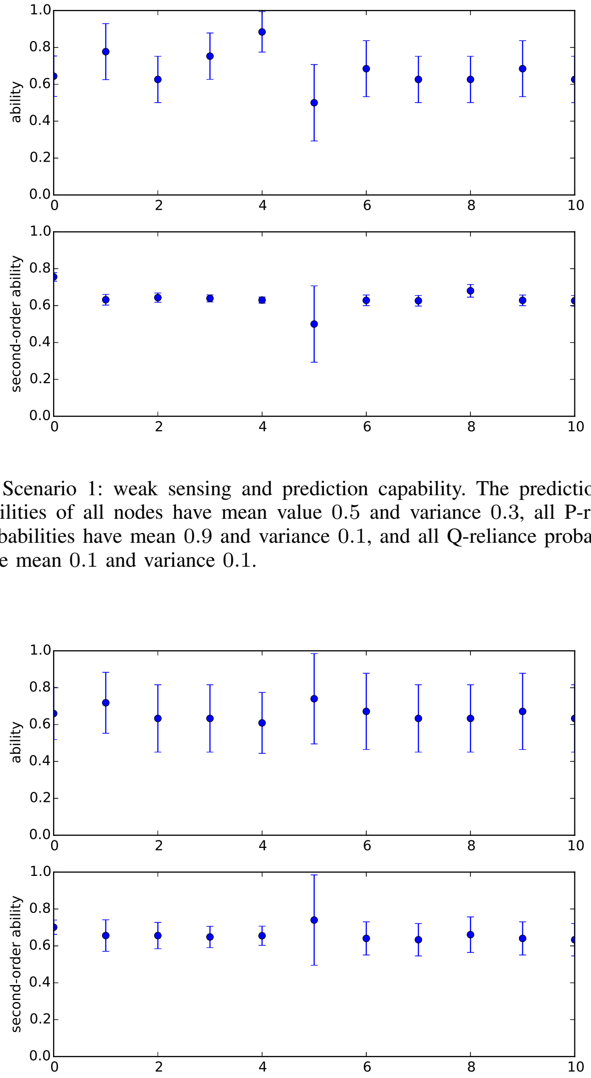


Fig. 4: A simple graph with 11 nodes



(a) Scenario 1: weak sensing and prediction capability. The prediction probabilities of all nodes have mean value 0.5 and variance 0.3, all P-reliance probabilities have mean 0.9 and variance 0.1, and all Q-reliance probabilities have mean 0.1 and variance 0.1.

(b) Scenario 2: strong sensing and prediction capability. The prediction probabilities of all nodes have mean value 0.9 and variance 0.1, all P-reliance probabilities have mean 0.5 and variance 0.3, and all Q-reliance probabilities have mean 0.5 and variance 0.3.

Fig. 5: The first- and second-order abilities of nodes in the model of Fig. 4 in two scenarios.

IV. Benevolence

Benevolence is a measure of the trustor's belief that how likely the trustee is motivated to do good to trustor, instead of for its own benefit. It captures the intention and motivation

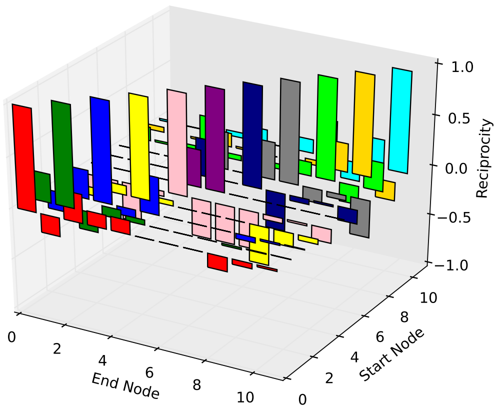
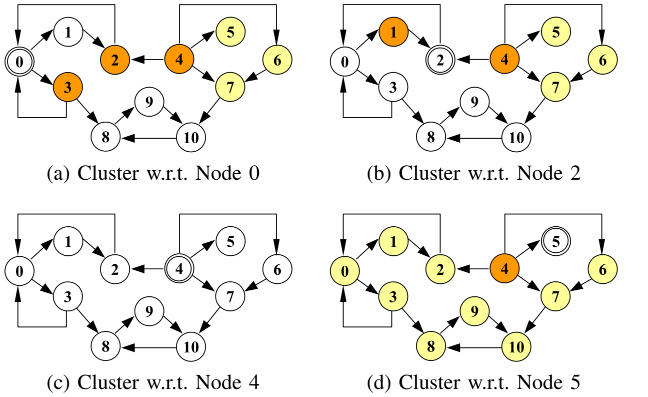


Fig. 6: The pairwise deterministic reciprocities for the graph model in Fig. 4.



(a) Cluster w.r.t. Node 0

(b) Cluster w.r.t. Node 2

(c) Cluster w.r.t. Node 4

(d) Cluster w.r.t. Node 5

Fig. 7: Trustworthy node clusters formed according to pairwise reciprocities with respect to (w.r.t.) Nodes 0, 2, 4, and 5 respectively. Heavily shaded nodes are the most trustworthy ones, lightly shaded nodes are neutral, and unshaded nodes are the least trustworthy ones.

of the trustee. The degree of benevolence is low if the motivation is originated from ergocentric gain, and high from mutual benefits. Benevolence between individuals is critical for information sharing. Without such aspect of trust, large-scale data sharing in CPS networks is not possible. *Reciprocity* is proposed here to measure the extent that the partners whom we share information with reciprocally share information with us. There are also some other characteristics associated with benevolence such as loyalty and dependability. *Motive* as the second metric proposed here is to measure the level of good intention and motivation for interactions within the community.

A. Deterministic reciprocity

The pairwise deterministic *reciprocity* of node j with respect to node i, $r_{i,j}$, is measured by the shortest topological distance, in terms of the number of hops in the network that node j shares information with node i, as

$$r_{i,j} = \exp(-h_{j\to i}) - \exp(-h_{i\to j}) + \exp(-h_{i\to j} - h_{j\to i}), \tag{36}$$

where $h_{j\to i}$ is the minimum number of hops or the shortest topological distance for information flow from node j to node i. Note that $h_{i\to i} = 0$ and $r_{i,i} = 1$. When $i \neq j$, $\partial r_{i,j}/\partial h_{j\to i} < 0$ and $\partial r_{i,j}/\partial h_{i\to j} > 0$. Hence, when the topological distance from node j to node i increases, the reciprocity of node j with respect to node i reduces. On the other hand, increasing the topological distance from node i to node j would increase the reciprocity of node j with respect to node i.

The calculation of reciprocity is straightforward. For instance, in the simple graph in Fig. 4, $r_{0,1} = e^{-2} - e^{-1} + e^{-3} = -0.18276$, $r_{0,2} = e^{-1} - e^{-2} + e^{-3} = 0.28233$, and $r_{0,5} = e^{-\infty} - e^{-\infty} + e^{-\infty} = 0$. All calculated pair-wise reciprocity results are shown in Fig. 6.

