Cognition on the Networked Data of Stochastic Topology

Tzu-Yu Chuang*, and Kwang-Cheng Chen†, *Fellow, IEEE*
Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan
Email : d99942022@ntu.edu.tw* and chenkc@cc.ee.ntu.edu.tw†

***Abstract*—In this paper we address the problem of cognition (inference and decision) on data collected from a system which is characterized by a stochastic network topology. The decision problem is formulated as a data fusion problem and a data model that only requires stochastic network topology, marginal probability densities and pairwise correlations is proposed. Based on such model, we can mitigate the difficulties of inference resulted from random topological networks and simplify the decision making in many data cases. The numerical simulation justifies our idea and the performance is comparable to the case of decision making with complete information and decision making under conventional independent data cases when the size of network is not too large. Furthermore, an experiment of classification of real-world data is also conducted to illustrate the potential applications of cognition of networked data on social networks.**

***Index Terms*—Cognition, stochastic network topology, statistical decision theory, minimax decision.**

I. Introduction

NETWORKED data processing is a new and challenging topic in the signal processing area [1], and is very useful in real world. Many applications need to make decisions based on data collected from a system with thousands or millions of components (agents), all capable of generating and communicating data, like computers, cell phones, sensors or people. For example, prediction of users' behavior in the online social networks is an immediate problem of this type [2]. In addition, the defense of intentional attacks in smart grid communication networks [3] or in complex networks [4] can be resolved by signal detection. Specific event detection in home-caring systems, such as fire alarm or elders accidents, is also another application of coordinating large number of sensors' information to obtain optimal decision. The decision making with networked data is challenging because it involves not only the information from data itself but also the information conveyed by the network structure, which tend to be complex as the rapid growth of network size.

One of the most useful theories behind decision making is *statistical decision theory* [5], which embeds the uncertainty of real world into a probability space and then we can utilize data to make a decision of the state of real world to optimize our payoff/risk statistically. From this point of view, it is expected that the more understanding of the statistical property of the system we have, the better decision we can make about the nature. The understanding of statistical property is also known as *statistical inference*. We refer the joint process of inference and decision making together as *Cognition*, which is of particular importance on networked data because it involves the inference of network structure, which might be exhausted if we do not have proper philosophy on it.

Cognition on the networked data has many different aspects from conventional data processing philosophy. In conventional setting, every component is treated or designed independently, and thus we can preserve very good statistical properties of data (independent and identical distributed); however, the system we are interested in now is composed of a large number of small interacting units which interact with each other and then generate correlated data. Even though we can understand the correlation pattern of data by the *correlation coefficient* or by the *joint probability density function* and make decisions based on them, to obtain these information indicates more inference and learning works on the data, which are still much more complicated than the conventional cases. Another aspect of difference is the discovery of network structure. It is often preferred to model the network topology of a large and complex networked system stochastically, such as the topology of social networks and WWW networks. This fact makes the discovery of deterministic knowledge of network not possible or not preferrable because it is very challenging to estimate every edge deterministically, and thus the inference of the statistical model becomes more challenging than ever only given the information of random topology of network. These different aspects stimulate the need of new ideas and thoughts on decision and inference on networked data collected from such complex system and networks.

In this work, we investigate the scenario where data about the state of real world is made available to a decision center by a set of *agents* in the network. Each agent has its own belief about the state of real world and can interact with its neighbors, and then, transmits a summery of its information about the state of real world to the decision center. Based on the received data, the decision center produces an estimate of the state of nature to minimize a predefined loss. Necessary inference and learning of data model and network should also be done with these raw data. This scenario is similar to the *decentralized detection* [6], [7], from which there are two major difference to our work. First, conventional decentralized detection focuses on deriving the optimal local decision rule and optimal fusion rule simultaneously, but our work deals

This research was supported by the National Science Council of Taiwan, R.O.C. and Intel Corp. under grants NSC 102-2917-I-002-079 and NSC 102-2219-E-002-019.

with the fusion rule only and assumes the local decision rule is fixed, unknown, and cannot be altered by the fusion center. Second, in the conventional scenario of decentralized detection, it is reasonable to assume that data from distributed source is independent and identical distributed (i.i.d.), but we aim to explore the modeling of networked data persisting random pairwise interactions among them according to the network structure and then derive optimal fusion rule based on such modeling.

