Proceedings of the ASME 2020
International Design Engineering Technical Conferences
and Computers and Information in Engineering Conference
IDETC/CIE2020
August 16-19, 2020, St. Louis, MO, USA

IDETC2020-22661

DESIGN OF TRUSTWORTHY CYBER-PHYSICAL-SOCIAL SYSTEMS WITH DISCRETE BAYESIAN OPTIMIZATION

Yan Wang
Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA

ABSTRACT

Cyber-physical-social systems (CPSS) with highly integrated functions of sensing, actuation, computation, and communication are becoming the mainstream consumer and commercial products. The performance of CPSS heavily relies on the information sharing between devices. Given the extensive data collection and sharing, security and privacy are of major concerns. Thus one major challenge of designing those CPSS is how to incorporate the perception of trust in product and systems design. Recently a trust quantification method was proposed to measure trustworthiness of CPSS by quantitative metrics of ability, benevolence, and integrity. In this paper, the applications of ability and benevolence metrics in design optimization of CPSS architecture are demonstrated. A Bayesian optimization method is developed to perform trust based CPSS network design, where the most trustworthy network with respect to a reference node can be selected to collaborate and share information with.

Keywords: Cyber-Physical-Social Systems; Probabilistic Graph Model; Trust; Ability; Benevolence; Integrity; Bayesian Optimization

1. INTRODUCTION

Cyber-physical systems (CPS) are physical devices that have highly integrated functions of sensing, actuation, computation, and communication. Currently both consumer and commercial products are becoming more intelligent with the implementations of them as CPS. These CPS devices have sensors embedded and can collect data of the surrounding environment. The data are shared between those devices, which help human users as well as the devices themselves to make individual decisions in a highly distributed fashion. The decisions will be executed with the control unit. These devices are the essential elements for smart home, smart city, intelligent manufacturing, personalized medicine, autonomous and safe transportation, omnipresent energy supplies, and many other applications. Given the ubiquity of CPS and their interaction and seamless integration with human society, they are also termed as cyber-physical-social systems (CPSS).

The design of CPSS is challenging because various factors and constraints in the cyber, physical, and social dimensions of design space need to be considered. There are unique challenges in CPSS design, such as sustainability, reliability, resilience, interoperability, adaptability, bio-compatibility, flexibility, and safety in the physical subspace. There are also principles of human-in-the-loop, data-driven design, co-design, scalability, usability, and security that need to be considered in the cyber subspace. In social subspace, the perceptions of risk, trust, and privacy, as well as memory capacity and emotion of users need to be incorporated.

The rapid growth of CPSS requires engineers to adopt a new *design for connectivity* principle. Different from tradition products, CPSS devices heavily rely on information sharing with each other to be functioning. Those devices form the Internet of Things (IoT). How to consider the connectivity related issues in product design therefore is new to engineers. Particularly, each CPSS device constantly collects data and shares them with other devices in the networks. Information security and privacy become critical issues in designing such massively networked systems. At the high-level application layer, decisions of what data can be collected, where data are stored, who can access the data, which portion of data can be shared, etc. need to be made during the software design. These decisions will simultaneously affect hardware and mechanism design as well as product safety. The effectiveness of their performance critically depends on what and how they share among each other. Trust is an important design feature for these systems to work together. Therefore, designing the intelligent decision making and decision support subsystems for CPSS need to incorporate the trust aspect in the social dimension, as trustworthiness can affect the design of the policies for security and privacy.

Furthermore, trust is critical for human users of these CPSS devices whose personal information are likely to be collected and shared by the devices. The users' perceptions of trust about the systems can affect the effectiveness of human-device interactions. Thus designing trustworthy CPSS devices and systems is an important task for design engineers.

Trust has been extensively studied in the domains of psychology, organizational behavior, marketing, and computer science. However, most studies remain conceptual and qualitative. Quantitative measurements of trustworthiness are needed when the concept is applied in engineering design and optimization. Some quantitative studies of trust have been conducted in computer science, where trustworthiness is mostly quantified by quality of service (QoS), e.g. success rate as well as consistency in packet forwarding and other transactions, in network communication. The reputations in user ratings and recommendations online were also used. These metrics are quantities only in cyber design space. There is still lack of trustworthiness metrics in both cyber and social design spaces, which are important to guide the design of trustworthy CPSS at the levels of network architecture and devices.

In this work, the perception of trust is quantified and applied in CPSS architecture design, where the collaboration network of a particular node can be optimized based on trustworthiness criteria. The quantitative trustworthiness metrics are based on the recently proposed ability-benevolence-integrity (A-B-I) model [1]-[3], where trustworthiness is quantified by the cyber-social metrics of ability, benevolence, and integrity. Ability shows how well a trustee party is capable of doing what it claims to perform. Benevolence indicates whether the motivation of the trustee is purely for the benefit of itself. Integrity measures if the trustee does what it claims to. Based on a mesoscale probabilistic graph model [4][5] of CPSS, the perceptions of ability, benevolence, and integrity can be quantified with the probabilities of good judgements for the nodes as well as the information dependencies among nodes. In this paper, we further demonstrate how to apply the quantitative trustworthy metrics as the design criteria in network architecture design and optimization. The design criteria are used as the utilities to identify an optimal subset of nodes in the network that one particular node can trust and collaborate with.

Here, a discrete Bayesian optimization method is developed to solve the combinatorial optimization problem. Bayesian optimization is a robust global optimization scheme that incorporates uncertainty in the searching process. Different from other global optimization approaches such as the commonly used genetic algorithms, simulated annealing, and other heuristic algorithms, Bayesian optimization performs search based on a surrogate model of the objective function. The surrogate, usually a Gaussian process regression model, keeps the search history in memory as opposed to other "memoryless" heuristic algorithms. In addition, an acquisition function is constructed and used to guide the searching or sequential sampling process. It is designed to strike a good balance between exploration and exploitation. During sequential sampling, the surrogate of objective function is continuously updated with new samples based on the Bayesian belief update. Therefore the searching process in Bayesian optimization can be accelerated with the properly designed surrogate model and acquisition function. This provides unique advantages in discrete optimization over traditional heuristic algorithms, especially for complex combinatorial problems where exhaustive search in the discrete solution space is computationally prohibitive. In our discrete Bayesian optimization method for the combinatorial problem of network optimization, a new distance kernel is developed to measure the similarity between networks.

In the remainder of this paper, the existing work of system-level design of CPSS, discrete Bayesian optimization, and trust quantification approaches are reviewed in Section 2, where the probabilistic graph model of CPSS is also introduced. In Section 3, the metrics of ability and benevolence in the A-B-I trust model are introduced. The discrete Bayesian optimization method is described in Section 4. The application of Bayesian optimization to the CPSS network architecture design is demonstrated with ability and benevolence metrics as the utilities.

2. BACKGROUND

Here an overview of CPSS system-level design is given. The existing research on discrete Bayesian optimization and trust quantification are reviewed. The probabilistic graph model of CPSS which the A-B-I model is based upon is also introduced.