Based on the degree of reciprocity, the trustworthy levels in terms of reciprocity can be ranked and clustered. For example, among all nodes from the perspective of node 0, three clusters can be formed. $\mathcal{T}_0 = \{2, 3, 4\}$ is the most trustworthy group of nodes with positive levels of reciprocity, $\mathcal{N}_0 = \{5, 6, 7\}$ is the neutral group with zero reciprocity, and $\mathcal{U}_0 = \{1, 8, 9, 10\}$ is the least trustworthy group with negative reciprocity. The network of node 0, denoted as $\mathcal{G}_0 = \{\mathcal{T}_0, \mathcal{N}_0, \mathcal{U}_0\}$, which is shown in Fig. 7a. The subgraph formed by $\mathcal{T}_1$ and $\mathcal{N}_1$ can be regarded as the trustable network for node 0 with the criterion of reciprocity. Similarly the respective trustable networks of nodes 2, 4 and 5 are shown in Figs. 7b, 7c, and 7d . The trustable networks for nodes 4 and 5 are two extreme cases. The one for node 4 does not include any other node, since node 4 does not rely on any other's information for its decision making. In contrast, the one for node 5 includes all nodes in the network, since node 5 does not give influence to any other nodes during their decision making process. Also notice that trust relationships are not necessarily mutual. It is seen in Figs. 7a and 7b that node 2 is included in the trustable network for node 0, whereas node 0 is not included in the one for node 2.

B. Perception of reciprocity

With the further consideration of reliance probabilities as weights of edges in probabilistic graphs, the expected value of the perceived reciprocity of node j with respect to node i is calculated as

$$\mathbb{E}(R_{i,j}) = D_{KL}(p_{i\to j}||p_{j\to i}) - D_{KL}(p_{j\to i}||p_{i\to j}) + b_0, \tag{37}$$

where $D_{KL}(P||Q)$ is the Kullback-Leibler divergence from probability distribution Q to P, $p_{j\to i} = \Pi_{j\to i} p_{ab}$ is the product of all P-reliance probabilities where information flows through along the shortest path from node j to node i, and p_{ab} corresponds to the P-reliance probability from node a to node b along the path. Similarly, we have $p_{i\to j} = \Pi_{i\to j} p_{cd}$. When path $j \to i$ does not exist, the principle of maximum entropy is applied, thereof $p_{i\to j} = 0.5$. b_0 is a reference threshold of neutral value, which is predetermined such that $\mathbb{E}(R_{i,j}) > b_0$ when node j has a high reciprocity with respect to node i, and $\mathbb{E}(R_{i,j}) < b_0$ otherwise. To make the value range of reciprocity be between 0 and 1, the reference threshold is typically set as $b_0 = 0.5$. Additional scaling can be applied if necessary to keep the value range. Notice that $\mathbb{E}(R_{i,i}) = b_0$ because $D_{KL}(p_{i\to i}||p_{i\to i}) = 0$.

When the P-reliance probabilities of all nodes in Fig. 4 take the same value, the expected value of perceived reciprocity can be calculated based on Eq.(37). With $b_0 = 0.5$ as the reference, all nodes can be similarly grouped to $\{\mathcal{T}, \mathcal{N}, \mathcal{U}\}$. The results are the same as the previous ones shown in Fig. 7 based on deterministic reciprocity. The two metrics are consistent.

The variance of the perceived reciprocity can be calculated from the variances of P-reliance probabilities. Assuming the independence between the perceptions of P-reliance probabilities, the variance will be associated with the high-dimensional Gaussian distribution formed by these perceptions.

High-dimensional Gaussian distributions are costly to calculate and use. If there are m hops in the path from node j to node i, the variance associated with $p_{j\to i}$ will be an m-dimensional Gaussian distribution. To simplify the calculation for ease of application, a one-dimensional Gaussian distribution is used here for estimation purpose. The variance associated with the perceived reciprocity is conservatively estimated as

$$\mathbb{V}(R_{i,j}) = \min(\sum_{j\to i} \tau_{ab}^{-1} + \sum_{i\to j} \tau_{cd}^{-1}, V_{max}), \tag{38}$$

where τ_{ab} and τ_{cd} are the precisions associated with the P-reliance probabilities along paths $j \to i$ and $i \to j$ respectively, and V_{max} is the theoretical maximum value of variance. As discussed in Section III-F, for a value range from 0 to 1 as probability, an upper bound of variance is around 0.5. The theoretical limit can be $V_{max} = 1.0$. When a path $j \to i$ does not exist, the associated variance is set to be V_{max}. At the same time, $\mathbb{V}(R_{i,i}) = 0$.

C. Motive

Motive is to measure the motivation and intention of information sharing in a community. A high level of motive for a node indicates that it shares high-quality information with neighbors for the purpose of improving the overall functionality and performance of the community, whereas a low level of motive shows an ergocentric purpose instead of community-oriented benefit.

In the context of probabilistic graph model, the expected value of the perceived *motive* of node j is defined as

$$\mathbb{E}(M_j) = {p_j}^{d_j}, \tag{39}$$

where p_j is the prediction probability associated with node j, and $d_j = |\mathcal{D}_j|$ is the number of destination nodes with respect to node j. The baseline of motive ($M_j = 1$) is when the node has no destination nodes and does not share information with others. Compared to those sharing accurate predictions with others, a node sharing low-quality predictions with others tends to have a lower level of motive. Particularly, the more neighboring nodes it shares inaccurate predictions with, the less trustable the node is. In this case, the expected value of motive reduces quickly for low p_j as d_j increases.

The variance associated with the perceived motive of node j is related to the precision τ_j of the perceived prediction probability p_j as

$$\mathbb{V}(M_j) = \tau_j^{-1}. \tag{40}$$

D. Overall benevolence

With the considerations of both reciprocity and motive, the expected overall benevolence perception of node j respect to node i is

$$\mathbb{E}(B_{i,j}) = \frac{\mathbb{V}^{-1}(R_{i,j})\mathbb{E}(R_{i,j}) + \mathbb{V}^{-1}(M_j)\mathbb{E}(M_j)}{\mathbb{V}^{-1}(R_{i,j}) + \mathbb{V}^{-1}(M_j)}, \tag{41}$$

according to Bayes' rule. The variance associated with the perception is

$$\mathbb{V}(B_{i,j}) = (\mathbb{V}^{-1}(R_{i,j}) + \mathbb{V}^{-1}(M_j))^{-1}. \tag{42}$$

Notice that $\mathbb{E}(B_{i,i}) = b_0$ and $\mathbb{V}(B_{i,i}) = 0$.

In dynamic systems, the perception of benevolence can be updated if any new information becomes available. The perception update is from the update of probabilities similar to Eqs.(34) and (35) according to Bayes' rule,.

