Knowledge-Based Systems 226 (2021) 107161

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Research on users' participation mechanisms in virtual tourism communities by Bayesian network

Yinghao Chen [a], Rong Chen [b], Jundong Hou [b], Muzhou Hou [a,*], Xiaoliang Xie [c]

[a] School of Mathematics and Statistics, Central South University, Changsha, Hunan, 410083, China
[b] School of Economics and Management, China University of Geosciences, Wuhan, Hubei, 430074, China
[c] School of Mathematics and Statistics, Hunan University of Technology and Business, Changsha, Hunan, 410205, China

ARTICLE INFO

Article history:
Received 2 November 2020
Received in revised form 19 March 2021
Accepted 17 May 2021
Available online 19 May 2021

Keywords:
Bayesian networks
Virtual tourism community
Causal inference
Grouped average Bayesian network model
Users' participation mechanism

ABSTRACT

In the context of the Internet age, the tourism industry has opened up new development opportunities with the help of Internet technology advancement. It has produced many tourism virtual communities such as TripAdvisor, Ctrip, Mafengwo Many studies have been conducted on user behavior's influencing factors in virtual communities (such as co-creation and participants' value-in-use). However, the studies on the mechanism of user participation in virtual communities are limited. This paper proposes a group average Bayesian network model, which is a data-driven method for obtaining the user participation mechanism's causal network. An induced Bayesian network is used to discover conditional dependence between factors and perform probabilistic inferences. Eleven main factors have been selected, including participation intensity, subjective norm, social identity, group norm, functional value, emotional value, social value, share, interaction, user experience and user satisfaction. We found that user experience, and functional value have the most significant impact on user satisfaction, and social identity plays an essential intermediary role in the participation mechanism. This study enriches the research methods of user participation mechanisms and provides a reference for the virtual tourism community's theoretical research and management practice.

1. Introduction

With the popularity of the Internet and mobile smart devices, people have entered the "Internet Plus" era. This has greatly changed people's traditional lifestyles. Today, tourism is one of the world's largest industries and created more than ten percent of the Global Gross Domestic Product (GDP) [1–3]. Driven by the digitalization, the tourism industry has developed into one of the world's largest online industries [4]. Internet users use dynamic information websites including social networks and virtual communities to express their beliefs and comments about services or products, among other behaviors [5]. The virtual tourism community is an online community based on user participation and interaction. Travel enthusiasts participate in community activities online, share travel guides, and answer confusions in the others' travel gathered in the tourism virtual community. Virtual tourism communities have different participants, depending on dummy network space, based on common interests, and under the constraints of community ecology, and has a knowledge cycle and value-added process formed by the exchange of tourism information. Tourism companies use the virtual tourism community to manage user relationships and understand users' needs to further satisfy user experience and satisfaction, thereby increasing user stickiness and affecting co-creating behavior. Therefore, exploring the causal relationship between the variables has practical significance and management significance.

In the previous research, some researchers proposed impact indicators of user participation in the virtual tourism community. Zeithaml defined perceived value as the overall evaluation of the product's utility by consumers based on perceived benefits and perceived effort [6]. In the study of the virtual tourism community members' participation degree, Wang divided members' participation behavior into two aspects, namely participation level and contribution of members [7]. In social influence studies, Bearden et al. pointed out that when individuals interact in a group, it causes changes in personal perception or behavior [8]. Oliver et al. thought that satisfaction is user satisfaction feedback in the "Expectation Disconfirmation" model [9]. Co-creation arises from Prahalad & Ramaswamy in the concept of co-creation of value, who believe that value is jointly created by the enterprise and the customer, from the perspective of customers [10]. Wang et al. researched knowledge sharing factors in the online travel community, which shows that sharing wish has a positive im-

* Corresponding author.
E-mail addresses: chenyinghao@csu.edu.cn (Y. Chen), houmuzhou@sina.com (M. Hou).

https://doi.org/10.1016/j.knosys.2021.107161

pact on sharing behavior, and trust and reciprocity have positive impacts on sharing wishes [11]. Peters proposed a theory of Proximal three-way decisions and applied it to large-scale social network datasets [12]. Faysal et al. compared some graphical network tools for large-scale network data visualization [13]. Cavallari et al. proposed a novel graph method that first combines nodes into communities and then embeds them into the graph. It effectively improves community detection and node classification [14]. Camacho et al. reviewed and summarized the research on social network analysis from four aspects: knowledge mining, extensibility, information fusion, and visualization [15].

Some relevant variables in tourism virtual communities were conducted through empirical research. Ruyter et al. found a correlation between the level of participation and user satisfaction [16]. Lusch [17] and Gronroos [18] proposed that users create tangible and intangible resources during the co-creation by participating in virtual communities. They will integrate these resources into their consumption or use process. Based on the integrated "push-pull" model, the research of Xu et al. shows that functional value, community usefulness, and ease of use are the main factors that affect user participation in virtual tourism communities [19]. Li et al. proposed a Ctrip model including involvement, statistics, trust, stickiness, and word-of-mouth in a B2C virtual community by empirical data [20]. Munoz-Leiva developed a structural equation model for studying the relationship between cognitive and behavioral variables in the Travel 2.0 website [21]. Integrating the factors proven in the conventional consumer behavior theory (such as word-of-mouth and satisfaction) with the behavior factors specific to the virtual domain (such as stickiness), Gao et al. established a structural equation model to study the relationship between beliefs, attitudes, and continuance behaviors in virtual tourism communities. [22]. Rubio et al. adopted the structural equation model to discuss the influence of the participants' value-in-use on the co-creation behaviors in the virtual community TripAdvisor. [23].

Causal analysis has been widely used both in natural science [24,25] and social science researches [26]. The covariance matrix or regression coefficient between variables was adopted. Besides structured equation model (SEM) [27–29] and multivariable linear regression model [30,31] was used to analyze the relationship between variables. Both methods can be regarded as a special case of the Bayesian network (BN). Different from SEM which requires a fixed structure in advance, BN can study the interrelation directly from the dataset. Traditional SEM methods require a lot of expert knowledge, which is unrealistic in a real-world problem with a large number of nodes. With the rise of artificial intelligence, these data-driven approaches have been increasingly used in research in the field of social sciences [32–39].

Since the Bayesian network was proposed in the 1990s [40], Bayesian networks have been widely used in research in various subjects. In 1996 Microsoft use the Bayesian network for printer troubleshooting [41]. Hewlett-Packard then further refined the diagnostic method and extended it to look-ahead analysis to resolve multiple problems at the same time [42]. Ruiz efficiency measurement of research groups using Bayesian networks [43,44]. Kim used the Bayesian network to study the impact of R&D collaboration on innovative performance in Korea [45]. Alfonso et al. studied 14 well-known bibliometric indices on computer science and artificial intelligence journals in 2011 [46]. Marco Scutari el al. proposed an Average Bayesian Network model to analyze the Malocclusion Data [47] and symptom recovery in chronic whiplash-associated disorders [48]. Varshney et al. used the Bayesian network method to discuss the probability of information dissemination in social networks. And they constructed a similarity model between user interests and content through potential topic information [49].

Our research's interest and originality lie in that it proposes a mixed correlation coefficient to filter the factors of user participation in the virtual tourism community and introduce the average Bayesian network to model these variables then analyze the causal relationship between them. We designed questionnaires and received 882 valid answers within two days. All studies were based on statistics from these answers. Based on expert knowledge to group some variables in the user participation mechanism. We also set up a Bayesian model to analyze the relationship between variable groups in this paper.

The main advantage of our work over earlier research, which analyzed only part of the factors, is that we can analyze the joint probability. Another advantage of our work is obtaining a causal model from data. Compared with prior studies which adopted the Interpretative Structural Modeling method, our methods needs a little expert knowledge but can get a model with higher correlation and higher interpretability.

The structure of this paper is as follows. The "Methods" section reviews some basic concepts about correlation, Bayesian network, and factors we select for studying the users' Participation Mechanism in the virtual tourism community, and also we proposed a method named grouped Average Bayesian Network in this section. The "Results" section presents the dataset we collected, the Bayesian networks are used to learn and analyze the arcs in models. Finally, the "Conclusions" part underlines the original contributions of this study and encourages more future research on this research orientation.

2. Methods

2.1. Analysis of correlation

Correlation analysis is a valuable method of data analysis. A potential relationship can be found by analyzing the connection between different features. The correlation coefficient is a numerical measure of some correlations, showing a statistical relationship between different variables. These variables can be two components of a random multivariable as a known distribution or two observation matrix columns. There are several types of the correlation coefficient, such as Pearson Correlation Coefficient [50], Spearman Correlation Coefficient [51,52], Kendall Correlation Coefficient [53,54], and Maximal Information Coefficient [55,56]. Each has its definition and scope of application, whose coefficients are usually from -1 to $+1$. The more significant the absolute value of the coefficient, the greater the correlation. 0 means the strongest independent.