2.1 Systems level design of CPSS

Network connectivity is essential for CPSS. A standalone CPSS device cannot perform the functions which it is designed for. Compared to traditional products, the design of CPSS requires engineers to have better understanding of the systems level behaviors [7], from conceptual design to design optimization of multidisciplinary and hierarchical architecture [8]. Given the evolution nature of cyber and physical technologies, adaptability that enables the capabilities of self-learning, self-organization, and context awareness is important to design open systems that can evolve along technology advancement [6]. With the complexity of the CPSS networks grows to billions of nodes, it is impossible to ensure all nodes are free from compromise or breakdown. Node compromise and subnet disruption should be treated as daily norms. Therefore the emphasis of CPSS networks and systems should be more on the ability to recover from breakdown, instead of preventing its breakdown. That is, resilience (the ability to recover) is more important than reliability (the ability to stay functioning) in designing systems of CPSS [4][5].

Some systems modeling methods and tools have been applied for CPSS design and analysis, such as hybrid discrete-event and continuous simulations [12]-[14], inductive constraint logic programming [15], abductive reasoning [16], hybrid timed automaton [17], ontologies [18], information schema [19], UML [20], SysML [21], and information dynamics modeling [22]. The high-dimensional design space of CPSS includes not only the cyber and physical subspaces, but also the social subspace. The modalities for human-system interaction [10], context awareness and personalized human-system communication [11], as well as trusted collaboration [1]-[3] have been studied.

2.2 Bayesian optimization for discrete problems

Bayesian optimization has been widely used in the continuous domain and only recently gained more attention in the discrete domains, such as in solving mixed-integer problems [41,42]. The straightforward extension of Bayesian optimization from continuous domain to discrete domain is just treating discrete variables as continuous ones and round the variable values to the closest integers during the searching process. For instance, Baptista and Poloczek [44] proposed a quadratic acquisition function for combinatorial problems and converted the binary variables to high-dimensional vectors during the searching process. The solutions are then projected back to the binary space. However, this approach may fail to identify the true optimum and be trapped in the local region because there is a mismatch between the true discontinuous objective function and the assumed continuous acquisition function. Zaefferer et al. [43] replaced the continuous distance with discrete distance measures and compared the performance using the expected improvement acquisition function. Garrido-Merchán and Hernández-Lobato [45] developed an input variable transformation to ensure the distance between any two discrete variables remain unchanged in evaluating kernels when the variables perturb into the continuous space. Zhang et al. [46] proposed a new kernel function based on the position distance for permutation problems and the prior knowledge about similarity in the problems. The sparse Gaussian process model is used to reduce the computational cost of kernel update.

2.3 Trust quantification for CPS

Conceptually, trust is the willingness to be vulnerable to another. It is a different concept from security. Security is critical for trust. However, security along cannot guarantee the trustworthiness. For instance, although security protocols can ensure data are not intercepted during transmission, they provide no guarantee against the misuse by the receiving party or against fraud by the transmitting party. In recent studies in computer science, trust was quantified with reputation, ratings, and user recommendations in information systems and social networks [23,24]. It was also measured by QoS, routing and delivery success rates, and consistency of data forwarding in computer networks and sensor networks [25,26]. Probability [27-29], imprecise probability [30,31], and fuzzy logic [32-34] have been applied to quantify the human perception of trust.

To quantify trustworthiness of CPS, Chen et al. [35] developed a fuzzy model of trust based on the reputation of communication efficiency. Al-Hamadi and Chen [36] calculated trust from user ratings aggregated from different time periods and different locations. Xu et al. [37] used the weighted average of direct user experiences and other's recommendations to evaluate the trust of edge computing devices. Tao et al. [38] measured the sensor data trustworthiness with the consistency with reference data sets. Junejo et al. [39] quantified trustworthiness of CPS nodes by QoS measurements.

Different from the above, we developed a quantitative approach with multi-faceted metrics of ability, benevolence, and integrity [1]-[3], which has been qualitatively studied in social organization [40]. In the quantitative A-B-I model, ability characterizes a node's capabilities of sensing, reasoning, and influence to other nodes based on its probability of correct predictions as well as those of other nodes due to the information shared by this node. Benevolence characterizes the motivation of a node for its information sharing. Integrity is related to the traditional cyber and physical security and can be quantified from QoS.

In order to build large-scale networks, trustworthiness should be treated as transferrable quantities so that it can be propagated in scalable systems. With the quantitative measures of trustworthiness, the risk of deploying CPSS can be quantified and assessed more thoroughly in highly complex networks where a global view of the networks is impossible to obtain.

2.4 Probabilistic graph model of CPSS

The probabilistic graph model [2][5] is an abstraction of CPSS networks at the mesoscale. It captures the sensing, computing, and communication capabilities of CPSS by the prediction probabilities for all nodes in a CPSS network and the pair-wise reliance probabilities between nodes as the extent of information dependency and mutual influences. The model is illustrated in Figure 1. The prediction and reliance probabilities of nodes are defined as follows.

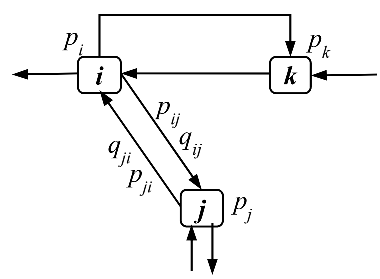


FIGURE 1: Probabilistic graph model of CPSS networks.

A probabilistic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{P}, \mathcal{R})$ consists of a set of vertices $\mathcal{V} = \{v_k\}$ and a set of directed edges $\mathcal{E} = \{(v_i, v_j)\}$. Each node v_k is associated with a prediction probability $p_k \in \mathcal{P}$, and each directed edge (v_i, v_j) is associated with a reliance probability $p_{ij} \in \mathcal{R}$. The prediction probability that the k-th node detects the true state of world θ is

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

where x_k is the state variable. Without loss of generality, here we only consider binary-valued state variables ($= \theta$ or $\neq \theta$). State variables with multiple discrete values can be easily extended. Continuous variables can be discretized in a digital computing environment.

With binary-valued state variables, we can define P-reliance probability

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

as the probability that the *j*-th node predicts the true state of world given that the *i*-th node predicts correctly. We also define Q-reliance probability

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

as the probability that the *j*-th node predicts the true state of world given that the *i*-th node does not predict the same.

The state variables contain the results from sensing. The values can be updated because of computing or reasoning. Therefore the prediction probabilities capture the sensing and computing functionalities, whereas the reliance probabilities indicate the functionality of communication. The random state variables with binary values can be extended to multiple values or continuous. For instance, one sensor measures a value which follows some distribution, as in prediction probability. If there are a finite set of possible values $\{\theta_1, \ldots, \theta_T\}$ for state variables. The prediction probability $P(x_k = \theta_n)$ and reliance probability $P\left(x_j = \theta_n \middle| x_i = \theta_m\right)$, where $1 \leq m, n \leq T$, can be enumerated similarly.

The edges in the probabilistic graph are directional. The neighbors of each node can be further differentiated as *source* nodes or *destination* nodes, as illustrated in Figure 2. For one node, its source nodes are those sending information to this node, whereas the destination nodes are those receiving information from it. When receiving different cues from source nodes, a CPS node can update its prediction probability to reflect its perception of the world. The aggregation of prediction probabilities sensitively depends on the rules of information fusion during the prediction update.