The random topological network model has received success in modeling the structure, function, and dynamics of complex networks [8], [9]. However, to the best of our knowledge, a satisfactory modeling of data embedded on random topological networks still remains an open problem. There is a detailed discussion of how to estimate network structure of data and detect target by network filtering with Gaussian graphical model in [10]. In [11], an in-depth analysis of social learning cooperated with stochastic network topology is explored. In the mean while, there are well understood of conventional decentralized detection throughout fundamentals to advanced topics [12], [13]. In recent researches, [14], [15] proposed solutions to the fusion rule of correlated local decisions based on the knowledge of correlation coefficient. [16] also proposed an adaptive fusion rule to achieve asymptotic optimality. [17] proposed an unified framework of distributed detection with conditionally dependent observations. These researches have dealt with some of most decentralized detection problems with deterministic topology. [18] extended this area to the adaptive network by exploring fully distributed and adaptive implementation of distributed detection based on diffusion strategies. These fascinating works inspire our idea of exploring cognition of networked data with stochastic topology as a data fusion problem. Our work is also related to the study of herd behavior in economics [19], which studies the sequential decision making problem in a market.

In this paper, we first formulate the decision problem as a binary hypothesis testing based on statistical decision theory and then a simple statistical model of networked data is proposed. We derive the optimal minimax decision rule based on the knowledge of joint probability density function. In order to mitigate the complexity of inference brought by the the randomness of network topology, a relatively simpler model is proposed and its parameters can be estimated simply in networked data. The performance of decision is compared with decision with complete information and with the performance of conventional case of i.i.d. data. At the end, an experiment based on data crawled from Epinions.com [20] is conducted to illustrate the possibility of applying our method to data analysis in social networks.

II. Problem Formulation

Generally, consider a network consisting of a countably infinite number of interacting agents (individuals) indexed by $i \in \mathbb{N}$. Every agent makes a single *individual decision* about an underlying state of the world θ. For simplicity, we assume that both the underlying state and individual decision are binary. In particular, the individual decision of agent i is denoted by

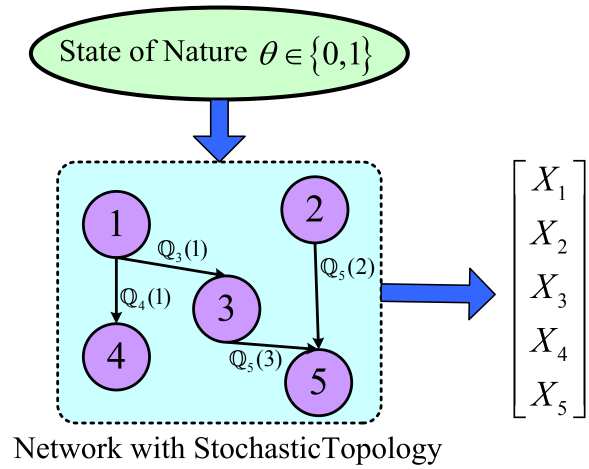

Fig. 1. Networked data of stochastic topology. Consider a small network with only five agents. The network is characterized by the network topology $\{\mathbb{Q}_i\}_{i=1}^5$. The edge set implies the pairwise correlations among data collected from the network. For instance, the edge between the pair of agents $(1,3)$ implies the correlation of (X_1, X_3).

$x_i \in \{1,0\}$ and the underlying state is $\theta \in \{1,0\}$. The goal of decision is to make a decision δ about the state of nature θ given a set of individual decisions collected from the network with respect to minimize some loss.

This decision problem can be regarded as a *hypothesis testing* with $\mathcal{H}_1 : \theta = 1$ versus $\mathcal{H}_0 : \theta = 0$. Given the observation $\mathbf{X} \in \mathsf{X}^n$ and loss function $L(\delta, \theta)$, the fusion center is faced with two possible courses of action $\mathcal{A} = \{a_1, a_0\}$. If $\theta = j$, then action a_j is appropriate. As usual, a fusion rule δ will be represented by a test function $\phi(\mathbf{x})$, where $\phi(\mathbf{x}) = \mathbb{P}\{\delta(\mathbf{x}) = a_1\}$ is the probability of accepting a_1 when $\mathbf{x}$ is observed. Define

$$\begin{aligned} P_F(\phi) &= \mathbb{E}_0[\phi(\mathbf{X})] = \text{false alarm probability.} \\ P_M(\phi) &= \mathbb{E}_1[1 - \phi(\mathbf{X})] = \text{missing probability.} \end{aligned} \tag{1}$$

Then, the risk function can be represented as $R(\mathcal{H}_0, \phi) = wP_F(\phi)$ and $R(\mathcal{H}_1, \phi) = P_M(\phi)$, where $w \in \mathbb{R}$.