Pearson Correlation Coefficient (PCC), also referred to as Pearson's r, is a measure of linear correlation between two variables developed by Karl Pearson according to the idea proposed by Francis Galton. For variables x_i and x_j, the PCC $r\left(x_i, x_j\right)$ is given by Eq. (1),

$$r\left(x_i, x_j\right) = \frac{cov\left(x_i, x_j\right)}{\sqrt{var\left(x_i\right) var\left(x_j\right)}} = \frac{\sum_{k=1}^{N}\left(x_i^{(k)} - \overline{x_i}\right)\left(x_j^{(k)} - \overline{x_j}\right)}{\sqrt{\sum_{k=1}^{N}\left(x_i^{(k)} - \overline{x_i}\right)^2}\sqrt{\sum_{k=1}^{N}\left(x_j^{(k)} - \overline{x_j}\right)^2}} \tag{1}$$

where N is the sample size, $x_i^{(k)}$ and $x_j^{(k)}$ is the individual sample point indexed with k; $\overline{x_i} = \frac{\sum_{k=1}^{N} x_i^{(k)}}{N}$ is the mean of sample x_i, and analogously for $\overline{x_j}$.

The Maximal Information Coefficient (MIC) belongs to the maximal information-based nonparametric exploration (MINE) class of statistics. It is a measure of the strength of both the linear and non-linear associations between two variables x_i and x_j. Based on the idea of information gain, David et al. proposed a low-complexity random variable measurement method as follows [55]:

$$MIC\left(x_i, x_j\right) = \max_{a \times b < C} \frac{I\left(x_i, x_j\right)}{\log\left(\min\{a, b\}\right)} \tag{2}$$

where, $I\left(x_i, x_j\right) = \int p\left(x_i, x_j\right) \log \frac{p(x_i x_j)}{p(x_i)p(x_j)} dx_i dx_j$ is the mutual information between random variable x_i and random variable x_j.

We often pay more attention to the linear correlation between variables while considering nonlinear correlation in some real-world problems. We introduce the variable $\theta \in [0, 1]$ for combining two correlation coefficients and propose the Mixed Correlation Coefficient (MCC) for measuring the degree of correlation between two variables. MCC is defined as,

$$MCC\left(x_i, x_j\right) = \theta \cdot r^2\left(x_i, x_j\right) + (1 - \theta)\, MIC\left(x_i, x_j\right) \tag{3}$$

Obviously, the value range of MCC is [0, 1]. R-square and MIC are the special cases of $\theta = 1$ and $\theta = 0$. To reduce the computational complexity, we do not calculate the correlation of the variables themselves and use 0 instead (the correlation of the variables themselves must be 1, but it is not valuable in our research).

Our research is based on the assumption that if there is a causal relationship between two variables, the two variables should have a high correlation. If the maximum value of a variable MCC less than 0.5, we say this variable is independent of others. We only choose the variables, which have a high correlation with at least one other variable.

2.2. Conditional relative average entropy

We use conditional relative average entropy (CARE) to determine the direction of arcs in the initial structure [57]. $CARE\left(x_i \to x_j\right)$ shows the average uncertainty of the random variable x_j with the variable x_i. It can be calculated as follows,

$$CARE\left(x_i \to x_j\right) = \frac{H\left(x_j | x_i\right)}{H\left(x_j\right) \cdot \left|x_j\right|} \tag{4}$$

Where $\left|x_j\right|$ represents the number of states or values of the x_j. If $CARE\left(x_i \to x_j\right) > CARE\left(x_j \to x_i\right)$, the arc's direction is from x_i to x_j, else the arc is from x_j to x_i.

2.3. Bayesian network

The BN model connects a set of variables $\mathcal{X} = \{X_1, X_2, \dots, X_N\}$ by a directed acyclic graph (DAG), where N is the number of variables. Arcs indicate a direct effect between variables, and paths represent indirect effects between variables. A Bayesian Network $BN(G, \Theta)$ consists of structure matrix $G = \langle \mathcal{X}, E \rangle$ and structural parameters $\Theta = \{\Theta_1, \Theta_2, \dots, \Theta_N\}$, where $E = \left\{X_i \to X_j, X_m \to X_n, \dots\right\}$ is the dependencies between nodes, $\Theta_i = P\left(X_i | pa\left(X_i\right)\right)$ is the conditional probability distribution of node X_i and $pa\left(X_i\right)$ represents the parent nodes of node X_i. An example of the Bayesian network is shown in Fig. 1. By the conditional probability formula, we can decompose the global distribution of the variable $\mathcal{X}$ into the local distribution of a single variable X_i, that is:

$$P(\mathcal{X}) = \prod_{i=1}^{N} P\left(X_i | pa\left(X_i\right)\right) \tag{5}$$

Estimating such a model from the dataset, the process of learning a Bayesian network can be divided into two steps:

- Structural learning i.e., learning which arcs are included in the network structure.
- Parameter learning i.e., learning to adjust these dependency intensity parameters.

2.4. Structural learning

In structure learning, learn which arcs are in the network. For a given dataset D, the constraint C and scoring function f for describing which the structural G matches better for the dataset D. The process of structural learning aims to find a network structure that makes the scoring function maximize, i.e.:

$$\max_{C \subset G} f\left(G, D\right) \tag{6}$$

For a Bayesian network with N nodes, Robinson proved that the number of all directed acyclic graphs satisfies the following iteration formula in 1977 [3],

$$\begin{cases} Num_{DAG}(1) = 1 \\ Num_{DAG}(N) = \sum_{i=1}^{N} (-1)^{i+1} \frac{N!}{(N-i)!i!} Num_{DAG}(N-1) \end{cases} \tag{7}$$

As $N = 11$, $Num_{DAG}(10) \approx 5 \times 10^{21}$. It means that when there are many nodes in the network, the traversal method cannot obtain the best model due to the high time complexity.

Some optimization algorithms are often used to search for the best structure of the network, such as k2 algorithm [58,59] and the Hill-Climbing algorithm [60,61]. However, the k2 algorithm needs to give the order of nodes also the maximum number of parent nodes in advance, which requires a large amount of expert knowledge and does not apply to some complex problems. Moreover, k2 algorithm is a greedy algorithm, which is easy to fall into the local optimum solution. The Hill-Climbing algorithm gives a random initial structure at first, usually an empty network or a complete network. The solution is obtained by iteratively selecting the optimal structure from the Bayesian network cluster after transforming one arc. The process of learning the Bayesian network structure by Hill-Climbing algorithm can be expressed as follows,

Step 1. Initialize a Bayesian network structure bs, and calculate its scoring $f\left(bs\right)$;

Step 2. Change one arc of the Bayesian network bs (increase an arc, delete an arc, or change direction) to get a cluster Bayesian networks BS;

Step 3. Calculate the score $f\left(bs^*\right)$ for each network bs^* in cluster BS separately.

Note $bs^m = arg\ \max_{bs^* \in BS} f\left(bs^*\right)$;

Step 4. If $f\left(bs^m\right) > f\left(bs\right)$, $bs \leftarrow bs^m$, return Step 2;

Step 5. Output the optimal Bayesian network structure bs.

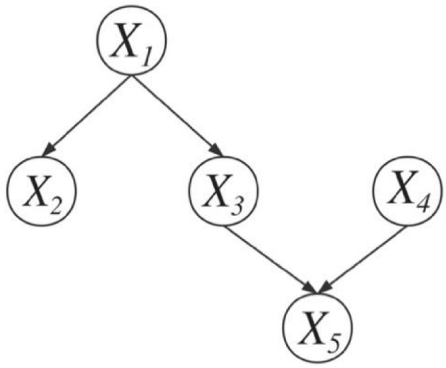
The Hill-Climbing algorithm is too sensitive to the initial solution. When the initial solution is poor, it is easy to fall into a local optimal solution and cause enormous time complexity. A better initial solution can effectively prevent the network from falling into a local maximum, and reduce the number of operations in the structure optimization process.

Some commonly used scoring functions such as Bayesian Information Criterion (BIC) [62,63], Minimum Description Length (MDL) [64,65], Akaike Information Criterion (AIC) [66–69], Verification Data Likelihood (HVL) [70,71], and so on. Expressions of these scoring functions are listed in Table 1.

Generally, the scoring function consists of two parts. The first part is the likelihood function, which describes how well the model structure fitting to the given data. And another part is the penalty term, which is used to prevent the model structure from being too complicated and leading to over-fitting.

2.5. Grouped Average Bayesian Network (GABN)

In some practical problems, there is only a limited sample. Typically, we estimate a single model directly from the data, then draw our conclusions that the model may ignore the fact that the estimated model is not "fixed" [47]. We use model averaging



$$p(X_1 = 0) = 0.1$$
$$p(X_2 = 0|X_2 = 0) = 0.1$$
$$p(X_2 = 0|X_2 = 1) = 0.5$$
$$p(X_3 = 0|X_2 = 0) = 0.8$$
$$p(X_3 = 0|X_2 = 1) = 0.6$$
$$p(X_4 = 0) = 0.7$$
$$p(X_5 = 0|X_3 = 0, X_4 = 0) = 0.6$$
$$p(X_5 = 0|X_3 = 1, X_4 = 0) = 0.4$$
$$p(X_5 = 0|X_3 = 0, X_4 = 1) = 0.3$$
$$p(X_5 = 0|X_3 = 1, X_4 = 1) = 0.6$$

Fig. 1. An example of the Bayesian network, DAG and its condition probability table.

Table 1
Some scoring functions for structure learning.