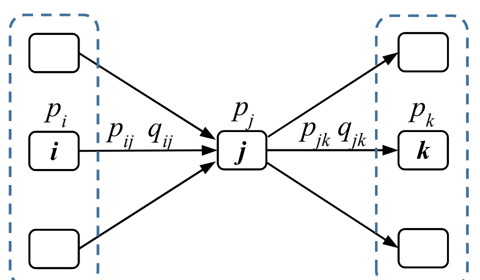


FIGURE 2: Source and destination nodes with respect to node *j* are differentiated.

If $P(x_k)$ and $P(x_k^C)$ denote the probabilities of a positive and a negative prediction from node *k* respectively, we define the *best-case* fusion rule as

$$P'^{(x_k)} = 1 - \left(1 - P(x_k)\right) \prod_{i=1}^{M_P} P(x_i)\left(1 - P(x_k|x_i)\right)$$
$$\prod_{j=1}^{M_N} P(x_j^C)\left(1 - P(x_k|x_j^C)\right) \tag{4}$$

where node k updates its prediction based on its own current prediction and those cues from its $M_P + M_N$ source nodes, out of which M_P of the source nodes provide positive predictions whereas M_N of them provide negative predictions, $P(x_k|x_i)$ indicates the probability that a positive message from node *i* leads to a positive prediction of node k, and $P(x_k|x_j^C)$ is the probability that a negative message from node j leads to a positive prediction of node *k*. Therefore, if any of the cues from the source nodes is positive, the prediction of the node is positive. Some variations of this fusion rules exist. For instance, the previous prediction from itself can be either included or excluded during the update.

Similarly, the *worst-case* fusion rule can be defined as

$$P'(x_k) = P(x_k) \prod_{i=1}^{M_P} P(x_i) P(x_k|x_i) \prod_{j=1}^{M_N} P(x_j^C) P(x_k|x_j^C) \tag{5}$$

That is, if any of the cues from the source nodes is negative, the prediction of the node is negative. The Bayesian fusion rule is defined as

$$P'(x_k) = \frac{P(x_k) \max_{\mathrm{P}}\left\{(P(x_k))^r \left(1 - P(x_k)\right)^{S-r}\right\}}{\int (P(x_k))^r \left(1 - P(x_k)\right)^{S-r} dP} \tag{6}$$

where the prediction of the node is updated to P' from prior prediction P, and out of S cues that the neighboring nodes provide, r of them provide are positive, if the maximum likelihood principle is taken.

The probabilistic graph model provides a system level abstraction and a mesoscale description of CPSS networks, where information exchange and aggregation are captured. Prediction and reliance probabilities can be easily obtained in a physical system from the collected historical data. The prediction probability of a 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. For instance, in sensor networks, the prediction probability associated with a node can be estimated as the ratio of the number of packets sent by this node to a baseline reference number that the best performer sends as the upper limit. The P-reliance probability for each path can be estimated as the ratio of the number of packets received by the destination to the number sent by the source [5]. If no experimental data are available, subjective estimations from domain experts can be elicited. Probability elicitation is well known in both practice and literature. Standard procedures are usually taken to elicit probabilities associated with some events from domain experts as subjective estimates.

3. THE A-B-I TRUST MODEL

Based on the probabilistic graph model, the trust metrics of ability and benevolence can be calculated. The ability of a CPSS node is measured with its capability of performing correct predictions and making right decisions from the perspectives of sensing and computation, as well as its influence to other nodes. The benevolence is measured by how willing it is to share information reciprocally and the motivation of sharing from the perspective of communication. The integrity of a CPSS node is closely related to the cybersecurity and can be evaluated with consistency, frequency of compromises, QoS, and other security measurements.

Here only the metrics of ability and benevolence are summarized. They will be used as the utilities to demonstrate the network optimization. Since integrity has been studied extensively in cybersecurity, ability and benevolence can show the uniqueness of our proposed trust measurements. The complete description of the A-B-I trust model as well as the illustrations of the metrics and their use for detecting malicious attacks can be found in Ref.[2].

3.1 Ability

The ability of a CPSS node is evaluated by its capability of prediction and its influence to other nodes. The capability of prediction for a node is measured by its capabilities of data collection and reasoning based on data obtained from its neighbors, quantified by the prediction probability and reliance probabilities perceived by others, as well as the precisions of the perceptions. The influence to others is quantified by how influential its information shared to others is in their decision making.

The perceived ability of node j with the consideration of its *prediction capability* is $A_j(\theta) = \mathbb{P}\left(P(x_j = \theta)\right)$, where $\mathbb{P}(\cdot)$ denotes perception. Suppose that all perceptions follow Gaussian distributions. The prediction capability can be quantified by its mean

$$\mathbb{E}\big(A_j(\theta)\big) = p_j, \tag{7}$$

and its variance

$$\mathbb{V}\big(A_j(\theta)\big) = \tau_j^{-1}. \tag{8}$$

That is, if a node has a higher prediction capability with less variability than others, it is more trustworthy.

Based on the directions of information sharing between nodes, the neighboring nodes for each node in the network are categorized as source nodes and destination nodes, as illustrated in Figure 2. With respect to node j, the set of source nodes that share information with node j is denoted as $\mathcal{S}_j = \{v_i | (v_i, v_j) \in \mathcal{E}\}$, and the set of destination nodes that receive information from node j is denoted as $\mathcal{D}_j = \{v_k | (v_j, v_k) \in \mathcal{E}\}$.

The perceptions about the P- and Q-reliance probabilities for nodes i and j are related to the information processing capability of node j. A high P-reliance probability indicates that node j can absorb knowledge quickly. A high Q-reliance probability shows that node j can have good judgement even in a noisy and uncertain situation. We simplify the notations as $L_{ij} = \mathbb{P}\left(P\left(x_j = \theta \middle| x_i = \theta\right)\right)$ and $L_{ij}^c = \mathbb{P}\left(P\left(x_j = \theta \middle| x_i \neq \theta\right)\right)$ respectively. They are assumed to follow Gaussian distributions with means $\mathbb{E}\big(L_{ij}|A_j\big) = p_{ij}$ and $\mathbb{E}\big(L_{ij}^c|A_j\big) = q_{ij}$, and variances $\mathbb{V}\big(L_{ij}|A_j\big) = \tau_{ij,p}^{-1}$ and $\mathbb{V}\big(L_{ij}^c|A_j\big) = \tau_{ij,q}^{-1}$, respectively.

The perceived ability of node j with the considerations of both capabilities of *prediction* and *information processing* is then quantified with mean

$$\mathbb{E}\big(A_j(\theta|\mathcal{L}^{(+j)})\big) = \frac{\tau_j p_j + \sum_{i\in\mathcal{S}_j} \tau_{ij,p} p_{ij} + \sum_{i\in\mathcal{S}_j} \tau_{ij,q} q_{ij}}{\tau_j + \sum_{i\in\mathcal{S}_j} \tau_{ij,p} + \sum_{i\in\mathcal{S}_j} \tau_{ij,q}} \tag{9}$$

and variance

$$\mathbb{V}\big(A_j(\theta|\mathcal{L}^{(+j)})\big) = \left(\tau_j + \textstyle\sum_{i\in\mathcal{S}_j} \tau_{ij,p} + \sum_{i\in\mathcal{S}_j} \tau_{ij,q}\right)^{-1} \tag{10}$$

based on Bayes' rule of belief update. Bayesian belief update is an intuitive way to combine multiple factors.