V. Integrity

Integrity is associated with the perceived characteristics of reliability, predictability, honesty, and consistency. Integrity is a relatively well studied topic in the context of cyber security. It is essential to protect the operation of CPSs and the networks. The quantification of integrity needs to consider the risk of deception attacks and replay attacks. In deception attack, adversary or compromised nodes send false information such as incorrect measurement, incorrect time of measurement, incorrect metadata (e.g. who measured the data), etc. to others. In replay attack, data transmitted between nodes are intercepted or delayed so that the decisions of the receiving nodes are maliciously manipulated.

The perception of integrity about a node can be measured from historical performance and behavior data of the node. The statistics of how often the information shared by the partner with different parties or at different time periods is inconsistent, and how the partner is rated or ranked by information consumers. Specifically data of how often the prediction from a node is changed or flipped as well as how often miscommunication occur (e.g. predicting False but sending True) can be collected.

Suppose that the prior belief of integrity for node j is

$$\mathbb{E}(I_j) = g_j \tag{43}$$

with imprecision or variance

$$\mathbb{V}(I_j) = \omega_j^{-1}. \tag{44}$$

The likelihood that node j is free from deception attack and maintains its integrity can be quantified as the deviation between its state variable value and the average state variable value in its neighborhood Ω_j where the same quantity of interest is measured and detected, as

$$P(x_j|x_{i\in\Omega_j}) = g_j^S \propto \exp\left[-\frac{(x_j - \hat{x}^{(j)})^2}{2\sigma_x^2}\right], \tag{45}$$

where $\hat{x}^{(j)} = \frac{1}{|\Omega_j|}\sum_{i\in\Omega_j} x_i$ is the average prediction of the neighboring nodes with respect to node j, and σ_x^2 indicates the natural variation between sensing units as the random error.

Based on Bayes' rule, the perception of integrity about node j can be updated to

$$\mathbb{E}(I_j|x_{i\in\Omega}) = \frac{g_j\omega_j + g_j^S\sigma_x^{-2}}{\omega_j + \sigma_x^{-2}}, \tag{46}$$

when new information about the behaviors of nodes is obtained.

The likelihood function can also be formulated to incorporate the temporal factor. If the $x_j(t_k)$ denotes the predicted state value by node j at time t_k, the likelihood can be extended to capture the node's own consistency as

$$P\left(x_j(t_k)|x_j(t_{k-1},\ldots,t_0)\right) = g_j^T \propto \exp\left[-\frac{(x_j(t_k)-\overline{x}_j)^2}{2\sigma_x^2}\right] \tag{47}$$

where $\overline{x}_j = \frac{1}{k}\sum_{i=0}^{k-1} x_j(t_i)$ is the average value of previous predictions by node j at time stamps from t_0 to t_{k-1}. The perception of integrity can be similarly updated to

$$\mathbb{E}(I_j|x_{i\in\Omega}, x_j(t_{k-1},\ldots,t_0)) = \frac{g_j\omega_j + g_j^S\sigma_x^{-2} + g_j^T\sigma_x^{-2}}{\omega_j + 2\sigma_x^{-2}}. \tag{48}$$

The proposed trustworthiness quantification based on perceptions of ability, benevolence, and integrity can be susceptible to trust attacks. Similar to other reputation or recommendation based quantifications, perception can be influenced and manipulated. Therefore it is susceptible to attacks on recommendation systems, particularly for those with centralized reputation management, such as *self-promoting attack* (provide good recommendations to selves), *ballot-stuffing attack* (provide good recommendations to bad nodes), *bad-mouthing attack* (provide bad recommendations to good ones), and *whitewashing attack* (disappear and rejoin the community with new identity) [47]. These attacks can change the perceptions of individuals. In addition, one's perception can also influence other's perceptions in a human society. Thus in a *perception attack*, attackers can broadcast their manipulated negative or positive perception and seek the ripple effect outside the recommendation system. Nevertheless, a perception based trust management system is less vulnerable to sudden attacks than a reputation based system. Because perception or belief update based on Bayes' rule has been known as a gradual process, particularly with the involvement of variance or imprecision associated with perception. Whenever new evidence arrives, the change of perception is affected by the precision associated with new evidence. Imprecise evidence brings little change to the perception. Even with precise ones, it takes iterations of updates to sway the perception of ability, benevolence, or integrity.

VI. Performance Evaluation

To evaluate the performance of the proposed trust metrics, simulation tests similar to Refs.[46], [47] are performed, where attacks on both sensing and communication capabilities of CPS nodes are simulated. The details of the tests are described as follows.

In the first test, the ability metric is evaluated. The simple network in Fig. 4 is used. Attacks on node 0 are simulated, where the prediction capability and communication are affected. Without loss of generality, it is assumed that the prediction and reliance probabilities of the nodes follow Dirichlet distributions. The initial or prior prediction probabilities of all 11 nodes are assumed to have a mean value of 0.9 and a variance of 8.99×10^{-5}, which was a result of 900 positive predictions out of 1000 from historical data. Again, the probabilities could also be subjective beliefs if no data are available. Similarly, the initial P-reliance probabilities of all nodes have mean 0.9 and variance 8.99×10^{-5}, and the initial Q-reliance probabilities of all nodes have mean 0.5 and variance 2.49×10^{-4}. When the sensing and reasoning units of node 0 are attacked, false predictions arise from the node. As a result, its perceived prediction probability is affected. Similarly, when the communication unit of the node is attacked, false predictions are sent to neighboring nodes. The perceived reliance probabilities are affected. During simulation, for each of the first 30 iterations, 99 false predictions out of a total of 100 samples are generated in a *light* attack scenario, and the false information is used to update the prediction probability of node 0 according to Bayes' rule. With the updated prediction probability, the overall ability of node 0 is calculated. Two policy update cases are tested. The first case is with *long memory*, where the node keeps the complete data history. For belief update in each iteration, the ratio of correct prediction (i.e. $p_j^{(new)}$ in Eq.(34)) is calculated from the number of correct predictions out of all samples accumulated from previous iterations. In the second case of *short memory*, the node does not remember the historical data. The ratio $p_j^{(new)}$ is calculated with the correct predictions within the particular iteration, which is $1/100$. A similar attack on communication is also simulated where P-reliance probabilities between node 0 and its neighboring nodes (1, 2, and 3) are affected and reduced toward zero when false information is shared among neighbors. The update is also based on Bayes' rule. In a *heavy* attack scenario, 999 false predictions out of 1000 are generated to update prediction probability of node 0 and P-reliance probabilities on edges between node 0 and its neighbors. The Bayesian update is similar to the first scenario. The changes of node 0's ability in these eight different attack modes are shown in Fig. 8a. The malicious attacks stop at the 30^{th} iteration, and recovery is followed for the next 60 iterations. For each iteration of recovery, the original prediction probability of 0.9 and P-reliance probability 0.9 are applied to update the perception of ability. That is, 900 correct predictions out of 1000 are applied to update the perceived prediction probability. The update policies of long and short memory are similarly applied. It is seen in Fig. 8a that the ability trustworthiness is gradually reduced during the attacks. The reduction speed in the heavy attack scenario is higher than the one in the light attack scenario. The re-establishment of trust during recovery is slower than when it is damaged, which is similar to natural human behavior. In addition, the belief update policy and memory of history affect the speed of change. Nodes with longer memory tend to be more resilient and less susceptible to attacks.

