IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 22, 2018, accepted May 11, 2018, date of publication May 17, 2018, date of current version June 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2837340

Structure Learning of Conditional Preference Networks Based on Dependent Degree of Attributes From Preference Database

ZHAOWEI LIU[1,2], ZHAOLIN ZHONG[2], KE LI[1], AND CHENGHUI ZHANG[1]
[1]School of Control Science and Engineering, Shandong University, Jinan 250061, China
[2]School of Computer Science and Control Engineering, Yantai University, Yantai 264005, China

Corresponding author: Chenghui Zhang (zchui@sdu.edu.cn)

This work was supported in part by the Natural Science Foundation of Shandong Province under Grant ZR2013FQ023 and Grant ZR2014FL009, in part by the Key Research and Development Project of Shandong Province under Grant 2015GSF115009, and in part by the National Natural Science Foundation of China under Grant 61403328 and Grant 61572419.

ABSTRACT Conditional preference networks (CP-nets) are used as intuitive graphical tools to represent conditional preference statements regarding the values of a set of attributes. Making use of CP-nets to solve certain learning problems has attracted growing attention from scholars and become a hot topic in the field of artificial intelligence. Therefore, many methods have been proposed to solve these learning problems. However, these approaches suffer from two main disadvantages: the amount of time required and their lack in concrete structures. To overcome these limitations, in this paper, we first provide theoretical support for the use of a conditional independent test for learning the structure of CP-nets. Second, we propose the dependent degree to calculate the dependency relationship among attributes. Finally, we present an algorithm to obtain the structures of CP-nets. Beyond that, a number of database samples have been reduced by filtering out insignificant or noise data, and a concrete structure of learned CP-nets with 18 attributes is given. The experiments show that our approach can obtain a better structure of CP-nets without materially increasing the time required for the process and put forward contrast to methods presented antecedently.

INDEX TERMS Conditional preference networks, structural learning, dependent degree, preference database.

I. INTRODUCTION

With the ever-increasing data in electronic commerce, the applications in recommender systems have become a hot topic of research [1]–[3]. In order to satisfy this demand, the artificial intelligence community is researching how to employ user preferences as a tool for customizing queries [4]. The ability to express preferences in a faithful but compact way, which can be handled efficiently, is essential in many reasoning tasks, including the above mentioned e-commerce, combinatorial optimization, multi-agent planning and agreement, and other scenarios where an agent need view and rank all the available choices [5]. Multi-attribute preference modeling and reasoning causes a combinatorial explosion, often leading to high computational cost [6]. The set of alternatives is often described as a product of multiple features, for example, a user's preferences over a set of cars, which can be described by their cost, colors, capacity, engine, cruise control system, sunroof, technical specifications, etc.

The tool we introduce in this paper is conditional preference networks (CP-nets) [7]–[9], which are typical qualitative formulations, to inform user preferences by means of the values of an attribute; these values are determined from the values of conditional or parent attributes.

The learning problem underlying the use of CP-nets is to extract a preference structure and multiple preference parameters by observing user queries [11]–[17]. Correctly extracting and collating these preferences is difficult in general, not just for acyclic CP-nets [18], [19], but also for binary CP-nets and separable CP-nets. The process involves employing methods for learning CP-nets, such as the hypothesis-testing learning solution [11], learning based on the induced graph of CP-nets [12], and other approaches that have been proposed [14], [20].

Compared to the approach of learning Bayesian Network with numerical probability [21]–[24], a qualitative method was proposed for mining conditional preferences expressed by a set of relational samples [25]. Our method allows us to construct CP-nets maintain scope and generality. The method receives a preference database as input (a set of pairs of tuples representing a sample of the user's past choices) and infers a preference network that will be used to infer new preferences about the values of new attributes.

The main contributions of our work are summarized as follows:

(1) We formalize the CP-nets learning problem based on preference database.

(2) Based on mutual information, we provide theoretical support for learning CP-nets using conditional independence tests.

(3) We propose the dependent degree among attributes, and present an algorithm to generate the structure from preference database based on it. Based on this approach, we can get a better level of performance without increasing the time required appreciably.