Scoring function	Expression
BIC	$f_{BIC}\,(G, D) = \sum_{i=1}^{n} \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} m_{ijk} log_2 \theta_{ijk} - \frac{1}{2} \sum_{i=1}^{n} q_i\,(r_i - 1)\, log_2 N$
MDL	$f_{MDL}\,(G, D) = \sum_{i=1}^{n} k_i log_2 n + \frac{log_2 N}{2} \sum_{i=1}^{n} (s_i - 1)\, \lvert\pi_i\rvert + N \sum_{i=1}^{n} H\,(x_i \lvert \pi_i)$
AIC	$f_{AIC}\,(G, D) = \sum_{i=1}^{n} \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} m_{ijk} log_2 \theta_{ijk} - \sum_{i=1}^{n} q_i\,(r_i - 1)$
HVL	$f_{HVL}\,(G, D) = \sum_{i=1}^{n} \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} m_{ijk} log_2 \theta_{ijk}$

Grouped Average Bayesian Network (GABN).

to improve the reliability of structure learning [63,72]. From the central limit theorem, the estimation of the population distribution can be obtained by estimating the subsamples. One hundred different datasets can be obtained by resampling the data using bootstrap. Another problem is that numerous nodes will cause substantial computational complexity. Grouping variables with expert knowledge can reduce the computation time. Firstly, we establish a Bayesian network for variable groups, then calculate the direction of each arc between groups separately. Causality between groups can be used as prior knowledge of the relationship between variables. To reduce the probability of the model falling into a local optimal solution, we use an initial structure based on the MCC and CARE. Given dataset $D = \{x_1, x_2, \ldots, x_N\}$ with M samples, x_i are variables. The flow chart of the Group Average Bayesian Network is shown in Fig. 2 and it can be implemented as follow steps.

Step1. Calculate the MCC between variables, and hold variables with maximum MMC greater than 0.5;
Step2. Resample dataset 100 times;
Step3. Group variable with expert knowledge, and recorded as Y_i;
Step4. For each dataset, calculate MCC for Y_i, and note the maximum for every row as $YMMCC_i$;
Step5. If the mixed correlation coefficient between two nodes satisfies the following conditions, there is an arc between the nodes,

$$\begin{cases} YMMCC_{i,j} > \alpha \cdot YMMCC_i \\ or \\ YMMCC_{i,j} > \alpha \cdot YMMCC_j \end{cases} \tag{8}$$

Step6. For arcs in Step5, determine the direction by CARE, and get the initial Bayesian network structure;
Step7. Optimize the structure by Hill-Climbing algorithm;
Step8. Compute the frequency with which each appears in N graphs, consensus DAG by choosing those arcs with a frequency above κ. Note this structure as YBN;
Step9. For arcs in YBN, calculate MCC for x_i and the maximum for every row as $xMMCC_i$ in each group;
Step10. If the MCC between two nodes x_i and x_j satisfies $\begin{cases} xMMCC_{i,j} > \hat{\alpha} \cdot xMMCC_i \\ or \\ xMMCC_{i,j} > \hat{\alpha} \cdot xMMCC_j \end{cases}$, there is an arc, if they are in different groups, the direction of arc $\widehat{x_i x_j}$ is same as their groups, or it can be determined by CARE.
Step11. Compute the frequency with each one that appears in 100 graphs, then choose arcs that have a high frequency above κ_i.
Step12. Combine all networks and use conditional independent tests to remove redundant arcs.

Remark for the process above, resampling 100 times is enough in most situations, and it can be chosen a bigger one when needed. α and $\hat{\alpha}$ in Step5 and Step10 are the thresholds respectively, and their value are between 0.7 and 0.9, which guarantees the obtained undirected graph contains the arcs that exist in a practical structure. κ and κ_i in Step8 & Step11 are thresholds that can be estimated from the data or just an arbitrary value between $[0.5, 1]$. They are used to obtain a sparse DAG that is easier to interpret.

In fact, if the nodes are not grouped with expert knowledge, that is, each group contains only one node. Then the proposed GABN method degenerates into an improved Hill-Climbing model. Unlike the conventional Hill-Climbing algorithm, GABN can provide a better initial structure by Mixed Correlation Coefficient and CARE, which will greatly improve Hill-Climbing algorithm optimization's efficiency from a random or an empty model. On the other hand, grouping nodes by expert experience can reduce the search space of the model and improve the optimization speed. Since GABN only needs to optimize the causal relationship between groups, and the number of groups is not more than (usually much less than) the number of nodes. Besides, the use of expert experience improves the interpretability of the obtained network structure, which is more conducive to further analysis and discussion by researchers. Moreover, GABN uses the averaging method, which can reduce the model errors caused by individual abnormal sample data.

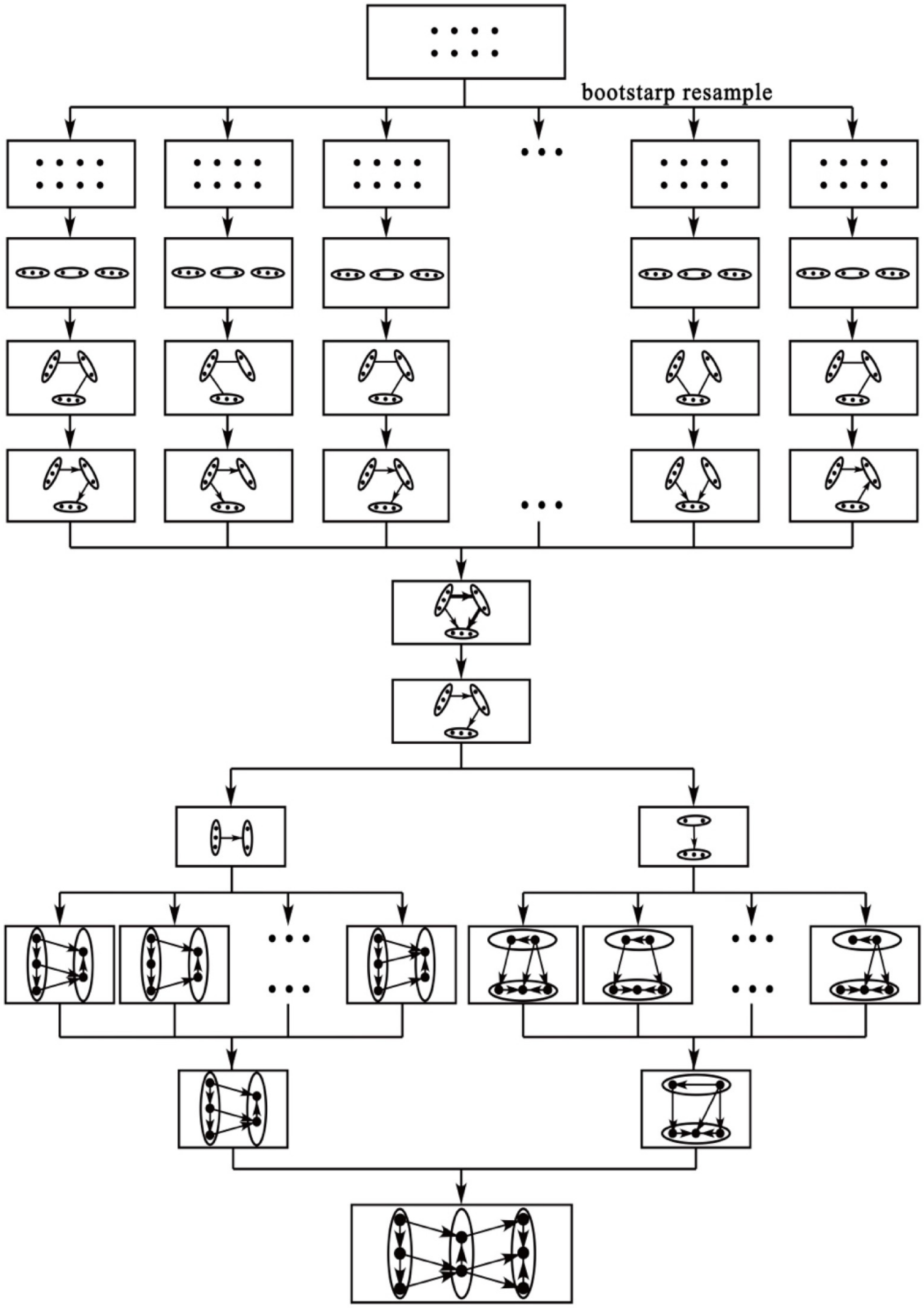


Fig. 2. A GABN algorithm flow chart with 3 groups of 8 nodes.

Main Variables in Bayesian Network

Our Bayesian networks represent relationships between factors in users' participation mechanisms of the virtual tourism community. In particular, each node X_i, in networks represents a specific factor, while arcs between nodes show the causal relationship among these variables. Through learning these Bayesian networks from data, we aim to discover a causal network among the set of variables that we selected. Users who frequently enter the virtual tourism community are often influenced by others' options, and perceive different types of values and experiences through the conflict of thoughts. Many factors in the participation process will cause the user to have a different level of satisfaction with the community. For example, when sharing traveling guidelines, users will pay more attention to popular topics and contents in the community, use the unique sharing form of this community to express their ideas, and hoping to meet community standards or gain recognition from other members. The virtual tourism community realizes its functional value by providing users with introductions to tourist attractions, travel planning strategies, and even some personalized travel route suggestions. Accurate answers and an easy-to-use user interface can improve user experience, thereby enhancing user satisfaction and user stickiness in the virtual tourism community. The main variables we chose include participation intensity, subjective norm, social identity, group norm, functional value, emotional value, social value, share, interaction, user experience, and user satisfaction.

Participation Intensity

The participation intensity in virtual tourism communities is an indicator of its contribution to the community and a reflection of user value. Participation intensity not only refers to the daily dynamics of individuals [7], but also includes the degree of individual interaction with members and community managers in community activities, such as the frequency of responding to comments, interactions, and sharing. Similar with Cho and Lai's research, the hypothesis below is proposed:

Participation intensity is the reason for other variables in virtual tourism communities' participation mechanism [73,74].