The information we have to fuse individual decision is the *joint probability distribution function (joint p.d.f.)* $\mathbb{P}(x_1, \ldots, x_n; \theta)$. In order to express joint p.d.f. explicitly, we have to explore the interaction among agents with a proper model. We restrict that agents can only interact with their neighbors in the network. This interaction can be associated with a graph model $G = (\mathcal{V}, E)$. Each agent corresponds to a vertex in the vertex set $\mathcal{V} = \{1, \ldots, n\}$ and define a (directed) edge set E such that $(i, j) \in E$ if and only if i is j's neighbor. An edge exists between (i, j) implies data sampled from i and j respectively are correlated.

However, the graph we are interested in is the one without deterministic edge set, or *random graph*. In the following subsections, we introduce stochastic network topology first and then discuss the statistical modeling of individual decision.

A. Stochastic Network Topology

We define the network topology to be stochastic, *i.e.*, *stochastic network topology* [11]. The network topology can be described in terms of neighborhood. Given the vertex set $\mathcal{V} = \{1, \ldots, n\}$, the neighborhood of $i \in \mathcal{V}$, denoted by $\mathcal{N}_i$, is a collection of vertices that i connects to in the network, which is often generated stochastically according to an arbitrary probability distribution $\mathbb{Q}_i$ over $\mathcal{V}^{i-1} = \{1, 2, \ldots, i-1\}$. We further assume that the draws from $\mathbb{Q}_i$ are independent from each other for all i. Denote $\{\mathbb{Q}_i\}_{i\in\mathcal{V}}$ by $\mathbb{Q}$. We call $\mathbb{Q}$ the *network topology* and the collection $\mathcal{G} = (\mathcal{V}, \mathbb{Q})$ the *network space*. The network topology is common knowledge, whereas the realized neighborhood $\mathcal{N}_i$ is private information of vertex i. We say that $\mathbb{Q}$ is a *deterministic* network topology if the probability distribution $\mathbb{Q}_i$ is a degenerate (Dirac) distribution for all i. Otherwise, if $\{\mathbb{Q}_i\}$ is not degenerate for some i, $\mathbb{Q}$ is a *stochastic* network topology.

B. Statistical Modeling of Networked Data

Based on stochastic network topology, networked data can be illustrated in Figure 1. The main challenge of statistical modeling of networked data is that, the randomness of stochastic network topology will leads to the uncertainty of joint probability density function, which plays an important role in decision. According to [14], the joint p.d.f. of data set $(X_1, X_2, \ldots, X_n)$ can be derived by defining following finite set of correlation coefficients:

$$\mathcal{C} = \left\{ \mathbb{E}_j \left[\prod_{i\in I} X_i \right] : I \subseteq \mathcal{V}, I \neq \varnothing, j \in \{0,1\} \right\} \tag{2}$$

Let $A_\mu = \{i : x_i = \mu\}$

$$\begin{aligned} &\mathbb{P}_j\{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\} \\ &= \mathbb{E}_j \left[\prod_{i\in A_1} X_i \prod_{i\in A_0} (1 - X_i) \right] \\ &= \sum_{I\subseteq A_0} (-1)^{|I|} \mathbb{E}_j \left[\prod_{i\in A_1\cup I} X_i \right] \end{aligned} \tag{3}$$

This representation suggests that, the decision rule can be implemented given the finite set of correlation coefficients. Therefore, the main challenge of networked data modelling becomes the derivation of the finite set of correlation coefficients with stochastic network topology.

III. Main Results

A. Derivation of Correlation Coefficient

For simplicity, we assume the network is homogeneous such data collected from it persists uniformity in marginal statistical property. That is

Assumption 1 (Uniformity)**.** *Given $(X_1, \ldots, X_n)$ sampling from $G = (\mathcal{V}, \mathbb{Q})$ and $j \in \{0, 1\}$, for all $i, k \in \mathcal{V}$ $i < k$, we have*

$$\begin{aligned} \mathbb{P}_j\{X_i = 1\} = \mathbb{P}_j\{X_k = 1\}; \\ \mathbb{P}\{i \to j\} = \alpha; \end{aligned} \tag{4}$$

Furthermore, if given $\mathcal{N}_i, \mathcal{N}_j \in \mathcal{V}$, and $|\mathcal{N}_i| = |\mathcal{N}_j|$, we have

$$\begin{aligned} &\mathbb{P}_j\{X_i = 1 | X_r = 1, r \in \mathcal{N}_i\} \\ =&\mathbb{P}_j\{X_k = 1 | X_r = 1, r \in \mathcal{N}_j\} \ \forall\ i, k \in \mathcal{V}, \end{aligned} \tag{5}$$

which implies constant conditional pairwise correlation coefficient.