Leadership should be regarded as one's ability. Here, it is estimated as its influence to others by sharing information. The perceived ability of node j with the considerations of its *prediction capability* and *influence* is quantified with mean

$$\mathbb{E}\big(A_j(\theta|\mathcal{L}^{(-j)})\big) = \frac{\tau_j p_j + \sum_{k\in\mathcal{D}_j} \tau_{jk,p} p_{jk} + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}(1-q_{jk})}{\tau_j + \sum_{k\in\mathcal{D}_j} \tau_{jk,p} + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}} \tag{11}$$

and variance

$$\mathbb{V}\left(A_j\left(\theta|\mathcal{L}^{(-j)}\right)\right) = \left(\tau_j + \textstyle\sum_{k\in\mathcal{D}_j} \tau_{jk,p} + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}\right)^{-1} \tag{12}$$

The overall and comprehensive ability perception with the simultaneous considerations of its capabilities of prediction, information processing, and influence is similarly calculated as

$$\mathbb{E}\left(A_j\left(\theta\middle|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)}\right)\right) = \frac{\tau_j p_j + \sum_{i\in\mathcal{S}_j} \tau_{ij,p} p_{ij} + \sum_{i\in\mathcal{S}_j} \tau_{ij,q} q_{ij} + \sum_{k\in\mathcal{D}_j} \tau_{jk,p} p_{jk} + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}(1-q_{jk})}{\tau_j + \sum_{i\in\mathcal{S}_j} \tau_{ij,p} + \sum_{i\in\mathcal{S}_j} \tau_{ij,q} + \sum_{k\in\mathcal{D}_j} \tau_{jk,p} + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}} \tag{13}$$

$$\mathbb{V}\left(A_j\left(\theta\middle|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)}\right)\right) = \left(\tau_j + \textstyle\sum_{i\in\mathcal{S}_j} \tau_{ij,p} + \sum_{i\in\mathcal{S}_j} \tau_{ij,q} + \sum_{k\in\mathcal{D}_j} \tau_{jk,p} + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}\right)^{-1} \tag{14}$$

Therefore, a node that gives accurate predictions, makes sound decisions, and brings positive influences to others is deemed to be trustworthy.

The perception of one's ability can also be dictated by the abilities of those ones that are closely associated. That is, if a neighbor or associate, who is influenced by a node, has high ability, the perception of this node's ability is also increased. Therefore higher-order perception of ability can be defined. If the ability in Eqs. (13) and (14) is first-order and has values of mean $\mathbb{E}\left(A_j\left(\theta\middle|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)}\right)\right) = E_j$ and variance $\mathbb{V}\left(A_j\left(\theta\middle|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)}\right)\right) = V_j$, the second-order ability is defined as

$$\mathbb{E}^{(2)}\left(A_j\left(\theta\middle|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)}\right)\right) = \frac{V_j^{-1} E_j + \sum_{k\in\mathcal{D}_j} \tau_{jk,p} p_{jk} (V_k^{-1} E_k) + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}(1-q_{jk})(V_k^{-1} E_k)}{\tau_j + \sum_{k\in\mathcal{D}_j} \tau_{jk,p} p_{jk} V_k^{-1} + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}(1-q_{jk}) V_k^{-1}} \tag{15}$$

$$\mathbb{V}^{(2)}\left(A_j\left(\theta\middle|\mathcal{L}^{(+j)}, \mathcal{L}^{(-j)}\right)\right) = \left(\tau_j + \textstyle\sum_{k\in\mathcal{D}_j} \tau_{jk,p} p_{jk} V_k^{-1} + \sum_{k\in\mathcal{D}_j} \tau_{jk,q}(1-q_{jk}) V_k^{-1}\right)^{-1} \tag{16}$$

Higher-order perceptions of ability can be similarly defined.

3.2 Benevolence

The benevolence of a CPSS node is evaluated by the reciprocity and motive. The perception of reciprocity is measured by the willingness of sharing information to others while receiving information simultaneously. The motive is quantified by the quality of information shared to others and the frequency of sharing.

The expected reciprocity for node j perceived by node i is defined as

$$\mathbb{E}\big(R_{i,j}\big) = D_{\mathrm{KL}}\big(p_{i\to j} || p_{j\to i}\big) - D_{\mathrm{KL}}\big(p_{j\to i} || p_{i\to j}\big) + b_0 \tag{17}$$

where $p_{j\to i} = \prod_{k=j}^{i-1} p_{k,k+1}$ is the product of all P-reliance probabilities $p_{k,k+1}$ corresponding to the shortest path from node *j* to node *i*, $D_{\mathrm{KL}}(P||Q) = \sum_i P_i \log(P_i/Q_i)$ is the Kullback-Leibler divergence from probability Q to P, and b_0 is a reference value such that $\mathbb{E}(R_{i,j}) > b_0$ when node *j* has a larger reciprocity with respect to node *i*. Intuitively, if node *j* is willing to share accurate information with node *i* without necessarily expecting node *i* to share information as a return, node *j* has a high reciprocity to node *i*. In other words, node *i* can trust node *j*. Here, $b_0 = 0.5$ such that reciprocity has a value between 0 and 1. A higher value of reciprocity indicates higher trustworthiness. Furthermore, $\mathbb{E}(R_{i,i}) = b_0$. The variance associated with the perceived reciprocity is conservatively estimated as

$$\mathbb{V}(R_{i,j}) = \min\left(\sum_{j\to i} \tau_{ab}^{-1} + \sum_{i\to j} \tau_{cd}^{-1}, V_{max}\right) \quad (18)$$

where τ_{ab} and τ_{cd} are the precisions associated with the P-reliance probabilities along paths *j*→*i* and *i*→*j*, respectively, and $V_{max} = 1.0$ is the theoretical maximum value of variance associated with probabilities. $\mathbb{V}(R_{i,i}) = 0$.

Motive measures the intention of information sharing within a community. Sharing high-quality information with neighbors indicates the good purpose of improving the overall functionality of the community. Thus perceived motive of node *j* is defined as

$$\mathbb{E}(M_j) = p_j^{d_j} \quad (19)$$

$$\mathbb{V}(M_j) = \tau_j^{-1} \quad (20)$$

where p_j is the prediction probability associated with node *j* with precision τ_j, and $d_j = |\mathcal{D}_j|$ is the number of destination nodes for node *j*.