Subjective Norm

Obedience means that important people around you think that an individual should do something, the purpose is to accept the influence and take actions to obtain the support of others or groups. The process of obedience is often expressed in subjective norms in researches [75].

Social Identity

The identification process is reflected through social identity, which shows that individuals identify with the community, considering themselves as members of the community, and having a sense of identity and belonging to the community [75].

Group Norm

Internalization means that individuals are willing to accept influence because their goals and values are similar to those of community members through group norms [75].

Functional Value

Functional value is the utility generated by members' perception of the service of the virtual community. When participating in the tourism community, users will be exposed to different tourism products and services. Individuals will have different evaluation feelings when they perceive whether the services they need are met or there is a gap with expectations. Functional value is a subjective perception [76].

Emotional Value

Emotional value refers to the value that community members get due to emotions in interacting with other members [77]. Like most virtual communities, the main characteristics of the virtual tourism community are the timeliness, instability, and vitality of the user community. Individuals have no sense of trust in the virtual identities of community members during the participation process. Only by increasing mutual understanding can they gradually build trust and get emotional satisfaction.

Social Value

Social value is the utility that members gain by using virtual communities to enhance their social status or self-realization [78]. The virtual tourism community can be manifested in the improvement of social relations and friend's intimacy.

Share

In an open virtual tourism community, sharing means that users recommend sharing community resources to others for use [79]. It is conducive to attract more new members and expand the influence of the virtual tourism community.

Interaction

Interaction refers to communication and feedback behavior between users and other subjects. Interactions are conducive to building harmonious community relationships between users and community members and are more conducive to community manager innovation in products and services [80]. Besides, interaction is the basis for co-creation, and it is necessary to be studied.

User Experience

User experience means that users participating in the virtual tourism community will experience different virtual community services, such as the convenience of operation, and the user's experience during the experience [81]. Through these experiences, the user will have a rough evaluation of the virtual tourism community. Based on this evaluation, the user will decide whether to continue to participate or leave.

User Satisfaction

User satisfaction shows the difference between actual feelings and expectations in the activities of the virtual tourism community. Cardozo et al. believe that improving customer satisfaction will lead to re-purchasing behavior without changing other products' perspective [46]. At the same time, user satisfaction is an important means to evaluate the performance of enterprise management systems [82].

Social Influence

In exploring the participation mechanism in the virtual tourism community, we found that social influence is a universal factor in different kinds of virtual network communities. Latané's research defines social influence as a common socio-psychological phenomenon in human life [83]. He believes that no matter humans or animals, their behaviors, emotions, and even thoughts will be more or less affected and changed by others. The influence of others on us is objective, especially based on the vitality of the virtual tourism community. This kind of influence cannot be ignored. Kelman defined the process of social influence from the perspective of categories, and he thought social influence is not indiscriminate [75]. The process mainly includes the mechanisms of obedience, identification, and internalization. This paper considers subjective norms, social identity, and group norms as the main social influence components.

Perceived Value

Customers participate in virtual communities mainly to obtain travel information, make friends, or purchase travel products. Users have clear needs and goals in the participation process, which reflect the role of perceived value. It mainly includes the relative relationship between the benefits obtained and the costs paid. It also consists of the value of the user's emotions and experience during the perception process [84]. Perceived value is a multi-dimensional variable [85,86], which has different division dimensions in various studies. In the virtual tourism community, the user's perceived value mainly includes functional value, emotional value, and social value.

Co-Creation

Co-creation in the virtual tourism community refers to how users actively cooperate with community members and community managers in using products or feeling services [87,88]. Different from previous concepts, the main body of co-creation is the user himself, which means that personalized user experience is beginning to be valued. From the ultimate purpose of co-creation, our primary research is the co-creation of value [89]. In the virtual tourism community, the behavior that best reflects the value creation of the user is the user's sharing and interaction behavior [90].

3. Experiments and results

Data collection

In this study, we designed a questionnaire and distributed it on the Internet (https://www.wjx.cn/jq/49696513.aspx). We

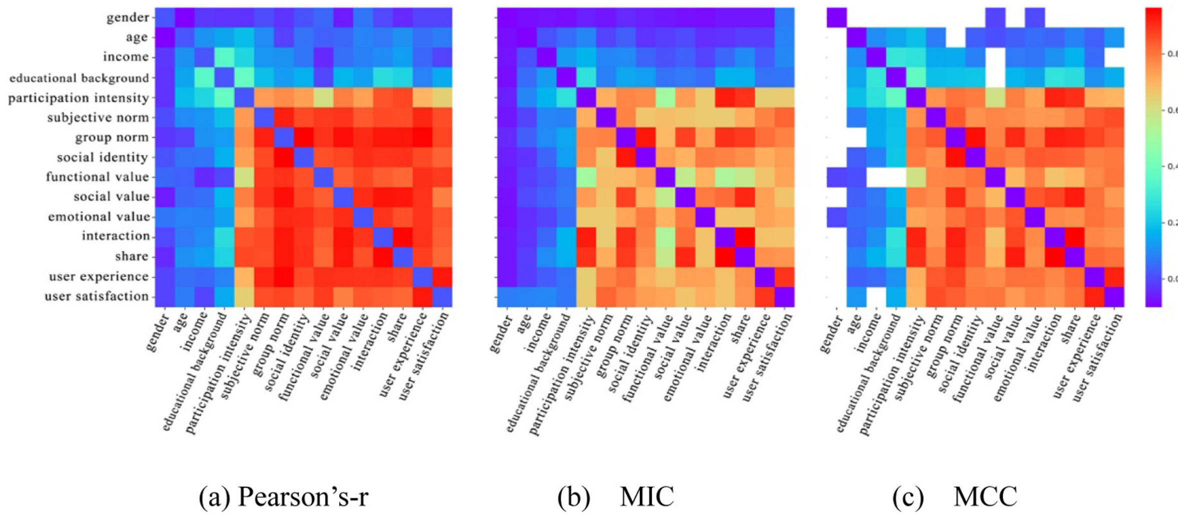


Fig. 3. Correlation coefficient matrix between variables.

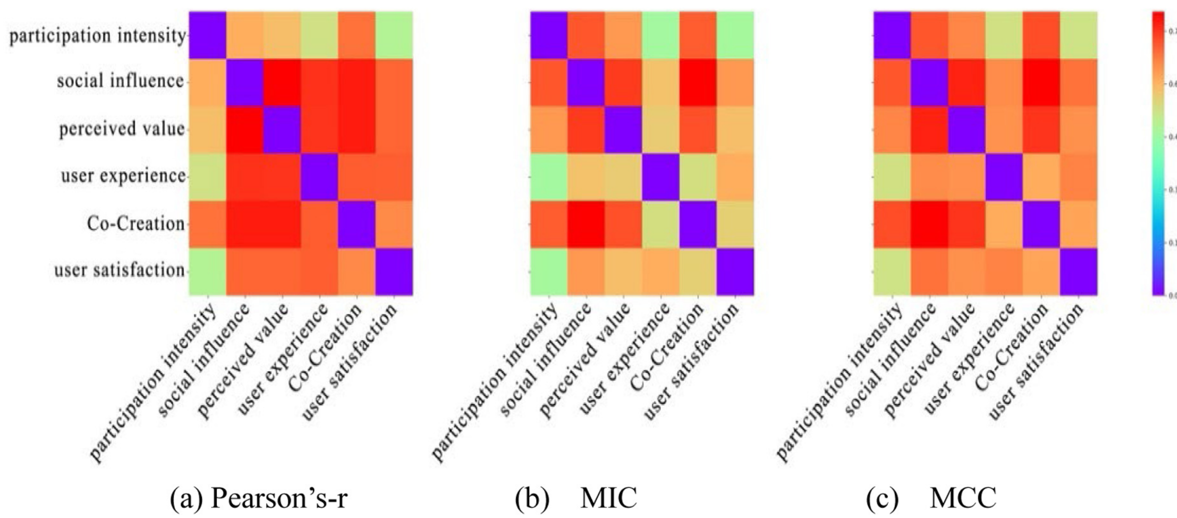


Fig. 4. Correlation coefficient matrix between variable groups.

received 900 users' responses, of which 882 were valid questionnaires. The *Cronbach*'*sAlpha* of our questionnaire is 0.969 and the KOM value is 0.982. We deleted samples that completed all questions within one minute (2.04%). The sample was female-dominated (59.86%). Most of them are young people, and 68% of the respondents are under 27 years old. In terms of education level, half of them are bachelor's degree holders, followed by the master' degree and Ph.D. (17.35%), high school (14.83%), junior college (13.60%) holders. The monthly income levels (in CNY) of the respondents are in the order of ￥ 1001-￥ 2000 (40.00%), ￥ 2001-￥ 3000 (21.03%), ￥ 3001-￥ 4000 (15.87%), ￥ 1-￥ 1000 (13.03%), others (10.07%). All of the respondents finished this questionnaire within ten minutes.

Model Building

In the first step, we calculated the correlation between all the variables we have, including gender, age, income, educational background, participation intensity, subjective norm, social identity, group norm, functional value, emotional value, social value, share, interaction, user experience, and user satisfaction. The correlation coefficient matrix between variables is shown below in Fig. 3. The weight of linear correlation is be chosen as $\theta = 0.3$.