(4) We contribute to a concrete structure with 18 attributes that has seldom been studied in previous work.

The paper is organized as follows.

In Section II, we briefly discuss some related works concerning conditional preference networks. And in Section III we give the concept, examples of CP-nets and preference database. In Section IV, we present the algorithm to calculate the degree of dependence betweenattributes. In Section V, the algorithm is researched to construct a conditional preference network structure from a set of tuple pairs as input. Our experiments and discussions are reported in Section VI. Finally, we conclude the paper and explore future work in Section VII.

II. RELATED WORK

CP-nets have been proposed as a simple and intuitive graphical tool for representing conditional ceteris paribus preference statements over the values of a set of variables. Learning CP-nets can be viewed as learning a preference order from pairwise comparisons, because CP-nets represent a partial order over outcome space. This problem is also called learning to order things [26] or supervised ordering [27]. Under these circumstances, the learner is provided with a set of outcomes and a set of pairwise comparisons over these outcomes, and the task is to predict the preference order for a new set of outcomes. If the predicted preference order is required to be a total order, the problem is also called the ranking problem [28]. CP-nets are proposed to solve the preference problem between variables. In the following, we introduce related work in two aspects: mode learning and structure learning.

A. MODE LEARNING

CP-nets learning modes can be divided into active learning [15], [16] and passive learning [17]. The former is required to correctly and efficiently extract preference networks in binary-valued domains in [15] and [16]. When membership queries are available, they provide attribute-efficient learning algorithms for which the query complexity is linear in the size of the minimal CP-net that identifies the target concept, and logarithmic in the total number of attributes. In addition, Koriche and Zanuttini [15], [16] propose an active learning algorithm. Different from active learners, passive learners do not interact with users. In the [17] point of view: although CP-nets, a compact, intuitive language for representing preferences, alleviate the burden of preference elicitation, but they cannot eliminate the burden. Indeed, deriving a CP-net in complex domains is a tedious and time-consuming. Moreover, in some domains, such as auctions and automated negotiations, users may be unwilling to reveal their preferences. Such pairwise comparisons may be gathered, for instance, by passively observing the choices of a user on a web page. In the simplest case, we seek to derive a CP-net N such that P is a subset of the relation induced by N. The problem is intractable even in the case of input comparison with different results on the two variables, and the in-degree of the nodes of the simplest CP-net that implies the comparisons (if it exists) is upper-bounded by 2 [17]. The passive learning method is also proposed in [11] and [12]. It is first proved that the learning problem is intractable, even under several simplifying assumptions.

B. STRUCTURE LEARNING

In [29], the problem of learning separable ceteris paribus structures (SCP-structures), the simplified CP-nets, is addressed. In this structure, the set of variables for each variable is empty, that is, each variable is independent of all other variables. The algorithm for generating an acyclic CP-net entailing all training samples is proposed in [17], and the algorithm for learning CP-nets over a fixed acyclic digraph is proposed in [18]. We finally show that the proposed algorithm is a PAC-learner, i.e., the CP-networks it induces accurately predict the user's preferences on previously unseen situations [17]. Liu *et al.* [12] proposes the model of learning CP-nets from inconsistent examples because the problem of learning CP-nets from inconsistent examples is seldom addressed. A two-step method to solve this problem is proposed. The method can learn general CP-nets over multi-valued variables from inconsistent examples. At the same time, Liu *et al.* [11] considers the case of the randomness of the users' behaviors or the observation errors, there exists some noises (errors) in the training samples, introduce a new model of learning CP-nets from noisy samples. The algorithm for solving this problem is based on the chi-square testing. Two problems are addressed in this paper: (1) a new model of learning CP network from noisy samples is proposed, and (2) a new method is proposed to solve the new model in polynomial time.