The overall benevolence of node *j* perceived by node *i* is

$$\mathbb{E}(B_{i,j}) = \frac{\mathbb{V}^{-1}(R_{i,j})\mathbb{E}(R_{i,j}) + \mathbb{V}^{-1}(M_j)\mathbb{E}(M_j)}{\mathbb{V}^{-1}(R_{i,j}) + \mathbb{V}^{-1}(M_j)} \quad (21)$$

$$\mathbb{V}(B_{i,j}) = \left(\mathbb{V}^{-1}(R_{i,j}) + \mathbb{V}^{-1}(M_j)\right)^{-1} \quad (22)$$

4. DISCRETE BAYESIAN OPTIMIZATION

The trust-based network optimization is to identify a subset of nodes in the network which are the most trustworthy with respect to a reference node. The optimization problem involves choosing the best subset of nodes and therefore is combinatorically complex. The traditional approach to solve these problems is using heuristic algorithms such as genetic algorithms and simulated annealing.

Here, a new discrete Bayesian optimization (dBO) method is developed to perform the CPSS network optimization. The design problem is to choose the optimum subgraph out of a graph with respect to a reference node such that the trustworthiness level perceived by the reference node is maximized. The proposed dBO method is a global optimization method to find the optimum combination of nodes.

Bayesian optimization is a class of surrogate based methods to search global optimum under uncertainty with Bayesian sequential sampling strategies. The search or sampling process is based on an acquisition function that is defined in the same input space of the objective function. In parallel, a surrogate model of the objective is also constructed and updated during the search. The most used surrogate is Gaussian process regression (GPR) model which is updated based on the Bayesian principle.

The sampling strategy of choosing the next sample is to maximize the acquisition function instead of the objective surrogate. One example of acquisition functions is the expected improvement (EI)

$$a_{EI}(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta) = \sigma(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta)\left(\gamma(\boldsymbol{x})\Phi(\gamma(\boldsymbol{x})) + \phi(\gamma(\boldsymbol{x}))\right) \quad (23)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function and cumulative distribution function of the standard normal distribution, $\gamma(\boldsymbol{x}) = (\mu(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta) - y_{best})/\sigma(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta)$ is the deviation away from the best solution y_{best} found so far, with posterior mean $\mu(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta)$ and posterior standard deviation $\sigma(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta)$, given the existing *D* samples $\{\boldsymbol{x}_i, y_i\}_{i=1}^D$ and GPR hyper-parameter θ.

Another example of acquisition function is upper confidence bound (UCB)

$$a_{UCB}(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta) = \mu(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta) + \kappa\sigma(\boldsymbol{x}; \{\boldsymbol{x}_i, y_i\}_{i=1}^D, \theta) \quad (24)$$

where κ is a hyper-parameter for the exploitation-exploration balance. To simply the optimization process, in this work we choose $\kappa = 1.5$ as a constant instead.

In the proposed dBO method for network design, the GPR surrogate of the objective function $f(\boldsymbol{z}) \sim \mathcal{GP}(m(\boldsymbol{z}), k(\boldsymbol{z}, \boldsymbol{z}'))$ has mean function $m(\boldsymbol{z})$ and covariance kernel function $k(\boldsymbol{z}, \boldsymbol{z}')$, where $\boldsymbol{z} = [z_1, \dots, z_N]$ is an index vector of *N* binary values ($z_i \in \{0,1\}, \forall i = 1, \dots, N$) for a graph with *N* nodes. A “1” indicates that the corresponding node is included in the subgraph as the solution, and a “0” indicates not. The major construct of the GPR model is the kernel function, defined as

$$k(\boldsymbol{z}, \boldsymbol{z}') = \exp\left(\sum_{i=1}^N d(z_i, z_i')/\theta_i\right), \quad (25)$$

where $d(\cdot)$ is a distance function defined in the discrete space such as the Hamming distance, and θ_i's are the hyper-parameters of scales. The advantage of one independent scale parameter being associated with each node comparison is that the different importance levels of nodes for trust quantification can be captured. In other words, not every node in a network is equally trustworthy with respect to a reference node. The scale parameters after the training can provide the weights of importance. The disadvantage of the kernel function in Eq. (25) is that the quickly increased number of hyper-parameters for large networks requires a large training datasets. The prediction will not be accurate otherwise. One easy way to mitigate the risk and reduce the computational load is to assume that all hyper-parameters have the same value, as

$$k(\boldsymbol{z}, \boldsymbol{z}') = \exp\left(\sum_{i=1}^N d(z_i, z_i')/\theta\right). \quad (26)$$

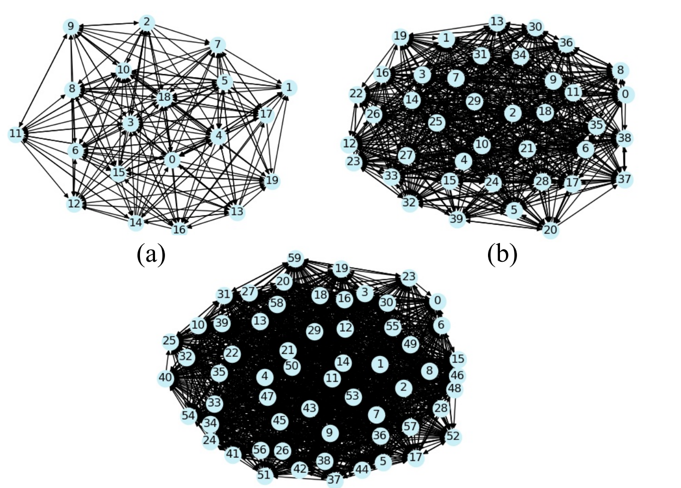
That is, there is only one hyper-parameter θ. This greatly simplifies the training process, at the expense of losing model granularity.

5. TRUST BASED STRATEGIC NETWORK DESIGN

A strategic network for a node is the most trustworthy network that the node can form the strategic collaboration relation. The design of such strategic network is to identify a subset of nodes within the complete network so that the node has the highest trustworthiness level. The trustworthiness metrics of ability and benevolence are used here to demonstrate the trust based strategic network design. The network optimization based on other metrics such as integrity can be done similarly.

5.1 Ability as the optimization criteria

Ability in Eq. (13) is first utilized as the metric to identify the most trustworthy network for a reference node. The strategic network of the reference node can be obtained by finding the network where the ability of the reference node is maximized. Three networks with 20, 40, and 60 nodes, shown in Figure 3, are generated with random connections for tests. The prediction and reliance probabilities are also randomly generated. Note that the random networks are generated to better test the robustness and scalability of the design optimization method than some deterministic ones.



(a) (b) (c)

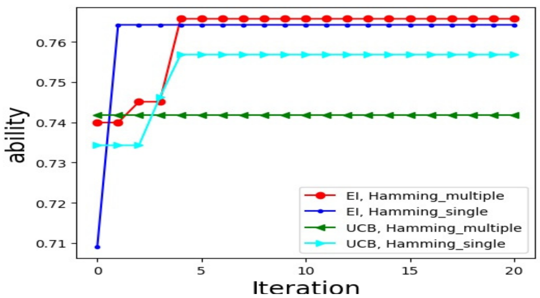
FIGURE 3: Three example networks for optimization tests, with (a) 20 nodes and 192 edges, (b) 40 nodes and 787 edges, and (c) 60 nodes and 1731 edges.