Obviously, from Fig. 3, gender, age, income, and educational background have a lower relationship between others, whether linear or non-linear correlation. We choose participation intensity, subjective norm, social identity, group norm, functional value, emotional value, social value, share, interaction, user experience, and user satisfaction as main factors in the following steps.

In order to lessen the search space of the optimal model and reduce the computational complexity, we group the variables based on expert knowledge. Fig. 4 shows the correlation coefficient matrix of variable groups. Most groups are in high linear correlation, at the same time some groups have a nonlinear correlation. MCC matrix integrated both linear and nonlinear correlations.

Resample dataset 100 times. Learning the structure from each dataset, establish the initial network by MCC-CARE with $\alpha = 0.85$, then optimize the model using the Hill-Climbing algorithm. Table 2 compares the average optimization steps used to obtain the best network structure using different methods to obtain the optimized solution. The average hamming distance between the initial network and the optimal structure is 1.93. In most subsets, it only takes not more than two steps of iteration during optimization. Compared to a fully connected network or an empty

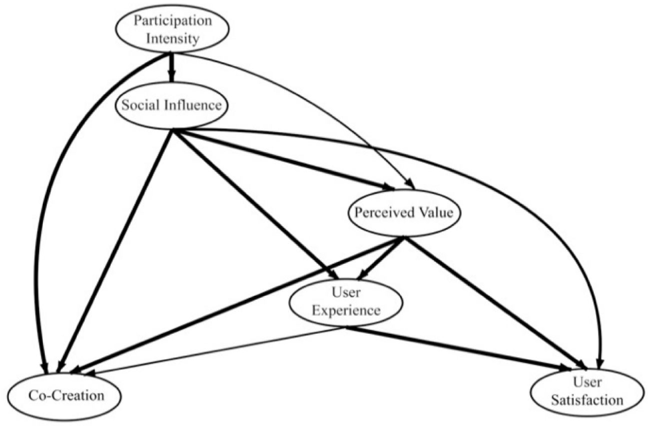


Fig. 5. The DAG underlying the consensus Bayesian network of variable groups was learned. The thickness of the arcs is in the proportion of their frequency. Only arcs with a frequency greater than 0.5 are included in this network.

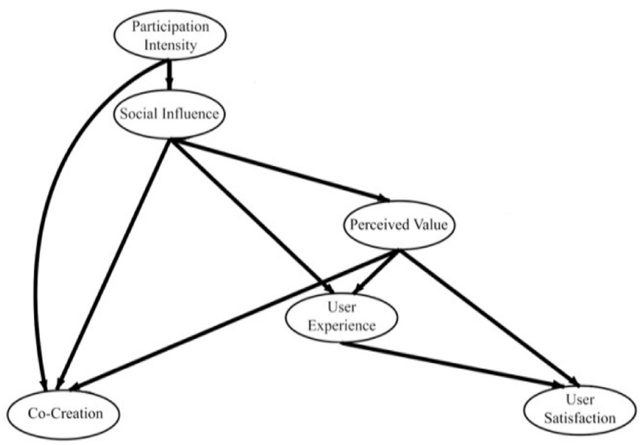


Fig. 6. A simplified DAG was derived from Fig. 5 after removing arcs with a strength smaller than 0.8.

network as the initial network, CARE-MCC is much faster than others.

The Bayesian network consensus is constructed by learning 100 networks from each resampled subset. Keeping arcs that appear at least in half of the model is shown in Fig. 5. All the direction of each arc is well established. Social influence, perceived value, and user experience seem like a mediation role in user satisfaction. Furthermore, deleting arcs with a frequency less than 0.8, a simplified DAG is shown in Fig. 6. In fact, the strength of the remaining arcs is greater than 0.9. The arc from social influence to perceived value and co-creation can be learned in every resampled dataset, which shows a high causal relationship. Social influence reflects the extent to which one's attitudes, beliefs, and behaviors are affected by others, so the users' value will also change with social influence. On the other hand, social influence and perceived value are manifestations of users' psychological activities. Generally, the creation and effect of shared behavior depend on many factors, including psychological activities.

Participation intensity directly affects social Influence and co-creation. Active users in the virtual community have more opportunities to develop friendships with other members, understand the activities of the community. They are more easily affected by community members and the community environment. The more engaged a user is, the more active he or she is in community activities. Responding to other community members for help or expresses views and opinions in the community, which enables users to work with the virtual community to create value together. User experience will be affected by perceived value. In virtual communities, besides satisfying utilitarian needs, users' experiences such as emotions and aesthetics during the consumption process are also important.

For each arc in variable groups Bayesian network, we learn its structure by MCC and CARE respectively. The directions are the same as their groups. $\hat{\alpha}$ is chosen as 0.75 and the threshold is $\kappa_i = 0.7$. The average Bayesian networks are shown in Fig. 7.

Similarly, we can find the mediation of sharing in community interaction in Fig. 7 (b, e, h). Users' perception of social value helps users gain a certain status and support in the community. Users will be more willing to participate in the co-creation of the community to enhance their value, while self-realizing. Socially valued users often attract fans through frequently sharing behaviors to increase their visibility and influence. On the other hand, interactions with community members will also better maintain their friendship with supporters. Sharing affects interaction because sharing provides an effective channel for other team members to interact with users. For example, fans can learn about travel destinations by browsing travel notes and can discuss further with users by posting comments. Therefore, sharing has a significant impact on interaction. When users participate in virtual tourism communities, they often browse information and search for content. Users often recommend practical and valuable knowledge content to their friends, hoping that they can also get information from it. Participation intensity is the active performance of users in the community. By sharing with other members, users can meet more members with common hobbies or interests, strengthen the relationship between friends, and interact more frequently. It also proves that sharing has a mediating effect on interaction.

Combining variable groups Bayesian networks, we obtain the causality of variables in the users' participation mechanism of the virtual tourism community in Fig. 8. Some arcs are removed because of their low correlation, such as social identity $\longrightarrow$ emotional value. Social identity includes both group identity and self-identity, based on the user's subjective feelings. But emotional value pays more attention to the perceived value generated by emotional communication, which shows that when users have a stronger sense of social identity, they will make more subjective judgments rather than "motional affairs".

At the same time, other edges were deleted due to intermediation, such as subject norm $\longrightarrow$ share. The purpose of users being affected by subjective norms is to gain support from others, integrate into the collective, and avoid being ignored. Subjective norms can affect sharing in order to be consistent with friends' sharing behavior. But users are free individuals and have their behavioral awareness. The impact of subjective norms gradually weakens as users become familiar with the community. Users are free individuals and have their sense of behavior. As users become more familiar with the community, the impact of subjective norms gradually diminishes. This shows that subjective norms and sharing behaviors are independent of respect to social value conditions.

Through the Bayesian causal reasoning network of user participation mechanism in the tourism virtual community, community managers can get some ideas, so as to improve the value and user satisfaction of the virtual community. The most important is participation intensity. Only when users join and participate in the virtual community will it be possible to co-create with the community. This requires the managers of the virtual tourism community to promote and allow more users to join the virtual community.

Another important result of the network is that social identity plays an important intermediary role. According to Table 3

Table 2
Comparison of average optimization steps of different methods.

Methods	GABN	Hill-Climbing form random model	Hill-Climbing form empty model	K2 algorithm
Average steps	1.93	10.51	9.27	10.19