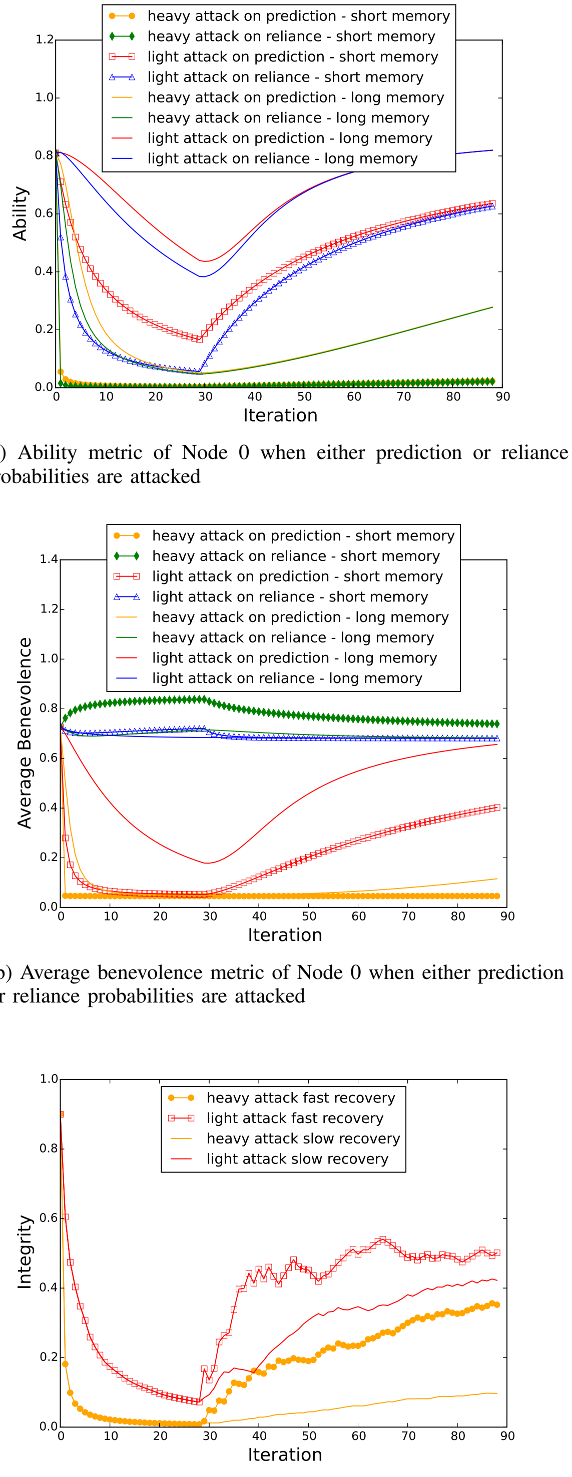
In the second test, the same attacks are simulated in the same network with average benevolence as the metric. The results are shown in Fig. 8b. The average benevolence of node 0 is calculated as $(1/11)\sum_{i=0}^{10} B_{i,0}$. It is seen that both ability and benevolence metrics are sensitive enough to detect attack. At the same time, the responses always have delays, instead of changing abruptly, which provides necessary robustness and stability to protect trust metrics against attacks. Similarly, nodes with longer memory are more resilient. During attacks, the damage of reputation is less severe and can be recovered faster. Nodes with shorter memory are more sensitive to detect attacks.

In the third test, the prior integrities of the nodes are assumed to follow a Dirichlet distribution with expected values of 0.9 and 0.1, from which the probabilities of accurate prediction are generated randomly. Random samples of predicted state variable values are generated. When node 0 is attacked in the *heavy* mode, 1 out of 1000 sample predictions is accurate and the rest of 999 are false predictions. The mean and variance of the state variable sample values from node 0 and those from its spatial neighbors (nodes $1, 2, 3, 4$) are used to estimate the likelihood in Eq.(45) and update the integrity based on Eq.(46). In the *light* attack mode, 1 out of 100 samples from node 0 is accurate, and the rest of 99 are false predictions. The changes of integrity metric are shown in Fig. 8c. During recovery, two modes are tested. For fast recovery mode, the sample size is 1000 for each iteration, out of which 900 predictions are accurate according to the original probability 0.9. For slow recovery, 100 samples are drawn during each iteration and 90 of them are accurate predictions. It is seen that the pattern of integrity attack is similar to those for ability and benevolence. Note that long memory scenario is not considered here since accumulated effect of integrity is not important from security perspective. When attacks occur, the metric of integrity drops quickly, which shows that the metric is sensitive. The quickly dropped integrity value become flat as it comes close to zero. During recovery, fast recovery helps increase the integrity value at a speed faster than slow recovery. Again, the drop of trust is rapid during attack, which is due to the nature of Bayesian update. The recovery of trust is much slower than when it is being damaged.

VII. Strategic Network Design

A strategic network for a CPS node is a trustable network or society that the node can collaborate and work with. The design of a strategic network with respect to a reference node is to maximize the expected utility by choosing the optimum combination of n (out of N) nodes to form its society. Different definitions of utility function U could lead to different strategic networks. In this section, two design criteria as utilities are described. In the first criterion, the utility function is defined as the node's reciprocity and benevolence. The optimum design is found by maximizing the individual's expected benevolence. In the second criterion, the utility function is individual's ability.

To exhaustively search the optimum combination of nodes is a NP-hard problem. An efficient alternative is searching with greedy algorithms. Starting from the source node, greedy algorithms selectively add nodes sequentially if the objective function value increases. Here, a breadth-first search (BFS) greedy algorithm is developed to demonstrate the design optimization approach. The algorithm is listed in Alg. 1. The algorithm adds nodes one-by-one using a BFS strategy. Star-



(a) Ability metric of Node 0 when either prediction or reliance probabilities are attacked

(b) Average benevolence metric of Node 0 when either prediction or reliance probabilities are attacked

(c) Integrity metric of Node 0 when predictions are attacked

Fig. 8: Trustworthy metrics of ability and benevolence when Node 0 in Fig. 4 is attacked. Recovery occurs at Iteration 30.

ting from the reference node i, the subgraph $\mathcal{G}^{(i)} = (\mathcal{V}^{(i)}, \mathcal{E}^{(i)})$ is constructed and updated sequentially by inserting additional nodes one at a time from the neighboring nodes of the existing subgraph. For each iteration, if the utility for the subgraph is non-decreasing from the previous iteration, the new node is accepted. The greedy algorithm allows for quick formation of the strategic network, but obviously could potentially miss the true optimum solution. Other optimization algorithms for combinatorial problems can also be applied.