Given an arbitrary subset $\{(1), \ldots, (q)\}$ of $\mathcal{V}$, the correlation coefficient can be expanded as

$$\begin{aligned} &\mathbb{E}_j \left[\prod_{i=1}^{q} X_{(i)} \right] \\ =&\mathbb{E}_j[X_{(1)}] \prod_{k=2}^{q} \mathbb{E}_j\{X_{(k)} | X_{(1)} = 1, \ldots, X_{(k-1)} = 1\} \end{aligned}$$

In particular, consider $\mathcal{A} = \{(1), \ldots, (q-1)\}$. We have

$$\begin{aligned} &\mathbb{E}_j[X_{(q)} | X_k = 1, k \in \mathcal{A}] \\ =& \mathbb{E}\mathbb{E}_j[X_{(q)} | X_k = 1, k \in \mathcal{A} \cap \mathcal{N}_{(q)} \subseteq \mathcal{A}] \\ =& \sum_{\mathcal{N}_{(q)}} \mathbb{P}\{\mathcal{N}_{(q)}\} \mathbb{E}_j[X_{(q)} | X_k = 1, k \in \mathcal{A} \cap \mathcal{N}_{(q)} \subseteq \mathcal{A}], \end{aligned} \tag{6}$$

where $\mathcal{N}_{(q)}$ is an realization of (q)'s neighborhood based on $\mathbb{Q}_{(q)}$. The key idea behind this expansion is that, $X_{(q)}$ is correlated with X_k if X_k is $X_{(q)}$'s neighbor, and the neighborhood composition is modelled stochastically based on $G = (\mathcal{V}, \mathbb{Q})$. Therefore, the correlation coefficient can be expressed as an expectation over all possible combination of neighborhood intersected with $\mathcal{A}$.

Based on **Assumption 1**, the correlation coefficient is independent of labeling but only depends on the cardinality of $\mathcal{A} \cap \mathcal{N}_{(q)}$. We can further derive the expression of correlation coefficient as follow. First, define following notations

$$\begin{aligned} &\epsilon_1^j \triangleq \mathbb{E}_j[X_1];\ q_k^j \triangleq \mathbb{E}_j[X_m | S_k], \\ &\epsilon_m^j \triangleq \mathbb{E}_j[X_m | X_k = 1, \ldots, X_{m-1} = 1]; \end{aligned} \tag{7}$$

where S_k is the event $\{X_r = 1, r \in \mathcal{N}_m, |\mathcal{N}_m| = k\}$. In particular, $S_0 = \varnothing$. Obviously, we have $\epsilon_1^j = q_0^j = \mathbb{E}_j[X_1]$ and

$$\epsilon_m^j = \sum_{k=0}^{m-1} \mathbb{P}\{|\mathcal{N}_m| = k\} q_k. \tag{8}$$

The expression of q_m can be derived by the *partial correlation coefficient* as follow. Let $\mathcal{N}_m = \{i_1, \ldots, i_k\}$ and let I_j represent the event $\{X_r = 1, r \in \{i_1, \ldots, i_{k-1}\}; \mathcal{H}_j\}$. We have

$$\begin{aligned} \rho_j &= \mathsf{Cov}\left(X_m, X_{i_k} | I_j\right) = \frac{A_j}{\sqrt{B_j C_j}} \\ A_j &= \mathbb{E}_j[X_m X_{i_k} | I_j] - \mathbb{E}_j[X_m | I_j] \mathbb{E}_j[X_{i_k} | I_j] \\ B_j &= \mathbb{E}_j[X_m | I_j](1 - \mathbb{E}_j[X_m | I_j]) \\ C_j &= \mathbb{E}_j[X_{i_k} | I_j](1 - \mathbb{E}_j[X_{i_k} | I_j]) \end{aligned} \tag{9}$$