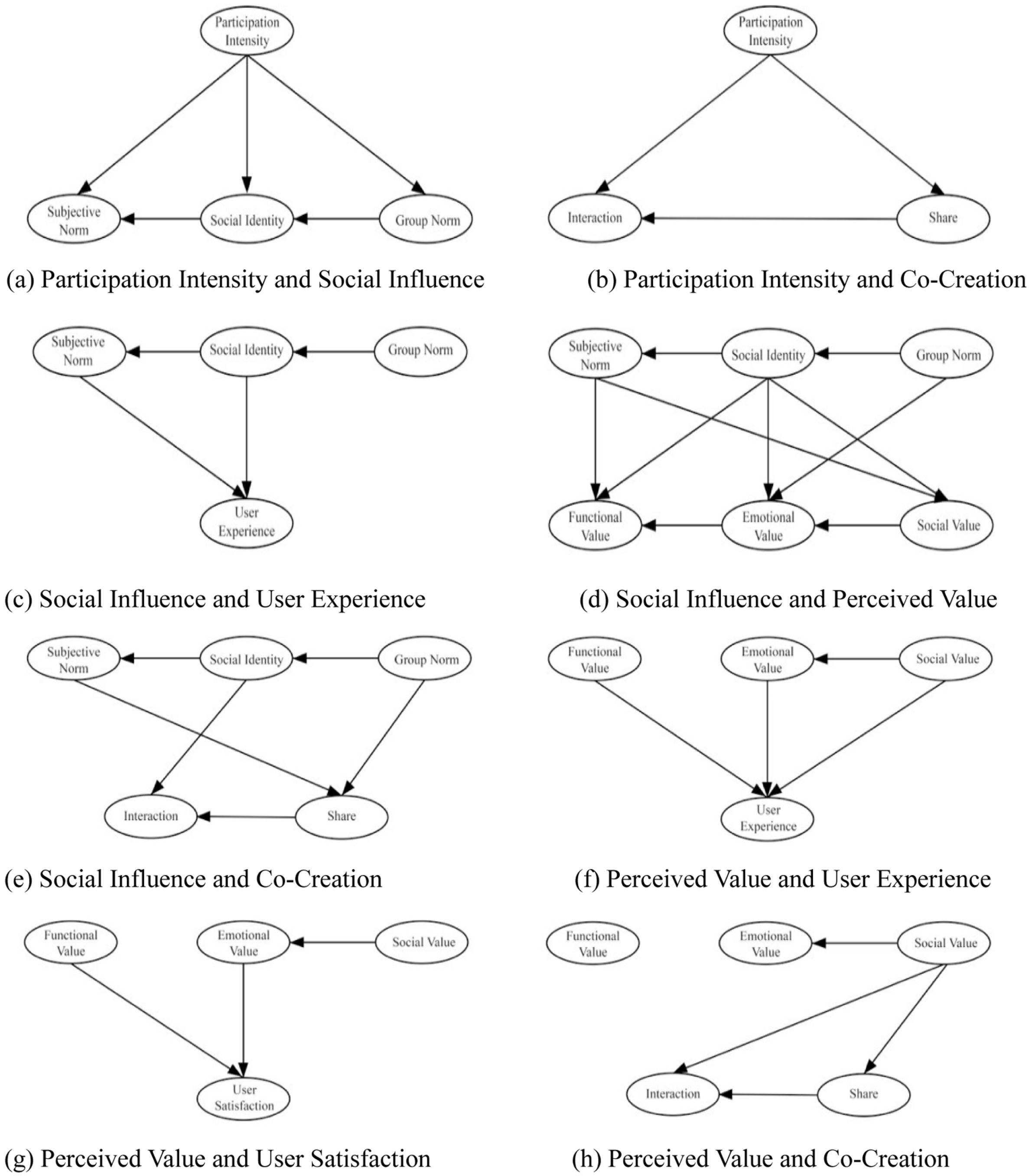


Fig. 7. The Bayesian network between variables of arcs in Fig. 6 respectively.

above, social identity has a lot of influence on other indices. Social identity can affect function value. The improvement of user identity will enhance people's self-satisfaction and the functional perception of the community will be more inclusive. In the meantime, the functional value will also affect the user experience. Users get information or share guides from the virtual tourism community, then gradually feel the community's convenience and usefulness. The user's perception of the functional value based on the services of the virtual community directly affects the user's evaluation of the service function. Users are more likely attention to the functionality and practicality of the community. Users feel that the virtual community can easily meet their needs, thereby increasing their satisfaction and willingness to continue using it. Participation Intensity and group norm have a positive impact on user identity in the virtual community. When users get a high sense of identity and belonging, this emotion gives

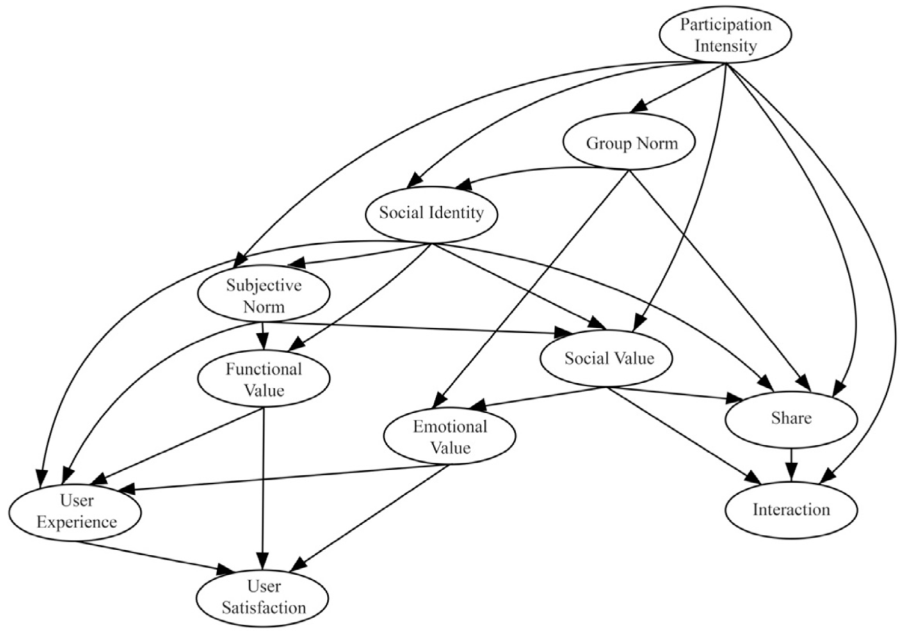


Fig. 8. The Bayesian network of users' participation mechanism.

Table 3
Measures of centrality for the indicators in users' participation mechanism network.

Variables	Indegree	Outdegree	Total
Participation intensity	0	**6**	6
Subjective norm	3	3	6
Social identity	2	5	**7**
Group norm	1	3	4
Functional value	2	2	4
Emotional value	2	2	4
Social value	3	3	6
User experience	4	1	5
Share	4	1	5
Interaction	3	0	3
User satisfaction	3	0	3

users the motivation to share in the community and helps users realize their social value. On the other hand, with the enhancement of user identity, their subjective consciousness will also increase, and subjective norms will not be disturbed by the external environment, enabling users to make subjective judgments on the value of community functions. Based on our causal network graph, operators of the virtual tourism community should take measures to improve users' sense of identity and belonging. Based on the characteristics of the virtual tourism community, providing users with a travel experience, promoting online and offline interaction and experience between users, enhance the cohesion and stickiness between users and the community and jointly co-create greater value for the community.

User experience comes from the user's spontaneous internal reaction. Through our research, we found that user experience can directly influence users' evaluation of virtual communities. It is related to the stickiness of users to communities and their willingness to continue using virtual communities. Managers should focus on the role of user experience in community development. Therefore, we recommend that operators of the virtual tourism community should continuously improve the functions of the community and establish simple user interfaces and operating methods. Protect user privacy, and thus users can operate and experience conveniently and comfortably. Increase the enthusiasm and community involvement of community members. Make it easier for users to obtain and publish information in the community, provide convenience for users to communicate with other members, and continuously improve user experience in the virtual tourism community

User satisfaction is the overall evaluation of the community, so the managers of the virtual tourism community should take measures to improve the level of user satisfaction. Based on the network we built, the improvement of user satisfaction needs to be achieved through user experience and perceived emotional value. Users will continue to enjoy the experience brought by the product or service when they perceive the results as being beneficial to themselves. It can be realized by material and spiritual incentives, such as the virtual tourism community, tourism talent honorary title, and so on. Users will enjoy sharing and improving their user satisfaction. In addition, managers need to provide a harmonious community communication environment and establish a good feedback channel. Users can make friends in the virtual tourism community and have a relaxed and pleasant mood, so as to deepen the emotional connection with the community.

4. Conclusion

In this paper, we proposed a Grouped Average Bayesian Network method to find the most suitable structure for analyzing users' participation mechanisms in the virtual tourism community. Grouping variables with expert knowledge can reduce computational complexity and improve the interpretability of the

Table 4
Questionnaire designed for this research.

Factors		Question
Personal information	Gender	My gender.
	Age	My age.
	Income	My income level.
	Educational	My education background.
	Community	Which virtual tourism community is being used.
Participation intensity		The average time you spend in the virtual tourism community each day. The average number of logins to the community per week. I often update my personal homepage. I am active in a virtual tourism community.
Social influence	Subjective norm	I have friends in this virtual travel community. People who have influence on me think I should use this community.
	Social identity	My personal image matches the image of the virtual tourism community. I feel that my world outlook is consistent with the values of community. I have a sense of belonging to this virtual tourism community. I am a valuable member of this virtual tourism community. I am an important part of this virtual tourism community.
	Group norm	I clearly understand the goals of other members of the community. The goals of participating in this community are same as other members.
Perceived value	Functional value	I can get valuable information and knowledge in this community. I can get services or help in this virtual tourism community in time.
	Social value	I feel trust in this virtual tourism community. I have friendship and keep in touch with members of this community.
	Emotional value	I feel relieved and relax in this virtual tourism community. Participating in the activities of the community will make me happy.
User experience		I think this virtual tourism community is easy to operate. I will use this virtual tourism community for a long time.
Co-Creation	Share	I often share and repost information on this community. I often post status and express views and opinions on this community.
	Interaction	I actively participate in the interactive activities of this community. I often help other members of this virtual tourism community.
User satisfaction		I am satisfaction with this virtual tourism community.

model. Considering the non-linear correlation between the indicators in practical problems, the linear correlation will receive more attention, we compound Pearson's r and MIC then gets the Mixed Correlation Coefficient to describe the relativity between indicators. Also, we initialize the network by MCC and CARE to improve the efficiency of model optimization. In order to get a stable model, an averaging method is used to consensus DAG by choosing high-frequency arcs. We simplify the network to better explain the causality between indicators.

The Bayesian network has some advantages as a tool for building a graphical causality model of users' participation mechanisms in the virtual tourism community. On the one hand, the Bayesian network is a graph-based model of the conditional probability distribution that captures correlation and conditional independence between variables. On the other hand, the Bayesian network is a data-driving model that can learn the structure just from the training samples. Moreover, Bayesian networks can easily incorporate expert information, which dramatically improves the reliability and interpretability of the model.

It is the first time that a Bayesian network-based method has been used to study user participation mechanisms in the virtual tourism community. Also, it's the first time to study 11 influential factors in the virtual tourism community's user participation mechanism in one model. In structural learning, the Bayesian network can be regarded as a promising tool, which can be used to model and analyze the causality relationship between indicators of users' participation mechanism in virtual tourism communities. The Bayesian method has no limit on the number of variables analyzed. Except for the eleven indicators we have chosen in this paper, many other variables of the virtual community participation mechanism can also be increased in future research.