Algorithm 1 The breadth-first search greedy algorithm

```
$\mathcal{V}^{(i)} = \emptyset$,$\mathcal{E}^{(i)} = \emptyset$;
$iteration = 1$
$Queue.append(i)$;
while ( $iteration < limit$ and $Queue.IsNotEmpty$ ) do    ▷ main iterations of search
    $j = Queue.pop()$;
    $\mathcal{V}^T = \mathcal{V}^{(i)} \cup \{j\}$;
    Construct $\mathcal{G}^T = \{\mathcal{V}^T, \mathcal{E}^{(i)}\}$ where $\mathcal{E}^{(i)} \subseteq \mathcal{E}$;
    Calculate $U[iteration]$;
    $\Delta U = U[iteration] - U[iteration - 1]$
    if $\Delta U \geq 0$ then
        Update $\mathcal{G}^{(i)} = \mathcal{G}^T$;
    end if
    for $\forall k \in$ neighbors of node j do
        $Queue.append(k)$;
    end for
    $iteration = iteration + 1$
end while
```

A. Deterministic reciprocity criterion

Individual reciprocities typically have conflict of interest. One node shares information with others without receiving information from others reciprocally tends to have lower perception of trust about its collaborators. It is individual nodes' interest to receive information as much as possible from others. At the same time, the willingness to share with others can be dampened without reciprocal treatment from others. In the optimum network with respect to a reference node, nodes are selected based on the individuals' reciprocities. For a 'selfish' approach, the reciprocity of the reference node is the only consideration. For an 'altruistic' approach, the reciprocities of nodes other than the reference node are only considered. Between the above two extreme scenarios, the weighted average of reciprocities among all nodes can be taken.

The utility as the overall weighted average reciprocity in the society with respect to node i is defined as

$$U^{(i)} = \sum_{j \in V^{(i)}} w_j \bar{r}_j \tag{49}$$

where $\bar{r}_j = (1/n_j) \sum_{k \in V^{(i)}} r_{j,k}$ is the average reciprocity of node j among its n_j neighboring nodes in the society of node i. The average reciprocity of a node indicates how well other nodes treat it reciprocally. Therefore, the utility function in Eq.(49) is the weighted average of all nodes in the society formed by node i. Determining the self-interest weights w_j's has an effect on how much emphasis on the reference node's benefit verses other nodes when forming strategic partnerships.

TABLE I: Strategic networks of different reference nodes in Fig. 4 with deterministic reciprocity as the criterion based on different self-interest weights

Refer.	Strategic Network		
	($w_i = 1.0$)	($w_i = 0.5$)	($w_i = 0$)
0	0, 2, 4	0, 1, 2, 3, 8	0, 1, 2
1	0, 1, 3	0, 1, 3, 8	1, 2, 4
2	1, 2, 4, 7	0, 1, 2, 4, 7	0, 2, 3, 8
3	0, 2, 3, 4	0, 2, 3	0, 3, 8, 10
4	4, 7	0, 1, 2, 4, 5, 6, 7, 10	2, 4, 5, 6, 7, 10
5	4, 5, 7	4, 5, 7	4, 5
6	4, 6, 7	4, 6, 7	4, 6, 7
7	4, 6, 7	4, 6, 7, 10	7, 9, 10
8	0, 3, 8, 10	0, 2, 3, 8, 9, 10	8, 9, 10
9	3, 8, 9	3, 8, 9	7, 9, 10
10	6, 7, 9, 10	6, 7, 8, 9, 10	3, 8, 10

Three simple cases of the 11-node example in Fig. 4 are tested for demonstration. The first case is the selfish approach where the self-interest weight associated with the reference node i is one ($w_i = 1$) and others are zeros. The second case is when $w_i = 0.5$ and all others are equally weighted and all self-interest weights still sum up to one. The third one is the altruistic approach with $w_i = 0$ and all others are equally weighted and sum up to one. The resulting strategic networks from the greedy algorithms with respect to each of the nodes as reference are shown in Table I. In above three cases, the P-reliance probabilities are not considered. The reciprocities are calculated based on Eq. (36), where only the topological effect of graphs is included.

B. Benevolence criterion

A further generalization is to consider the prediction and reliance probabilities and use benevolence as the criterion. The perception of benevolence is calculated as in Eq.(41) and replaces the deterministic reciprocity in Eq.(49) for utility. Assume that all expected values of P- and Q-reliance probabilities between all nodes in Fig. 4 are 0.5, and their variances are 0.1. The expected value and variance of prediction probability for all nodes are 0.9 and 0.1 respectively. The strategic networks with respect to each of the nodes with self-interest weights $w_i = 1$ and $w_i = 0$ are listed in Table II.

C. Ability criterion

The second criterion that can be used for network optimization is ability. In addition to the prediction capability, ability also measures how influential a node is in a society. Therefore, the natural objective of a node to build a strategic network around itself is to maximize its influence within the network if its prediction capability is fixed.

The utility based on the second-order ability in Eq.(32) with respect to node i can be defined as

$$U^{(i)} = \mathbb{E}^{(2)}(A_j(\theta|+,-)) \tag{50}$$

TABLE II: Strategic networks for different reference nodes in Fig. 4 with two self-interest weights ($w_i = 1$ and $w_i = 0$) when benevolence is used as the criterion

Refer.	Strategic Network	
	($w_i = 1.0$)	($w_i = 0$)
0	0, 1, 2, 4	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
1	0, 1, 2, 4, 5, 6, 7	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
2	0, 1, 2, 3	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
3	0, 1, 2, 3, 4, 5, 6, 7	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
4	0, 2, 3, 4, 5, 6	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
5	0, 2, 3, 4, 5, 5	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
6	0, 2, 3, 4, 6, 7	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
7	0, 3, 7, 8, 10	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
8	7, 8, 9, 10	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
9	0, 2, 3, 4, 5, 7, 8, 9, 10	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
10	0, 1, 2, 3, 4, 7, 8, 9, 10	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

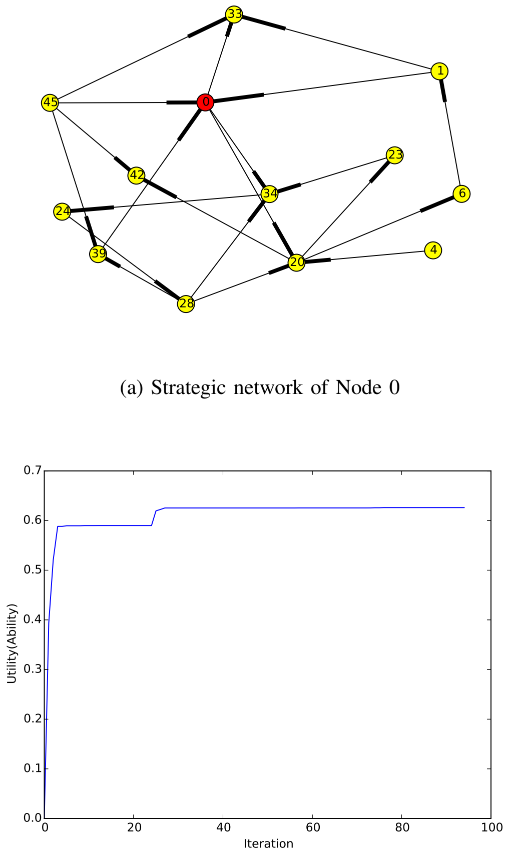
TABLE III: Strategic networks for different reference nodes with ability as criterion