III. PRELIMINARIES

A. CP-nets

Definition 1 (CP-nets): A Conditional Preference Network (CP-net) is a graph model $< G, T >$ that consists of:

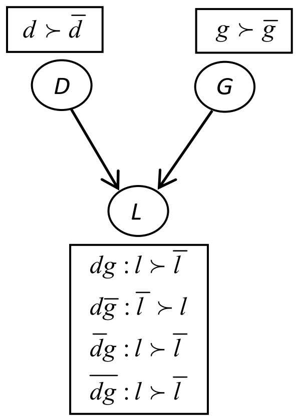


FIGURE 1. A CP-net for movie recommendations with attributes director (*D*), genre (*G*) and language (*L*).

–A graph G with vertices V and attributes $\{X1, X2, \ldots, Xn\}$; the edges of E stand for dependent relationships among attributes.

–Each vertex V is associated with a conditional preference table (CPT) T that qualifies the preference from the parent vertices to the children vertices. If a vertex has no incoming vertices, then its preference does not involve any other attributes from the graph G.

If the domain of an arbitrary attribute U is binary, like u and $\bar{u}$, it is called a binary-valued CP-net. Similarly, the domain of another attribute V is multiple, like v_i, v_j and v_k, and it is called a multi-valued CP-net. In this paper, we focus on binary-valued CP-net.

The CP-net may be acyclic or cyclic, distinguished structurally by the classic Bayesian network. Coincidentally, the term "CPT" is the same abbreviation as used to denote the conditional probabilistic table in Bayesian network.

Example 1 (Movie Recommendation): FIGURE 1 illustrates a CP-net that shows movie recommendations to a user named Alice. It consists of three attributes D, G and L, which stands for director, genre and language respectively. D, G and L are all binary values. In this example, d denotes Spielberg, and $\bar{d}$ denotes Lucas; g denotes action, and $\bar{g}$ denotes war; l denotes English, and $\bar{l}$ denotes foreign language. Alice unconditionally prefers Spielberg to Lucas and action to war. When Alice selects a language, her preference between English and a non-English (foreign) language is conditioned by the combination of director and genre: if Spielberg and action or Lucas and war is selected, she prefers an English language film. Otherwise, if Spielberg and war or Lucas and action are selected, she prefers a foreign language film.

What calls for special attention is that we can obtain the preferred outcome on the basis of the conditional attributes that it prefers or violates. The uvw outcome violates none of the preference constraints. The $dg\bar{l}$ outcome violates the conditional preference of L. The $d\overline{gl}$ outcome violates attributes G and L. The $\overline{dgl}$ outcome infringes all of the preference. Given this, we know that the semantics of CP-nets implies that violating one child attribute has a higher priority than violating the parent attribute. However, we cannot determine whether violating one or a few parent attributes is better than violating two or more child attributes. This dominance is an important property of CP-nets.

Definition 2 (Dominance of CP-Nets): Given a CP-net N and two outcomes o and o', we must decide whether $o \prec o'$ in N; that is, whether o' dominates$o$ in N.

When a binary-valued CP-net is direct-path singly connected (i.e., there is at most one direct path between any pair of nodes), then dominance testing is NP-complete [9], [10]. The problem remains difficult even the number of alternative paths between any pair of nodes in a CP-net is polynomially bounded.

Definition 3(Conditional Preference Independence): Let U, V and W be nonempty sets of P, where U is conditionally preferentially independent of V given an assignment w for W, for all $u_1, u_2 \in U$ and $v_1, v_2 \in V$. This satisfies:

$u_1 v_1 w \succ u_2 v_1 w$ if and only if $u_1 v_2 w \succ u_2 v_2 w$.

The relationship of independence statement $I(U, W, V)$—tested by performing a CI (Conditional Independence) test using a statistical hypothesis testing procedure.

B. PREFERENCE DATABASE

The main goal of conditional preference network learning is to provide a preference relationship set for a given preference dataset. The preference relationship of a finite set of objects follows a strict partial order. Typically, a strict partial order is represented by the symbol $\succ$. We denote by this formulation: $a_1 \succ a_2$, which meaning a_1 is preferred to a_2.

Definition 4 (Preference Database): A preference database is expressed with $\{I, < O_{tri}, O_{bri} >\}$. Let $I = R(1, 2, \ldots, n)$ be a preference schema with n attributes, and $Dom(R)$ be the set of all tuples over R. Let preference configuration outcomes be $I \subset Dom(R) \times Dom(R)$ and. The relation $< O_{tri}$, $O_{bri} >$ denotes the fact that the user prefers outcome O_{tri} to O_{bri}.