Analyzing the Bayesian network of users' participation mechanism, we found that participation intensity and social identity are the most important factors in the users' participation mechanism of the virtual tourism community. On the one hand, it can strengthen the vitality of the community, and improve the sense of belonging and identity of users by encouraging users to participate and improve their participation. On the other hand, with the improvement of user identity, the possibility of user stickiness and value co-creation will be significantly improved. The research shows that by encouraging users to share their views, users can participate in the interaction more frequently and promote the development of the virtual community. Besides modeling a network for all variables, we also obtain the causal network between groups. The causal relationship between the variable groups can be confirmed with the hypothesis in the previous study. On the other hand, this shows that the models we built in this paper are reasonable.

Based on the causality network we obtained, we found that user satisfaction and co-creation both are outcome variables of user participation mechanisms in the virtual tourism community. Simultaneously, emotional value and functional value will directly affect user satisfaction, proving that the customer pays more attention to practicality and perceived value. Both social identity and subjective norms have a significant impact on co-creation, while subjective norms do not. This not only provides a new research direction for the participation mechanism of value co-creation, but also provides a two-way expansion for further research. Our model is not only suitable for the virtual tourism community, but also can provide references for other virtual communities.

In the future studies, we intend to select more indicators for analysis, optimize the setting of questions and collect feedback from more users of various ages in different countries, to optimize the structure of the network, hoping to obtain a more reliable and explanatory model. Besides, we will provide references for the development of virtual communities, especially virtual tourism communities, to help them provide better services for the users.

Data transparency: The data used to support the findings of this study are available from the corresponding author upon request. Our questionnaire can be found on the internet: https://www.wjx.cn/jq/49696513.aspx. The questions of the questionnaire have attached in the appendix. All programs in this paper are implemented based on python.

CRediT authorship contribution statement

Yinghao Chen: Study conception and design, Conceptualization, Methodology, Formal analysis, Writing - original draft, Read and approved the final manuscript. **Rong Chen:** Study conception and design, Conceptualization, Methodology, Visualization, Writing - original draft, Read and approved the final manuscript. **Jundong Hou:** Study conception and design, Writing - review & editing, Supervision, Read and approved the final manuscript. **Muzhou Hou:** Study conception and design, Conceptualization, Writing - review & editing, Supervision, Funding acquisition, Read and approved the final manuscript. **Xiaoliang Xie:** Study conception and design, Writing - review & editing, Project administration, Funding acquisition, Read and approved the final manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was funded by Projects of the National Social Science Foundation of China (No: 19BTJ011). The author would like to thank the anonymous reviewers for their valuable comments that helped improve the quality of this paper, and thank Mr. Yan Hui for his language editing service. All authors approved the version of the manuscript to be published.

Appendix

See Table 4.

References

[1] Y. Xing, B. Dangerfield, Modelling the sustainability of mass tourism in island tourist economies, J. Oper. Res. Soc. 62 (9) (2011) 1742–1752.
[2] J.O. Meynecke, R. Richards, O. Sahin, Whale watch or no watch: the Australian whale watching tourism industry and climate change, Reg. Environ. Change (2016) 1–12.
[3] R.W. Robinson, Counting unlabeled acyclic digraphs, 1977.
[4] A.P. Hosamani, J.S. Vadiraj, An empirical study on customer engagement practices in tourism industry, Int. J. Mark. Hum. Resour. Manag. (2013).
[5] A. Patino, D.A. Pitta, R. Quinones, Social media\s emerging importance in market research, J. Consum. Mark. 29 (3) (2012) 233–237.
[6] V.A. Zeithaml, Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence, J. Mark. 52 (3) (1988) 2–22.
[7] Y.C. Wang, D.R. Fesenmaier, Towards understanding members' general participation in and active contribution to an online travel community, Tour. Manag. 25 (6) (2004) 709–722.
[8] W.O. Bearden, R.G. Netemeyer, J.E. Teel, Measurement of consumer susceptibility to interpersonal influence, J. Consum. Res. 15 (4) (1989) 473.
[9] R.L. Oliver, J.E. Swan, Equity and disconfirmation perceptions as influences on merchant and product satisfaction, J. Consum. Res. 16 (3) (1989) 372.
[10] C.K. Prahalad, V. Ramaswamy, Co-creating unique value with customers, Strategy Leadersh. 32 (2004) 4–9.
[11] J. Wang, et al., Empirical research on knowledge sharing factors in travel online community, Pakistan J. Statist. 30 (5) (2014) 815–826.
[12] J.F. Peters, S. Ramanna, Proximal three-way decisions: Theory and applications in social networks, Knowl.-Based Syst. 91 (2016) 4–15.
[13] M.A.M. Faysal, S. Arifuzzaman, A comparative analysis of large-scale network visualization tools, in: N. Abe, et al. (Eds.) 2018 IEEE International Conference on Big Data, 2018, pp. 4837-4843.
[14] S. Cavallari, et al., Embedding both finite and infinite communities on graphs, IEEE Comput. Intell. Mag. 14 (3) (2019) 39–50.
[15] D. Camacho, et al., The four dimensions of social network analysis: An overview of research methods, applications, and software tools, Inf. Fusion 63 (2020) 88–120.
[16] k. Ruyter, Customer loyalty in extended service settings, Int. J. Serv. Ind. Manag. 10 (3) (1999) 320–336.
[17] R. Lusch, S. Vargo, Service-dominant logic: Reactions, reflections and refinements, Mark. Theory 6 (2006) 281.
[18] C. Gronroos, A. Ravald, Service as business logic: implications for value creation and marketing, J. Serv. Manag. 22 (1) (2011) 5–22.
[19] M.K. Hsu, Y. Huang, F. Huang, Exploring the determinants of virtual tourist community participation, in: X. Feng, L. Sun (Eds.) Proceedings of International Symposium on Green Hospitality and Tourism Management, 2010, pp. 389-397.
[20] G. Li, S. Elliot, C. Choi, Electronic word-of-mouth in B2C virtual communities:An empirical study from ctrip.com, 20, (3) 2010, pp. 262–268.
[21] F. Munoz-Leiva, J. Hernandez, J. Sanchez, Generalising user behaviour in online travel sites through the Travel 2.0 website acceptance model, Online Infor. Rev. 36 (6) (2012) 879–902.
[22] L. Gao, X. Bai, A. Park, Understanding sustained participation in virtual travel communities from the perspectives of is success model and flow theory, J. Hosp. Tour. Res. 41 (4) (2017) 475–509.
[23] N. Rubio, N. Villasenor, M. Jesus Yague, Does use of different platforms influence the relationship between cocreation value-in-use and participants' cocreation behaviors? An application in third-party managed virtual communities, Complexity (2019).
[24] T.N. Mane, M.B. Nagori, S.A. Agrawal, FMRI data analysis with dynamic causal modeling and Bayesian networks, in: C.S. Zhang (Ed.), Materials Science and Information Technology, Pts 1-8, 2012, pp. 5303–5307.
[25] W.W. Wu, Linking Bayesian networks and PLS path modeling for causal analysis, Expert Syst. Appl. 37 (1) (2010) 134–139.
[26] L.C. McCandless, J.M. Somers, Bayesian sensitivity analysis for unmeasured confounding in causal mediation analysis, Stat. Methods Med. Res. 28 (2) (2019) 515–531.
[27] J.L. Howard, et al., Using bifactor exploratory structural equation modeling to test for a continuum structure of motivation, J. Manag. 44 (7) (2018) 2638–2664.
[28] H.W. Marsh, et al., A new look at the big five factor structure through exploratory structural equation modeling, Psychol. Assess. 22 (3) (2010) 471–491.
[29] D. Pepe, J.H. Do, Analyzing apomorphine-mediated effects in a cell model for parkinson's disease with partial least squares structure equation modeling, J. Comput. Biol. (2019).
[30] R.A. Kleiv, K.L. Sandvik, Modelling copper adsorption on olivine process dust using a simple linear multivariable regression model, Miner. Eng. 15 (10) (2002) 737–744.
[31] Y. Xu, et al., Predicting available transfer capability for power system with large wind farms based on multivariable linear regression models, in: 2014 IEEE Pes Asia-Pacific Power and Energy Engineering Conference, 2014.
[32] Y. Chen, et al., A deep residual compensation extreme learning machine and applications, J. Forecast. 39 (6) (2020) 986–999.
[33] Y. Chen, et al., Solution of ruin probability for continuous time model based on block trigonometric exponential neural network, Symmetry-Basel 12 (6) (2020).
[34] Y. Chen, et al., Numerical solving of the generalized black-scholes differential equation using Laguerre neural network, Digit. Signal Process. 112 (2021) 103003.
[35] P. Wang, Y. Lou, L. Lei, Research on stock price prediction based on BP wavelet neural network with mexico hat wavelet basis, in: S.B. Tsai and J. Liu (Eds.) Proceedings of the 2017 International Conference on Education, Economics and Management Research, 2017, pp. 99-102.
[36] F. Yakuwa, et al., Novel time series analysis and prediction of stock trading using fractal theory and time delayed neural network, in: 2003 IEEE International Conference on Systems, Man and Cybernetics, Vols. 1-5, Conference Proceedings, 2003, pp. 134-141.
[37] M. Baur, et al., Software for visual social network analysis, in: P. Mutzel, M. Junger, S. Leipert (Eds.), Graph Drawing, 2002, pp. 463–464.
[38] S. Min, et al., STGSN - a spatial-temporal graph neural network framework for time-evolving social networks, Knowl.-Based Syst. (2021) 214.