Refer.	Strategic Network	Abilities
0	0, 1, 2, 4	0.765, 0.689, 0.597, 0.0
1	0, 1, 2, 4	0.765, 0.689, 0.597, 0.0
2	0, 1, 2, 3, 8	0.720, 0.689, 0.645, 0.645, 0.0
3	0, 1, 2, 3	0.681, 0.671, 0., 0.689
4	0, 2, 4, 5, 6, 7, 10	0.5, 0.517, 0.645, 0.5, 0.628, 0.517, 0.5
5	4, 5	0.0, 0.5
6	6, 7, 9, 10	0.689, 0.517, 0.0, 0.5
7	3, 7, 8, 10	0.0, 0.689, 0.5, 0.517
8	7, 8, 9, 10	0.0, 0.684, 0.684, 0.684
9	3, 8, 9, 10	0.0, 0.684, 0.684, 0.684
10	3, 8, 9, 10	0.0, 0.684, 0.684, 0.684

The strategic network of node i can be obtained by finding the network where the ability of the reference node is maximized. In the simple example in Fig. 4, the strategic network with respect to each of the 11 nodes are obtained by applying the greedy algorithm, and the results are shown in Table III. The abilities corresponding to the nodes in the final networks are also listed.

The optimization process is also applied to the random graph model in Fig. 2. The optimum network with respect to node 0 is shown in Fig. 9a, and the evolution of utility during the search is shown in Fig. 9b. The ability of node 0 is 0.62604317 in its final strategic network. Notice that the proposed greedy optimization algorithm is heuristic and does not guarantee a global optimum solution. The solution can be searching path dependent. When a node is selected to be included in the solution, it may affect which nodes to be included in the following searching steps. That is the reason that some plateaus are observed in 9b. There might be no improvement for quite a few iterations without breakthrough once some nodes are selected in the optimum network.

Higher-order abilities instead of the second-order one in Eq. (50) can be similarly used as the criterion in network optimization. When higher-order abilities are used, the influence of a node in the network gains more weights in calculating abilities, which is also emphasized more in obtaining the optimum network.



(a) Strategic network of Node 0

(b) Utility evolution during the searching process

Fig. 9: Trustworthy network of Node 0 from random graph in Fig. 2 where the second-order ability is applied as the utility for optimization.

Note that the integrity of nodes is not used in designing a node's strategic network. Because the integrity of an individual node is mostly independent from the topological relationship between those nodes. The network rarely has effects on how an individual node behaves or how it is compromised when attacked. The goal of the strategic network with respect to a reference node is building a trustable community which the reference node can rely on and work with. Nevertheless, if the integrity of networks instead of individual nodes is concerned and the goal is to maintain the integrity of a networked system, the design optimization is relatively straightforward and is to increase the size of the network as much as resources allow. Introducing redundancy can increase the reliability, resilience, and thus the integrity of the system. This can also be seen in the perception based integrity measure in Eq.(48), where the large variation among nodes, caused by individual compromised nodes, helps reduce the impact of the individual's swing and keeps the overall perception stable.

VIII. Concluding Remarks

In this paper, a perception based trust framework is described in order to include human user aspects in trust. The trustworthiness of CPS nodes in a networked environment is quantified by three independent metrics, including ability, benevolence, and integrity. Ability indicates how capable a CPS node is to provide accurate sensing, reasoning, and prediction, and how influential a node is in affecting other's decision making process. Benevolence measures the motivation of information sharing and how much reciprocity a node may receive from its neighbors during information and data exchange. Integrity shows the level of reliability, predictability and security of a node in the network.

The three quantitative metrics can be obtained objectively from the statistical data of performance as well as perceptual reputation, including prediction and reliance probability values. The perceptual models can also be applied when beliefs are elicited from experts as subjective probabilities. The calculation of trustworthiness metrics is all based on the Bayesian approach. The only assumption made in the model is the Gaussian distributions of perceptions. Therefore, the generality of the metrics is maintained, and human and social behaviors can be captured.

In the current graph model setting, the state variables are simplified to binary, where only two values are taken. In more generic analog systems, sensing and reasoning are based on continuous variables. The continuous variables with analog values need to be digitalized. It has been shown that the probabilistic graph model can be easily extended to handle multi-valued state variables. The probabilistic metrics for ability can also be generalized so that the expected values of perceptions become multi-valued functions or distributions. Then the same Bayesian belief update can be applied. The Kullback-Leibler divergence in reciprocity metric is general and can be directly applied to multi-valued probability distributions. Nevertheless, the implementation of such general framework and the assessment of computational complexity need to be included in future work. Further performance evaluation is also needed.

In addition, the proposed modeling method can be regarded as a mesoscale model of networks, where detailed network communication protocols between nodes is not considered, nor detailed sensing and control mechanisms within each node. The mesoscale model needs to be compared with fine-grained bottom-up models in the future. For instance, hidden variables may be included in detailed models for measurement and internal reasoning in each node. Control variables can also be introduced for actuation. Furthermore, network design optimization based on multiple objectives requires more studies, given that multiple metrics are used in quantification.

ACKNOWLEDGMENT

This work is supported in part by National Science Foundation under grant CMMI-1663227.

REFERENCES

[1] I. Horvath and B. H. Gerritsen, "Cyber-physical systems: Concepts, technologies and implementation principles," in *Proceedings of The 9th International Symposium on Tools and Methods of Competitive Engineering (TMCE2012)*, pp. 19–36, 2012.

[2] M. Grimm, R. Anderl, and Y. Wang, "Conceptual approach for multi-disciplinary cyber physical systems design and engineering," in *Proceedings of The 10th International Symposium on Tools and Methods of Competitive Engineering (TMCE2014)*, pp. 61–72, 2014.

[3] S. Grabner-Kräuter and E. A. Kaluscha, "Empirical research in on-line trust: a review and critical assessment," *International Journal of Human-Computer Studies*, vol. 58, no. 6, pp. 783–812, 2003.

[4] J. Sabater and C. Sierra, "Review on computational trust and reputation models," *Artificial intelligence review*, vol. 24, no. 1, pp. 33–60, 2005.

[5] D. Artz and Y. Gil, "A survey of trust in computer science and the semantic web," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5, no. 2, pp. 58–71, 2007.