Therefore,

$$\begin{aligned} q_k^j &= \mathbb{E}_j\{X_m = 1 | X_{i_k} = 1, I_j\} = \frac{\mathbb{E}_j[X_m X_{i_k} | I_j]}{\mathbb{E}_j[X_{i_k} | I_j]} \\ &= q_{k-1}^j + \frac{\rho_j \sqrt{q_{k-1}^j \epsilon_{k-1}^j (1 - q_{k-1}^j)(1 - \epsilon_{k-1}^j)}}{\epsilon_k^j} \end{aligned} \tag{10}$$

Since the conditional probability is independent of labeling, we have $\mathbb{E}_j\left[\prod_{i=1}^{q} X_{(i)}\right] = \mathbb{E}_j\left[\prod_{i=1}^{q} X_i\right] = \prod_{i=1}^{q} \epsilon_i^j$. Combining

(8) and (10), we have a iterative method to derive the finite set of correlation coefficients $\mathcal{C}$. A special case of $\mathcal{C}$ is the case with complete graph. We will have $\epsilon_k = q_{k-1} = \epsilon_{k-1} + \rho(1-\epsilon_{k-1}) = 1-(1-\epsilon_1)(1-\rho)^{k-1}$, which leads to the consistent result in correlated binomial distribution.

B. Minimax Decision

To explore more information about θ, we can utilize *individual decisions* collected from the network. Assume

$$\mathbb{P}\{X_i = 1|\theta = j\} = p_j \ \forall\ j \in \{0,1\} \tag{11}$$

Now that we have obtained an expression for the likelihood ratio, the randomization constant γ and the threshold t in terms of minimax criterion can be determined.

Without loss of generality, we enumerate all the possible realizations of $\mathbf{X}$, $\mathbf{x}_i$, $1 \le i \le N = 2^n$ such that $T(\mathbf{x}_1) \le T(\mathbf{x}_2) \le \cdots \le T(\mathbf{x}_N)$. Then, it follows that

$$P_j(T(\mathbf{X}) < t) = \begin{cases} 0, & \text{if } t < T(\mathbf{x}_1) \\ \sum_{k=1}^{i} P_j(\mathbf{x}_k), & \text{if } T(\mathbf{x}_i) \le t < T(\mathbf{x}_{i+1}) \\ 1, & \text{if } t \ge T(\mathbf{x}_N) \end{cases}$$

with $j \in \{0,1\}$, $i \in \{1,\dots,N-1\}$. It is readily shown that the Bayes tests can be written as

$$\phi(\mathbf{x}) = \begin{cases} 1, & \text{if } T(\mathbf{x}) > t_j \\ \gamma, & \text{if } T(\mathbf{x}) = t_j \\ 0, & \text{if } T(\mathbf{x}) < t_j \end{cases} \tag{12}$$

where $t_j = T(\mathbf{x}_j)$. For a test of this form,

$$\begin{aligned} P_F(\phi) &= P_0\{T(\mathbf{X}) > t_j\} + \gamma P_0\{T(\mathbf{X}) = j\}, \\ P_M(\phi) &= P_1\{T(\mathbf{X}) < t_j\} + (1-\gamma) P_1\{T(\mathbf{X}) = j\}. \end{aligned} \tag{13}$$

By equalizer rule, the minimax rule is the decision rule with (t_j, γ) satisfying $wP_F(\phi) = P_M(\phi)$, which can be solved numerically, at least.

IV. Numerical Simulation and Experiment

A. Numerical Simulation

In this numerical simulation, we assume uniform randomness in stochastic network topology. That is, $\mathbb{Q}_i(j) = p_{\text{ER}}$ for all $i \in \mathcal{V}$ and for all $j \in \mathcal{V}^{i-1}$. The size of network is fixed at $|\mathcal{V}| = 50$. The data on stochastic network is generated by the following linear model

$$\mathbb{P}_j\{X_i = +1|x_k, k \in \mathcal{N}_i\} = p_j + \frac{\Delta_p}{i-1}\left[2\sum_{j\in\mathcal{N}_i} x_j - |\mathcal{N}_i|\right]$$

where $\Delta_j \le \min\{p_j, 1-p_j\}$. This model captures the idea that, when i observes an a_1 in its neighborhood, it will increase its probability of choosing a_1, but decreases its probability of choosing a_0 when an a_0 is observed in its neighborhood.