The EI acquisition in Eq. (23) and UCB acquisition in Eq. (24) along with the two kernel functions in Eqs. (25) and (26) are tested for the 20-node-192-edge example. The Hamming distance is used in the kernels. When searching for the optimum network to maximize the ability of node 0, they have different convergence rates, as compared in Figure 4(a). The optimum solution, as shown in Figure 4(b), is found with the EI acquisition in combination with the multi-parameter kernel. During the search, a simulated annealing algorithm is applied to maximize the acquisition to decide the next sample. It is seen that the search can be trapped at the local optimum when the single-parameter kernel function in Eq. (26) is used. The single-parameter kernel function does not provide the as much granularity as the multi-parameter kernel and does not differentiate much about the different contributions between nodes for the ability of node 0. Therefore, the parameter training tends to be not optimal. The UCB acquisition function emphasizes more on exploitation than the EI acquisition. Thus the search tends to get trapped in local optima.

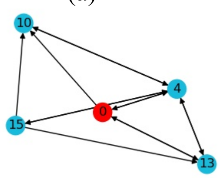
The convergence speeds for the networks of different sizes are further tested. The results are shown in Figure 5. It is seen that as the size of network increases, more iterations are required to find the global optimum. The reason is two-fold. First, larger networks result in the higher dimension of the searching space. The searching complexity for the possible solutions grows exponentially. Second, as the dimension of searching space increases, more samples are required to construct reliable surrogate models. Therefore, more iterations are necessary to ensure the convergence to the global optimum.

To compare the performance of the dBO method with the commonly used heuristic algorithms, simulated annealing is applied for the same network optimization problems. For each of the three examples with 20, 40, and 60 nodes, the simulated annealing algorithm to maximize the ability metric is run 5 times with different annealing steps ranging from 50 to 300. The means and standard deviations of the obtained optimal ability values for those test runs are listed in Table 1, Table 2, and Table 3 respectively. The means and standard deviations of results for 5 runs of the dBO algorithm after 50 iterations are also listed in these tables, where EI acquisition and multi-parameter kernel are used. The number of annealing steps indicates the computational cost where each step involves one evaluation of the original objective function. In the dBO searching, 50 initial samples with the evaluations of the objective function were obtained to construct the initial GPR surrogate. Additional samples are added for each of the iterations in Figure 4 and Figure 5. Each iteration involves one evaluation of the objective function, whereas the evaluation of the acquisition function in Bayesian optimization is based on the surrogate and usually costs much less, especially when the original objective function requires heavy computation. Therefore, the cost of dBO for 50 iterations is approximately equivalent to the cost of simulated annealing for 100 steps in these examples. From the comparisons, it is seen that the dBO method can find better solutions than the simulated annealing with the similar cost. Furthermore, the results of the dBO method have much less variability. In other words, the dBO algorithm is also more robust than the heuristic simulated annealing.

Besides the comprehensive ability metric, capacity in Eq. (9) and influence in Eq. (11) can also be applied individually as the criteria to perform design optimization based on specific interests. In addition, the second-order ability in Eq. (15) can also be used as the optimization criterion. The respective optimum networks based on these three criteria for node 0 in the 20-node example are shown in Figure 6. It is seen that different criteria lead to different optimum networks. If multiple criteria are used simultaneously, multi-objective optimization methods are needed.

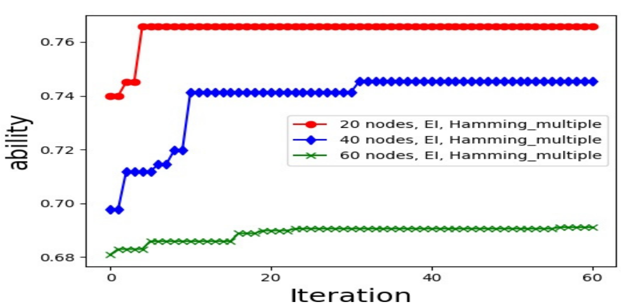


(a)

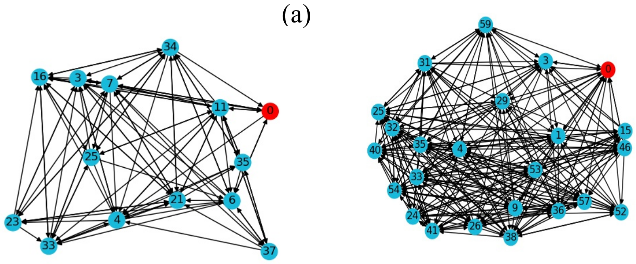


(b)

FIGURE 4: (a) The convergence speed of four cases with EI and UCB acquisition functions, along with single-parameter and multiple-parameter kernel functions, are com-pared for the 20-node-192-edge example. (b) The optimum network with the ability of node 0 maximized is found with the EI acquisition and multiple-parameter kernel.



(a)



(b) (c)

FIGURE 5: (a) The convergence speeds when searching in the 20-, 40-, and 60-node networks, with the EI acquisition and multi-parameter kernel functions. (b) The optimum in the 40-node network. (c) The optimum in the 60-node network.

TABLE 1: The means and standard deviations of the maximum ability for the 20-node network using simulated annealing with different annealing steps, in comparison with the dBO of 50 iterations

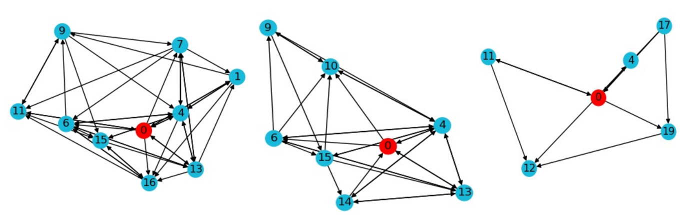
Steps	Mean	Standard Deviation
50	0.704128758	0.024803099
100	**0.717732062**	**0.01618725**
150	0.724677974	0.021446642
200	0.738149753	0.026914332
250	0.72842703	0.018894042
300	0.726842286	0.014625707
dBO	**0.763904996**	**0.002614458**

TABLE 2: The means and standard deviations of the maximum ability for the 40-node network using simulated annealing with different annealing steps, in comparison with the dBO of 50 iterations

Steps	Mean	Standard Deviation
50	0.638595221	0.060644109
100	**0.684115767**	**0.035342407**
150	0.696934409	0.028088683
200	0.68054112	0.023215712
250	0.709194429	0.031983543
300	0.70440341	0.023225232
dBO	**0.746661792**	**0.00340882**

TABLE 3: The means and standard deviations of the maximum ability for the 60-node network using simulated annealing with different annealing steps, in comparison with the dBO of 50 iterations

Steps	Mean	Standard Deviation
50	0.623391013	0.056150683
100	**0.65012841**	**0.039877341**
150	0.657217419	0.046396371
200	0.679789337	0.005860135
250	0.678678903	0.005974927
300	0.676195812	0.00793658
dBO	**0.692554458**	**0.003021649**



(a) (b) (c)

FIGURE 6: Optimum networks with respect to node 0 in the 20-node-192-edge example by different ability metrics: (a) capability as criterion, (b) influence as criterion, and (c) second-order ability as criterion.