[6] S. N. L. C. Keung and N. Griffiths, *Trust and reputation*, pp. 189–224. London: Springer, 2010.

[7] J. Golbeck *et al.*, "Trust on the world wide web: a survey," *Foundations and Trends® in Web Science*, vol. 1, no. 2, pp. 131–197, 2008.

[8] Y. Ruan and A. Durresi, "A survey of trust management systems for online social communities–trust modeling, trust inference and attacks," *Knowledge-Based Systems*, vol. 106, pp. 150–163, 2016.

[9] J. Sabater and C. Sierra, "Reputation and social network analysis in multi-agent systems," in *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: Part 1*, pp. 475–482, ACM, 2002.

[10] H. Jameel, L. X. Hung, U. Kalim, A. Sajjad, S. Lee, and Y.-K. Lee, "A trust model for ubiquitous systems based on vectors of trust values," in *Multimedia, Seventh IEEE International Symposium on*, pp. 6–pp, IEEE, 2005.

[11] T. D. Huynh, N. R. Jennings, and N. R. Shadbolt, "An integrated trust and reputation model for open multi-agent systems," *Autonomous Agents and Multi-Agent Systems*, vol. 13, no. 2, pp. 119–154, 2006.

[12] S. Songsiri, "Mtrust: a reputation-based trust model for a mobile agent system," in *International Conference on Autonomic and Trusted Computing*, pp. 374–385, Springer, 2006.

[13] S. H. Houmb, I. Ray, and I. Ray, "Estimating the relative trustworthiness of information sources in security solution evaluation," in *iTrust*, pp. 135–149, Springer, 2006.

[14] T. DuBois, J. Golbeck, J. Kleint, and A. Srinivasan, "Improving recommendation accuracy by clustering social networks with trust," *Recommender Systems & the Social Web*, vol. 532, pp. 1–8, 2009.

[15] K. Nordheimer, T. Schulze, and D. Veit, "Trustworthiness in networks: A simulation approach for approximating local trust and distrust values," in *IFIP International Conference on Trust Management*, pp. 157–171, Springer, 2010.

[16] G. Pitsilis, X. Zhang, and W. Wang, "Clustering recommenders in collaborative filtering using explicit trust information," in *IFIP International Conference on Trust Management*, pp. 82–97, Springer, 2011.

[17] L.-H. Vu and K. Aberer, "Effective usage of computational trust models in rational environments," *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, vol. 6, no. 4, p. 24, 2011.

[18] X. Ma, H. Lu, and Z. Gan, "Improving recommendation accuracy by combining trust communities and collaborative filtering," in *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pp. 1951–1954, ACM, 2014.

[19] T. Beth, M. Borcherding, and B. Klein, "Valuation of trust in open networks," *Computer SecurityESORICS 94*, pp. 1–18, 1994.

[20] B. Yu and M. Singh, "A social mechanism of reputation management in electronic communities," pp. 154–165, 2000.

[21] S. Lee, R. Sherwood, and B. Bhattacharjee, "Cooperative peer groups in nice," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 2, pp. 1272–1282, IEEE, 2003.

[22] D. O'Doherty, S. Jouili, and P. Van Roy, "Towards trust inference from bipartite social networks," in *Proceedings of the 2nd ACM SIGMOD Workshop on Databases and Social Networks*, pp. 13–18, ACM, 2012.

[23] K. Govindan and P. Mohapatra, "Trust computations and trust dynamics in mobile adhoc networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 279–298, 2012.

[24] R. R. Sahoo, A. R. Sardar, M. Singh, S. Ray, and S. K. Sarkar, "A bio inspired and trust based approach for clustering in wsn," *Natural Computing*, vol. 15, no. 3, pp. 423–434, 2016.

[25] M. Singh, A. R. Sardar, K. Majumder, and S. K. Sarkar, "A lightweight trust mechanism and overhead analysis for clustered wsn," *IETE Journal of Research*, pp. 1–12, 2017.

[26] X. Li, F. Zhou, and J. Du, "Ldts: A lightweight and dependable trust system for clustered wireless sensor networks," *IEEE transactions on information forensics and security*, vol. 8, no. 6, pp. 924–935, 2013.

[27] P. Zhou, S. Jiang, A. Irissappane, J. Zhang, J. Zhou, and J. C. M. Teo, "Toward energy-efficient trust system through watchdog optimization for wsns," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 613–625, 2015.

[28] Z. Chen, L. Tian, and C. Lin, "Trust model of wireless sensor networks and its application in data fusion," *Sensors*, vol. 17, no. 4, p. 703, 2017.

[29] V. B. Reddy, S. Venkataraman, and A. Negi, "Communication and data trust for wireless sensor networks using d–s theory," *IEEE Sensors Journal*, vol. 17, no. 12, pp. 3921–3929, 2017.

[30] K. S. Barber and J. Kim, "Belief revision process based on trust: Agents evaluating reputation of information sources," in *Trust in Cyber-societies* (R. Falcone, M. Singh, and Y.-H. Tan, eds.), vol. 2246, pp. 73–82, Springer, 2001.

[31] B. Yu and M. P. Singh, "Distributed reputation management for electronic commerce," *Computational Intelligence*, vol. 18, no. 4, pp. 535–549, 2002.

[32] J. Patel, W. L. Teacy, N. R. Jennings, and M. Luck, "A probabilistic trust model for handling inaccurate reputation sources," in *Trust Management* (P. Herrmann, V. Issarny, and S. Shiu, eds.), (Berlin, Heidelberg), pp. 193–209, Springer, Springer, 2005.

[33] Y. Wang, V. Cahill, E. Gray, C. Harris, and L. Liao, "Bayesian network based trust management," in *Autonomic and Trusted Computing* (L. T. Yang, H. Jin, J. Ma, and T. Ungerer, eds.), (Berlin, Heidelberg), pp. 246–257, Springer, 2006.

[34] Y. Wang, M. Li, E. Dillon, L.-g. Cui, J.-j. Hu, and L.-j. Liao, "A context-aware computational trust model for multi-agent systems," in *Networking, Sensing and Control, 2008. ICNSC 2008. IEEE International Conference on*, pp. 1119–1124, IEEE, 2008.

[35] H. Kim, H. Lee, W. Kim, and Y. Kim, "A trust evaluation model for qos guarantee in cloud systems," *International Journal of Grid and Distributed Computing*, vol. 3, no. 1, pp. 1–10, 2010.

[36] X. Li, H. Ma, F. Zhou, and X. Gui, "Service operator-aware trust scheme for resource matchmaking across multiple clouds," *IEEE transactions on parallel and distributed systems*, vol. 26, no. 5, pp. 1419–1429, 2015.

[37] V. Kant and K. K. Bharadwaj, "Fuzzy computational models of trust and distrust for enhanced recommendations," *International Journal of Intelligent Systems*, vol. 28, no. 4, pp. 332–365, 2013.