We use this rule to generate 5000 sets of data. The parameter of p_1, p_0, and ρ is estimated from sample mean and sample covariance, respectively. Since each set of data are mutually independent. The sample mean will be the maximum likelihood estimator of its mean, respectively. We simulate the performance of minimax decision with complete information and with estimated parameter, respectively. The decision with complete information means that we compute the likelihood function of data based on the signal model and the parameter p_1, p_0 is known. The decision with estimated parameter is based on (3) with p_1, p_0, and ρ estimated from raw data. We also plot performance of the decision with data follows independent and identical distributions, which can be computed analytically. The simulated performance is depicted in Figure 2. We can discover that, when the number of nodes is low, the performance of decision with estimated parameter is comparable to the one with complete information. However, when the number of nodes goes beyond about 45, the performance of decision with estimated parameter become much worse than the others. The possible reason of this failure is that, the estimated parameter and corresponding derived joint PDF cannot capture the true distribution of high number of data because of the complexity of network composition introduced by high number of nodes. In our formulation, we only focus on expression of joint PDF with pairwise correlation. However, when the number of nodes increase, the correlation patterns become so complex that an averaged correlation coefficient may not be able to capture well.

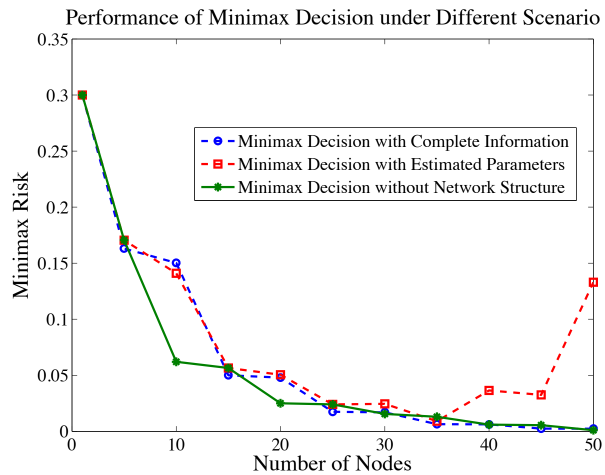

Fig. 2. The simulated minimax risk under different scenario. The performance of our proposed method (decision with estimated parameters) is comparable to the one with complete knowledge of network and parameters when the size of network is not too high, but fails when the number of node goes beyond about 45. The parameter is listed as follow: $p_1 = 0.7$, $p_0 = 0.3$, $w = 1$, $\Delta_p = 0.3$, $p_{\text{ER}} = 0.5$, $\hat{\rho} = 0.22$

B. Experiment

We analyze the dataset collected by Paolo Massa in a 5-week crawl (November/December 2003) from the Epinions.com Web site [20]. There are two type of dataset: `trust data` and `rating data`. In `trust data`, there are total 49,290 users showing their mutual trust relationships. In `rating data`, there are total 139,738 items being rated. First we construct the trust network based on `trust data` and accumulate its empirical distribution. The result indicates that the degree distribution is tend to be *power law* with coefficient $\alpha \cong 1.7$ (Figure 3). This coefficient is used as

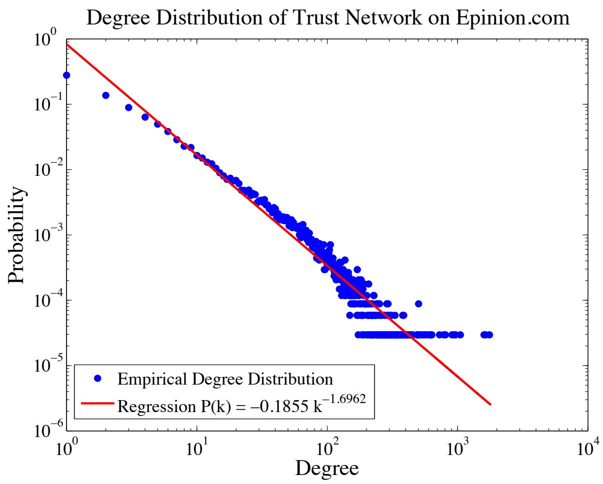

Fig. 3. The empirical degree distribution of trust network in Epinion.com. The fitting curve is derived by linear regression method.