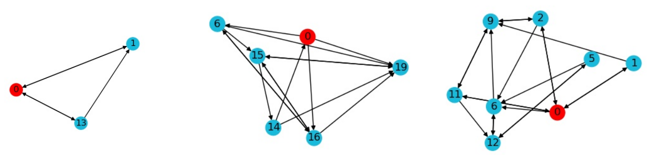
5.2 Benevolence as the optimization criteria

The design optimization procedure can be similarly applied with benevolence as the criterion. Because the reciprocity in Eq. (17) and benevolence in Eq. (21) are defined as pair-wise metrics, the optimization can be based on the weighted average benevolence perceived by node i as

$$U^{(i)} = \sum_{j\in\mathcal{V}^{(i)}} w_j \bar{B}_j \quad (27)$$

for all neighboring nodes $\mathcal{V}^{(i)}$ of node i, where $\bar{B}_j = (1/n_j)\sum_{k\in\mathcal{V}^{(i)}} B_{j,k}$ is the average benevolence of node j among its n_j neighbors, and weights w_j's ($0 \le w_j \le 1$) indicate the self-interest level. When $w_i = 1$ and $w_j = 0$ ($\forall j \neq i$) with respect to node i, it is a "selfish" mode. Only the benevolence of node i is considered as the criterion to find the optimum network for node i. On the other hand, when $w_i = 0$ and $\sum_{j\neq i} w_j = 1$, it is considered to be a "altruistic" mode. The weighted average reciprocity can be calculated similarly.

In the 20-node-192-edge example, the optimum networks for node 0 with the benevolence criteria are shown in Figure 7. It is seen when the self-interest weight w_0 is lower it is easier to build a larger trustworthy network.



(a) (b) (c)

FIGURE 7: Optimum networks with respect to node 0 in the 20-node-192-edge example by different benevolence metrics: (a) weighted average benevolence as criterion with w_0=1; (b) weighted average benevolence as criterion with w_0=1/2 and all other weights are 1/38; (c) weighted average reciprocity as criterion with w_0=1/2 and all other weights are 1/38.

6. CONCLUSION

In this paper, quantitative trustworthiness metrics are used as the design criteria to perform optimization of cyber-physical-social system networks. Each node can choose its own most trusted strategic network so that they can collaborate and share information. The trustworthiness is quantified as multi-facet quantities in both cyber and social spaces, including the dimensions of ability, benevolence, and integrity. In CPSS, the ability and benevolence can be calculated based on statistics from their working history to measure the capacities of information gathering, reasoning, and information sharing. The most trusted strategic network for a node is the subnet that maximizes the ability of the node if ability is used as the criterion. A node that has the high capacities of observing the state of world accurately, making sound decisions based on available information, and bringing positive impacts to others is deemed to possess a high level of ability and thus a trustworthy individual. Similarly, a node that is willing to share accurate information with others is also regarded as trustworthy. The strategic network is the one that leads to the maximum level of ability for the reference node, or consists of a group of collaborators that are the most willing to collaborate with the reference node.

It has been shown [2] that the new ability, benevolence, and integrity metrics are sensitive to trust attacks. When a malicious node generates false predictions and sends them to other nodes, its perceived trustworthiness will drop quickly. When the attack stops, the perceived trustworthiness will gradually increase and recover. This matches well with human social behaviors. It usually takes time to establish a trust relation, whereas the damage can be done much more quickly. When designing the trusted strategic network, the risks of attacks also need to be considered. Instead of targeting at the maximum trust level as shown in this paper, additional criteria for robustness need to be incorporated in future work.

The proposed discrete Bayesian optimization performs reasonably well for the combinatorial problem of network design. For the kernel function based on the Hamming distance, more hyper-parameters can help increase the flexibility of the kernel, whereas a small number of hyper-parameters is not robust enough for optimization. The limitation of using multiple hyper-parameters is the training efficiency. More samples are required to train a larger number of hyper-parameters, which makes it not feasible for small problems. Combinatorial problems usually have very large searching space. Introducing additional hyper-parameters can potentially bring the benefit of faster convergence.

In this work only single-objective optimization is applied. The multi-facet trustworthiness metrics eventually will need a multi-objective optimization approach for trust based design, where multiple metrics are considered simultaneously and tradeoffs need to be made. The scalability of the discrete Bayesian optimization also requires further investigation, given that the Bayesian update procedure in GPR is computationally expensive when the number of samples is large. The proposed scheme for large-scale networks will require further tests. Enhancement such as sparse GPR is likely to bring better scalability.

ACKNOWLEDGEMENTS

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

REFERENCES

1. Wang, Y. (2018). Trust Based Cyber-Physical Systems Network Design. In *Proc. ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2018)*, pp. V01AT02A037.
2. Wang, Y. (2018). Trust quantification for networked cyber-physical systems. *IEEE Internet of Things Journal*, 5(3): 2055-2070.
3. Wang, Y. (2018). Trustworthiness in designing cyber-physical systems. In *Proc. 12th International Symposium on Tools and Methods of Competitive Engineering (TMCE2018)*, pp.27–40.