[38] K. W. Nafi, T. S. Kar, M. A. Hossain, and M. Hashem, "A fuzzy logic based certain trust model for e-commerce," in *Informatics, Electronics & Vision (ICIEV), 2013 International Conference on*, pp. 1–6, IEEE, 2013.

[39] R. Falcone, G. Pezzulo, and C. Castelfranchi, "A fuzzy approach to a belief-based trust computation," in *Workshop on Deception, Fraud and Trust in Agent Societies*, pp. 73–86, Springer, 2002.

[40] M. Alhamad, T. Dillon, and E. Chang, "A trust-evaluation metric for cloud applications," *International Journal of Machine Learning and Computing*, vol. 1, no. 4, p. 416, 2011.

[41] M. Ashtiani and M. A. Azgomi, "Trust modeling based on a combination of fuzzy analytic hierarchy process and fuzzy vikor," *Soft Computing*, vol. 20, no. 1, pp. 399–421, 2016.

[42] R. K. Chahal and S. Singh, "Fuzzy rule-based expert system for determining trustworthiness of cloud service providers," *International Journal of Fuzzy Systems*, vol. 19, no. 2, pp. 338–354, 2017.

[43] D. Chen, G. Chang, D. Sun, J. Li, J. Jia, and X. Wang, "Trm-iot: A trust management model based on fuzzy reputation for internet of things," *Computer Science and Information Systems*, vol. 8, no. 4, pp. 1207–1228, 2011.

[44] Y. B. Saied, A. Olivereau, D. Zeghlache, and M. Laurent, "Trust management system design for the internet of things: A context-aware and multi-service approach," *Computers & Security*, vol. 39, pp. 351–365, 2013.

[45] M. Nitti, R. Girau, and L. Atzori, "Trustworthiness management in the social internet of things," *IEEE Transactions on knowledge and data engineering*, vol. 26, no. 5, pp. 1253–1266, 2014.

[46] R. Chen, J. Guo, and F. Bao, "Trust management for soa-based iot and its application to service composition," *IEEE Transactions on Services Computing*, vol. 9, no. 3, pp. 482–495, 2016.

[47] R. Chen, F. Bao, and J. Guo, "Trust-based service management for social internet of things systems," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 6, pp. 684–696, 2016.

[48] H. Al-Hamadi and R. Chen, "Trust-based decision making for health iot systems," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1408–1419, 2017.

[49] U. Jayasinghe, A. Otebolaku, T.-W. Um, and G. M. Lee, "Data centric trust evaluation and predication framework for iot," in *ITU Kaleidoscope 2017, 27 November 2017 - 29 November 2017, Nanjing, China*.

[50] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Academy of management review*, vol. 20, no. 3, pp. 709–734, 1995.

[51] J. A. Colquitt, B. A. Scott, and J. A. Lepine, "Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance," *Journal of Applied Psychology*, vol. 92, no. 4, pp. 909–927, 2007.

[52] M.-H. Yang, B. Lin, N. Chandlrees, and H.-Y. Chao, "The effect of perceived ethical performance of shopping websites on consumer trust," *Journal of computer information systems*, vol. 50, no. 1, pp. 15–24, 2009.

[53] M. K. Lee and E. Turban, "A trust model for consumer internet shopping," *International Journal of electronic commerce*, vol. 6, no. 1, pp. 75–91, 2001.

[54] F. B. Tan and P. Sutherland, "Online consumer trust: a multi-dimensional model," *Journal of Electronic Commerce in Organizations (JECO)*, vol. 2, no. 3, pp. 40–58, 2004.

[55] O. B. Büttner and A. S. Göritz, "Perceived trustworthiness of online shops," *Journal of Consumer Behaviour*, vol. 7, no. 1, pp. 35–50, 2008.

[56] S. Y. Yousafzai, J. Pallister, G. R. Foxall, *et al.*, "Strategies for building and communicating trust in electronic banking: A field experiment," *Psychology & Marketing*, vol. 22, no. 2, pp. 181–201, 2005.

[57] J. Benamati, M. A. Serva, and M. A. Fuller, "Are trust and distrust distinct constructs? an empirical study of the effects of trust and distrust among online banking users," in *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on*, vol. 6, pp. 121b–121b, IEEE, 2006.

[58] T. W. Lauer and X. Deng, "Building online trust through privacy practices," *International Journal of Information Security*, vol. 6, no. 5, p. 323, 2007.

[59] S. Akter, J. D'Ambra, and P. Ray, "Trustworthiness in mhealth information services: an assessment of a hierarchical model with mediating and moderating effects using partial least squares (pls)," *Journal of the Association for Information Science and Technology*, vol. 62, no. 1, pp. 100–116, 2011.

[60] H. Chen, "The influence of perceived value and trust on online buying intention.," *Journal of Computers*, vol. 7, no. 7, pp. 1655–1662, 2012.

[61] S. Scherer and M. A. Wimmer, "Conceptualising trust in e-participation contexts," in *International Conference on Electronic Participation*, pp. 64–77, Springer, 2014.

[62] X. Li, T. J. Hess, and J. S. Valacich, "Using attitude and social influence to develop an extended trust model for information systems," *ACM sigmis database*, vol. 37, no. 2-3, pp. 108–124, 2006.

[63] M. A. Fuller, M. A. Serva, J. Baroudi, *et al.*, "Clarifying the integration of trust and tam in e-commerce environments: implications for systems design and management," *IEEE Transactions on Engineering Management*, vol. 57, no. 3, pp. 380–393, 2010.

[64] Y. Wang, "System resilience quantification for probabilistic design of internet-of-things architecture," in *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, p. V01BT02A011, ASME, 2016.

[65] Y. Wang, "Resilience quantification for probabilistic design of cyber-physical system networks," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B*, vol. 4, no. 3, p. 031006, 2018.

[66] Y. Wang, "Trustworthiness in designing cyber-physical systems," in *Proceedings of The 12th International Symposium on Tools and Methods of Competitive Engineering (TMCE2018)*, pp. 1–12, 2018.

[67] A. O'Hagan, C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley, and T. Rakow, *Uncertain judgements: eliciting experts' probabilities*. John Wiley & Sons, 2006.

PLACE
PHOTO
HERE

Yan Wang (M'07) received the B.S. degree in electrical engineering from the Tsinghua University, Beijing, in 1996, the M.S. degree in electrical engineering from the Chinese Academy of Sciences, Beijing, in 1998, and the Ph.D. degree in industrial engineering from the University of Pittsburgh, Pittsburgh, Pennsylvania, in 2003. He currently is an Associate Professor at the Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta. He is interested in engineering design, modeling and simulation, and uncertainty quantification, and has over 100 peer-reviewed publications. His research work was recognized with multiple conference Best Paper Awards from societies of ASME, IISE, and TMS, as well as a National Science Foundation Early Career Award. He was the Chair of ASME Advanced Modeling & Simulation Technical Committee, and serves as the Vice Chair of the Executive Committee for ASME Computers & Information in Engineering Division.