TABLE I
MINIMAX DECISION RESULT OF DATA FROM EPINION.COM

	Hight Rating Item	Low Rating Item
Minimax [a]	107,519(77%)	32,219(23%)
Reference [b]	102,544(73%)	37,194(27%)

[a] $\hat{p}_1 = 0.92$, $\hat{\rho}_1 = 0.06$; $\hat{p}_0 = 0.46$, $\hat{\rho}_0 = 0.18$
[b] Number of mismatch items $n_1 = 1,486$, $n_0 = 6,460$

a reference to generate stochastic network topology in our modeling of data.

We use `rating data` to classify items into two categories: *high rating item* ($\theta = 1$). A reference classifier is use the total average rating, which is $TAR = 3.9$. The item with average rating higher than TAR will be considered as *high rating item*, while the one with lower rating will be classified into *low rating item*. Then we estimate the parameters for our minimax decision method in each group, which is shown in the Table I.

From the experiment result, even though it is hard to say that which method is better, we can still conclude that our proposed minimax decision method is comparable to the reference method, while our method can provide more statistical insights on data. An interesting observation is that, the number of mismatch result between two method is $n_1 = 1,486(1.4\%)$ and $n_2 = 6,460(17.3\%)$, respectively. It indicates that this two method tend to not agree with each other on the group of low rating items. According to the parameter we estimated, in the group of low rating items, $\hat{p}_0 = 0.46$ is not a low value. This fact indicates that many items are classified into low rating item by the reference method because an abnormally low value of rating. For example, an item is rated as $(4,4,4,4,1)$ with averaged rating 3.4, which is lower than the TAR. However, it will be classified into the group of high rating item when it by our method when it is transformed into binary value. This fact is also known as the *outlier* analysis, which is no discussed in this paper.

V. Conclusion and Future Works

In this paper we explored statistical method of decision making with data collected from random topological networks. We focus on how to derive joint probability density functions based on marginal probability densities and pairwise correlations. The performance of our proposed method is comparable to the decision with compete knowledge of network structure and parameters when the number of nodes is not too high, but fails when we increase the size of network, which also introduces more complexity in network structures and correlation patterns. Beside, an experiment based on real world data is also conducted to illustrate possible application of the minimax decision method we proposed. We expect more delicate analysis of performance and more possible solutions to the problem of cognition on networked data with stochastic topology, which we addressed in this paper.

References

[1] A. H. Sayed, S. Barbarossa, S. Theodoridis, and I. Yamada, "Introduction to the issue on adaptation and learning over complex networks," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 2, pp. 161–162, 2013.
[2] T.-Y. Chuang, "Trust with social network learning in e-commerce," in *2010 IEEE International Conference on Communications Workshops (ICC)*, May 2010, pp. 1–6.
[3] P.-Y. Chen, S.-M. Cheng, and K.-C. Chen, "Smart attacks in smart grid communication networks," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 24–29, 2012.
[4] P.-Y. Chen and K.-C. Chen, "Intentional attack and fusion-based defense strategy in complex networks," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, 2011, pp. 1–5.
[5] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. New York: Springer, 1985.
[6] J. N. Tsitsiklis, "Decentralized detection," in *In Advances in Statistical Signal Processing*. JAI Press, 1993, pp. 297–344.
[7] J. F. Chamberland and V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, 2003.
[8] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, pp. 47–97, Jan. 2002.
[9] M. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.
[10] S. Yang and E. D. Kolaczyk, "Target detection via network filtering," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2502–2515, May 2010.
[11] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *LIDS report 2780, Review of Economic Studies*, 2011.
[12] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors i. fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, 1997.
[13] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed detection with multiple sensors i. advanced topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, 1997.
[14] E. Drakopoulos and C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593–606, 1991.
[15] M. Kam, Q. Zhu, and W. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 916–920, 1992.
[16] J. G. Chen and N. Ansari, "Adaptive fusion of correlated local decisions," *IEEE Trans. Syst., Man, Cybern. C*, vol. 28, no. 2, pp. 276–281, 1998.
[17] H. Chen, B. Chen, and P. Varshney, "A new framework for distributed detection with conditionally dependent observations," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1409–1419, 2012.
[18] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 1917–1932, 2011.
[19] A. V. Banerjee, "A simple model of herd behavior," *The Quarterly Journal of Economics*, vol. 107, no. 3, pp. 797–817, 1992.
[20] P. Massa and P. Avesani, "Trust-aware bootstrapping of recommender systems," in *Proceedings of ECAI 2006 Workshop on Recommender Systems*, 2006, pp. 29–33.