[39] A. SeyedHassani, M.S. Haghighi, A. Khonsari, Bayesian inference of private social network links using prior information and propagated data, J. Parallel Distrib. Comput. 125 (2019) 72–80.
[40] E.A. Wan, Neural network classification: a Bayesian interpretation, IEEE Trans. Neural Netw. 1 (4) (1990) 303–305.
[41] J.S. Breese, D. Heckerman, Decision-theoretic troubleshooting: A framework for repair and experiment, 38, (3) 1996, pp. 49–57.
[42] F.V. Jensen, et al., The SACSO methodology for troubleshooting complex systems, in: Ai Edam-Artificial Intelligence for Engineering Design Analysis and Manufacturing, 15, (4) 2001, pp. 321–333.
[43] C.F. Ruiz, et al., Efficiency measurement of research groups using data envelopment analysis and Bayesian networks, in: B. Larsen, J. Leta (Eds.) Proceedings of Issi 2009-12th International Conference of the International Society for Scientometrics and Informetrics, Vol. 1, 2009, pp. 296-300.
[44] C. Fabian Ruiz, et al., Efficiency measurement of research groups using Data Envelopment Analysis and Bayesian networks, Scientometrics 83 (3) (2010) 711–721.
[45] H. Kim, Y. Park, The impact of R & D collaboration on innovative performance in Korea: A Bayesian network approach, Scientometrics 75 (3) (2008) 535–554.
[46] A. Ibanez, P. Larranaga, C. Bielza, Using Bayesian networks to discover relationships between bibliometric indices. a case study of computer science and artificial intelligence journals, Scientometrics 89 (2) (2011) 523–551.
[47] M. Scutari, et al., Bayesian networks analysis of malocclusion data, Sci. Rep. (2017) 7.
[48] B.X.W. Liew, et al., Investigating the causal mechanisms of symptom recovery in chronic whiplash-associated disorders using Bayesian networks, Clin. J. Pain 35 (8) (2019) 647–655.
[49] D. Varshney, S. Kumar, V. Gupta, Predicting information diffusion probabilities in social networks: A Bayesian networks based approach, Knowl.-Based Syst. 133 (2017) 66–76.
[50] Francis, Galton, Regression towards mediocrity in hereditary stature, 1886.
[51] W.-Y. Zhang, et al., Measuring mixing patterns in complex networks by Spearman rank correlation coefficient, Physica A 451 (2016) 440–450.
[52] A.H. Pripp, Pearsons or Spearman's correlation coefficients, Tidsskr Nor Laegeforen 138 (8) (2018) 749.
[53] D. Valencia, R.E. Lillo, J. Romo, A Kendall correlation coefficient between functional data, Adv. Data Anal. Classif. 13 (4) (2019) 1083–1103.
[54] J. van Doorn, et al., Bayesian inference for Kendall's rank correlation coefficient, Am. Stat. 72 (4) (2018) 303–308.
[55] D.N. Reshef, et al., Detecting novel associations in large data sets, Science 334 (6062) (2011) 1518–1524.
[56] J.B. Kinney, G.S. Atwal, Equitability, mutual information, and the maximal information coefficient, Proc. Natl. Acad. Sci. USA 111 (9) (2014) 3354–3359.
[57] R.M. Gray, Entropy and information theory, 2000.
[58] X.-W. Chen, G. Anantha, X. Lin, Improving Bayesian network structure learning with mutual information-based node ordering in the K2 algorithm, IEEE Trans. Knowl. Data Eng. 20 (5) (2008) 628–640.
[59] S. Kiyomoto, T. Tanaka, K. Sakurai, K2: A stream cipher algorithm using dynamic feedback control, in: J. Hernando, E. FernandezMedina, M. Malek (Eds.) Secrypt 2007: Proceedings of the Second International Conference on Security and Cryptography, 2007. 204-+.
[60] I. Tsamardinos, L.E. Brown, C.F. Aliferis, The max-min hill-climbing Bayesian network structure learning algorithm, Mach. Learn. 65 (1) (2006) 31–78.
[61] H. Wang, D. Wang, S. Yang, A memetic algorithm with adaptive hill climbing strategy for dynamic optimization problems, Soft Comput. 13 (8–9) (2009) 763–780.
[62] C.T. Volinsky, A.E. Raftery, Bayesian information criterion for censored survival models, Biometrics 56 (1) (2000) 256–262.
[63] D. Posada, T.R. Buckley, Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests, Syst. Biol. 53 (5) (2004) 793–808.
[64] P.M.B. Vitanyi, M. Li, Minimum description length induction, Bayesianism, and Kolmogorov complexity, IEEE Trans. Inform. Theory 46 (2) (2000) 446–464.
[65] M.H. Hansen, B. Yu, Model selection and the principle of minimum description length, J. Amer. Statist. Assoc. 96 (454) (2001) 746–774.
[66] S.I. Vrieze, Model selection and psychological theory: A discussion of the differences between the akaike information criterion (AIC) and the Bayesian information criterion (BIC), Psychol. Methods 17 (2) (2012) 228–243.
[67] K. Yamaoka, T. Nakagawa, T. Uno, Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations, J. Pharmacokinet. Biopharm. 6 (2) (1978) 165–175.
[68] H. Muzhou, et al., A new hybrid constructive neural network method for impacting and its application on tungsten price prediction, Appl. Intell. 47 (1) (2017) 28–43.
[69] F. Weng, et al., Gold price forecasting research based on an improved online extreme learning machine algorithm, J. Ambient Intell. Humaniz. Comput. (2020).
[70] S. Srihari, et al., Combining evidence using likelihood ratios in writer verification, in: R. Zanibbi, B. Couasnon (Eds.), Document Recognition and Retrieval Xx, 2013.
[71] T. Hasan, J.H.L. Hansen, Maximum likelihood acoustic factor analysis models for robust speaker verification in noise, IEEE-Acm Trans. Audio Speech Language Process. 22 (2) (2014) 381–391.
[72] L. Wasserman, Bayesian model selection and model averaging, J. Math. Psych. 44 (1) (2000) 92–107.
[73] S. Cho, A causal relationship among nonverbal communication of dance instructor, customer participation, customer satisfaction and customer loyalty, Korean J. Sport Sci. 24 (1) (2013) 105–115.
[74] T.J. Lai, C.Y. Chen, IEEE, virtual community and customer participations in user centric internet service ventures, in: 2008 Portland International Conference on Management of Engineering & Technology, Vols. 1-5, IEEE, New York, 2008, pp. 1020–1027.
[75] H.C. Kelman, Interests, relationships, identities: Three central issues for individuals and groups in negotiating their social environment, Ann. Rev. Psychol. (2006) 1–26.
[76] Y. Wang, Q. Yu, D.R. Fesenmaier, Defining the virtual tourist community: Implications for tourism marketing, Tour. Manag. 23 (4) (2002) 407–417.
[77] J.N. Sheth, B.I. Newman, B.L. Gross, Why we buy what we buy: A theory of consumption values, J. Bus. Res. 22 (2) (1991) 159–170.
[78] M.B. Holbrook, Consumption experience, customer value, and subjective personal introspection: An illustrative photographic essay, J. Bus. Res. 59 (6) (2006) 714–725.
[79] X. Bao, F. Bouthillier, Information sharing: As a type of information behavior, Can. J. Inf. Libr. Sci.-Revue Canadienne Sci. Inf. Bibl. 30 (1–2) (2007) 91–92.
[80] Y.-F. Yang, H.-C. Yeh, W.-K. Wong, The influence of social interaction on meaning construction in a virtual community, Br. J. Educ. Technol. 41 (2) (2010) 287–306.
[81] O. Iglesias, S. Markovic, J. Rialp, How does sensory brand experience influence brand equity? Considering the roles of customer satisfaction, customer affective commitment, and employee empathy, J. Bus. Res. 96 (2019) 343–354.
[82] D. Paddeu, G. Fancello, P. Fadda, An experimental customer satisfaction index to evaluate the performance of city logistics services, Transport 32 (3) (2017) 262–271.
[83] B. Latané, The psychology of social impact, Am. Psychol. 36 (1981) 343–356.
[84] Z.J. Gong, W.J. Zhao, S.C. Xiong, Research on members' perceived value of knowledge sharing in virtual community based on grounded theory, Inf. Sci. (2014).
[85] H. Wilkins, Souvenirs: What and why we buy, J. Travel Res. 50 (3) (2011) 239–247.
[86] C. Achar, et al., What we feel and why we buy: the influence of emotions on consumer decision-making, Curr. Opinion Psychol. 10 (2016) 166–170.
[87] J. Camilleri, B. Neuhofer, Value co-creation and co-destruction in the Airbnb sharing economy, Int. J. Contemp. Hosp. Manag. 29 (9) (2017) 2322–2340.
[88] V. Ramaswamy, K. Ozcan, What is co-creation? An interactional creation framework and its implications for value creation, J. Bus. Res. 84 (2018) 196–205.
[89] B. Zhang, Co-creation:A new innovation mode, Sci. Technol. Progress Policy 27 (7) (2010) 22–24.
[90] A.C. Campos, et al., Co-creation of tourist experiences: a literature review, Curr. Issues Tourism 21 (4) (2018) 369–400.