4. Wang, Y. (2016) System resilience quantification for probabilistic design of Internet-of-Things architecture. In *Proc. 2016 ASME International Design Engineering Technical Conferences & The Computer and Information in Engineering Conference (IDETC/CIE2016)*, pp. V01BT02A011.
5. Wang, Y. (2018). Resilience Quantification for Probabilistic Design of Cyber-Physical System Networks. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*. 4(3), 031006.
6. Horváth, I., & Gerritsen, B. H. (2012). Cyber-physical systems: Concepts, technologies and implementation principles. In *Proc. 9th International Symposium on Tools and Methods of Competitive Engineering (TMCE2012)*, pp. 19–36.
7. Tavčar, J., & Horváth, I. (2018). A review of the principles of designing smart cyber-physical systems for run-time adaptation: Learned lessons and open issues. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(1), 145-158.
8. Grimm, M., Anderl, R., & Wang, Y. (2014). Cyber-physical augmentation: An exploration. In *Proc. 10th International Symposium on Tools and Methods of Competitive Engineering (TMCE2014)*, pp. 61-72.
9. Vroom, R. W., & Horváth, I. (2014). Cyber-physical augmentation: An exploration. In *Proc. 10th International Symposium on Tools and Methods of Competitive Engineering (TMCE2014)*, pp. 1509-1520.
10. Horváth, I., & Wang, J. (2015). Towards a comprehensive theory of multi-aspect interaction with cyber physical systems. In *Proc. ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2015)*, pp. V01BT02A009.
11. Li, Y., Horváth, I., & Rusák, Z. (2018). Constructing Personalized Messages for Informing Cyber-Physical Systems based on Dynamic Context Information Processing. In *Proc. 12th International Symposium on Tools and Methods of Competitive Engineering (TMCE2018)*, pp.105–120.
12. Jeon, J., Chun, I., & Kim, W. (2012). Metamodel-based CPS modeling tool. In *Embedded and Multimedia Computing Technology and Service*, *Lecture Notes in Electrical Engineering*, Vol. 181 (pp. 285-291). Springer.
13. Lee, K. H., Hong, J. H., & Kim, T. G. (2015). System of Systems Approach to Formal Modeling of CPS for Simulation-Based Analysis. *ETRI Journal*, 37(1), 175-185.
14. Lee, E. A., Niknami, M., Nouidui, T. S., & Wetter, M. (2015, October). Modeling and simulating cyber-physical systems using CyPhySim. In *Proc. the 12th IEEE International Conference on Embedded Software*, pp. 115-124.
15. Saeedloei, N., & Gupta, G. (2011). A logic-based modeling and verification of CPS. *ACM SIGBED Review*, 8(2), 31-34.
16. Horváth, I. (2019). A Computational Framework for Procedural Abduction Done by Smart Cyber-Physical Systems. *Designs*, 3(1), 1.
17. Burmester, M., Magkos, E., & Chrissikopoulos, V. (2012). Modeling security in cyber–physical systems. *International Journal of Critical Infrastructure Protection*, 5(3-4), 118-126.
18. Petnga, L., & Austin, M. (2016). An ontological framework for knowledge modeling and decision support in cyber-physical systems. *Advanced Engineering Informatics*, 30(1), 77-94
19. Pourtalebi, S., & Horváth, I. (2017). Information schema constructs for instantiation and composition of system manifestation features. *Frontiers of Information Technology & Electronic Engineering*, 18(9), 1396-1415.
20. Magureanu, G., Gavrilescu, M., Pescaru, D., & Doboli, A. (2010, September). Towards UML modeling of cyber-physical systems: A case study for gas distribution. In *Proc. IEEE 8th International Symposium on Intelligent Systems and Informatics*, pp. 471-476.
21. Palachi, E., Cohen, C., & Takashi, S. (2013, April). Simulation of cyber physical models using SysML and numerical solvers. In *Proc. 2013 IEEE International Systems Conference (SysCon)*, pp. 671-675.
22. Wang, Y. (2020). Information dynamics in the network of cyber-physical systems. In *Proc. 13th International Symposium on Tools and Methods of Competitive Engineering (TMCE2020)*.
23. Vu, L. H., & Aberer, K. (2011). Effective usage of computational trust models in rational environments. *ACM Transactions on Autonomous and Adaptive Systems*, 6(4): 24.
24. Ruan, Y., & Durresi, A. (2016). A survey of trust management systems for online social communities–trust modeling, trust inference and attacks. *Knowledge-Based Systems*, 106: 150-163.
25. Li, X., Zhou, F., & Du, J. (2013). LDTS: A lightweight and dependable trust system for clustered wireless sensor networks. *IEEE Transactions on Information Forensics and Security*, 8(6): 924-935.
26. Chen, Z., Tian, L., & Lin, C. (2017). Trust model of wireless sensor networks and its application in data fusion. *Sensors*, 17(4): 703.
27. Barber, K. S., & Kim, J. (2001). Belief revision process based on trust: Agents evaluating reputation of information sources. In: *Trust in Cyber-societies* (pp. 73-82). Springer.
28. Kim, H., Lee, H., Kim, W., & Kim, Y. (2010). A trust evaluation model for QoS guarantee in cloud systems. *International Journal of Grid and Distributed Computing*, 3(1): 1-10.
29. Li, X., Ma, H., Zhou, F., & Gui, X. (2014). Service operator-aware trust scheme for resource matchmaking across multiple clouds. *IEEE Transactions on Parallel and Distributed Systems*, 26(5): 1419-1429.

30. Yu, B., & Singh, M. P. (2002). Distributed reputation management for electronic commerce. *Computational Intelligence*, 18(4): 535-549.
31. Reddy, V. B., Venkataraman, S., & Negi, A. (2017). Communication and data trust for wireless sensor networks using D–S theory. *IEEE Sensors Journal*, 17(12): 3921-3929.
32. Falcone, R., Pezzulo, G., & Castelfranchi, C. (2002). A fuzzy approach to a belief-based trust computation. In: *Proc. Workshop on Deception, Fraud and Trust in Agent Societies* (pp. 73-86). Springer.
33. Alhamad, M., Dillon, T., & Chang, E. (2011). A trust-evaluation metric for cloud applications. *International Journal of Machine Learning and Computing*, 1(4): 416.
34. Ashtiani, M., & Azgomi, M. A. (2016). Trust modeling based on a combination of fuzzy analytic hierarchy process and fuzzy VIKOR. *Soft Computing*, 20(1): 399-421.
35. Chen, D., Chang, G., Sun, D., Li, J., Jia, J., & Wang, X. (2011). TRM-IoT: A trust management model based on fuzzy reputation for internet of things. *Computer Science and Information Systems*, 8(4), 1207-1228.
36. Al-Hamadi, H., & Chen, R. (2017). Trust-based decision making for health IoT systems. *IEEE Internet of Things Journal*, 4(5), 1408-1419.
37. Xu, Q., Su, Z., Wang, Y., & Dai, M. (2018). A trustworthy content caching and bandwidth allocation scheme with edge computing for smart campus. *IEEE Access*, 6, 63868-63879.
38. Tao, H., Bhuiyan, M. Z. A., Rahman, M. A., Wang, T., Wu, J., Salih, S. Q., Li, Y., & Hayajneh, T. (2020). TrustData: Trustworthy and Secured Data Collection for Event Detection in Industrial Cyber-Physical System. *IEEE Transactions on Industrial Informatics*, 16(5), 3311-3321.
39. Junejo, A. K., Komninos, N., Sathiyanarayanan, M., & Chowdhry, B. S. (2020). Trustee: A Trust Management System for Fog-enabled Cyber Physical Systems. *IEEE Transactions on Emerging Topics in Computing*.
40. Mayer, R. C., Davis, J. H., and Schoorman, F. D. (1995). An integrative model of organizational trust. *Acad. of Management Review*, 20(3): 709–734.
41. Tran A.V., Tran M., and Wang Y. (2019) Constrained mixed-integer Gaussian mixture Bayesian optimization and its applications in designing fractal and auxetic metamaterials. *Structural & Multidisciplinary Optimization*, 59(6): 2131–2154
42. Iyer, A., Zhang, Y., Prasad, A., Tao, S., Wang, Y., Schadler, L., Brinson, L.C. & Chen, W. (2019). Data-centric mixed-variable Bayesian optimization for materials design. In: *Proc. ASME 2019 IDETC/CIE Conferences*, pp.V02AT03A066.
43. Zaefferer, M., Stork, J., Friese, M., Fischbach, A., Naujoks, B., & Bartz-Beielstein, T. (2014). Efficient global optimization for combinatorial problems. In: *Proc. 2014 ACM Annual Conference on Genetic and Evolutionary Computation*, pp. 871-878.
44. Baptista, R., & Poloczek, M. (2018) Bayesian Optimization of Combinatorial Structures. In: *Proc. PMLR 35th International Conference on Machine Learning, PMLR*, 80, pp.462-471.
45. Garrido-Merchán, E. C., & Hernández-Lobato, D. (2020). Dealing with categorical and integer-valued variables in Bayesian optimization with Gaussian processes. *Neurocomputing*, 380: 20-35.
46. Zhang, J., Yao, X., Liu, M., & Wang, Y. (2019) A Bayesian discrete optimization algorithm for permutation problems. In: *Proc. 2019 IEEE Symposium Series on Computational Intelligence (SSCI 2019), pp.871-881*.