Example 2 (Preference Database): Let $R(A, B, C)$ be a preference schema with attribute domains given by $Dom(X) = \{x_0, x_1\}$ with $X = A, B$ and C, respectively.

In TABLE 1, the preference data is divided into the top part and the bottom part. The preference relation is fixed at $O_{tri} \succ O_{bri}$. When the instant is chosen as r_1, immediate benefits $O_{tr1} \succ O_{br1}$, that is, $a_1 b_1 c_1 \succ a_2 b_1 c_1$. Similarly, $a_1 b_1 c_1 \succ a_2 b_1 c_2$, $a_2 b_2 c_2 \succ a_2 b_1 c_2$ and $a_2 b_1 c_2 \succ a_2 b_1 c_1$.

To learn the CP-nets, we must overcome the following two problems: (1) generating the structure G *with* conditional independent testingthat reflects the dependencies among the attributes according to preference samples and (2) producing the conditional preference table T that reflects how some attributes influence preferences over the values of other attributes. The second problem can be naturally solved using Maximum Likelihood Estimation based on G derived from the first question; we omit this calculation due to space constraints in this paper.

The CP-nets learning problem is formalized as follows with TABLE 2:

TABLE 1. Preference database.

O	Attribute	Instant					
		r_1	r_2	r_3	r_4	...	r_n
O_{tri}	A	a_1	a_1	a_2	a_2	...	a_i
	B	b_1	b_1	b_2	b_1		b_i
	C	c_1	c_1	c_2	c_2		c_i
O_{bri}	A	a_2	a_2	a_2	a_2	...	a_j
	B	b_1	b_1	b_1	b_1		b_j
	C	c_1	c_2	c_2	c_1		c_j

TABLE 2. The CP-net structure learning problem.

The CP-net structure learning problem	
Input:	A relational schema $R(X_1,\ldots,X_n)$ of a training preference samples set *TRS* over *R*.
Output:	A conditional preference network (N).
Task:	Find a directed acyclic graph N such that $N \in \arg\max_{N} \sum_{i=1}^{n} similarity(Pare(X_i))$ Where $Pare(X_i)$ is the parent set of X_i.

Input: A relational schema $R(X_1, \ldots, X_n)$ is composed of a training preference samples set (*TRS*) over R and a testing preference sample (*TES*) of R. The relation Re_{uv} is provided by the user, which means that the user prefers outcome u to v.

Output: A conditional preference network (G, T) over the relational schema $R(X_1, \ldots, X_n)$ with a good similarity [11] with respect to R.

IV. DEPENDENT DEGREE OF ATTRIBUTE

In the dependence analysis approach, a CI test is typically used to check the validity of the conditional independence assertion $I(U, W, V)$ of any two given nodes U, V and a conditioning set W.

In the dependency analysis approach, we introduce mutual information to measure the dependence degree between the sets of variables U and V given W, which is simply the Kullback-Leibler divergence [30] between the joint distribution for U and V and the product of the corresponding marginal distribution. This value of mutual information is null when the two sets of variables are independent, and it becomes maximum when they are functionally dependent. Given a probability distribution p defined over two sets of variables U and V in the preference database, the mutual information between U and V is:

$$M(U, V|W) = \sum_{u,v} p(u, v|w) \log(\frac{p(u, v|w)}{p(u|w)p(v|w)}) \quad (1)$$

Theorem 1 (Kullback, 1968): Given a data set D with N attributes, if U and V are conditionally independent given W, the statistic $2NM_D(U, V|W)$ approximates to a distribution $\chi^2(l)$(Chi-square)[31] with $l = (c_U - 1)(c_V - 1)c_W$ degrees of freedom, where c_U, c_V and c_W represents the number of configurations for the sets of variables U, V and W, respectively.

If $W = \emptyset$, the result can be provided with the following corollary.

Corollary 1: Given a data set D with N attributes, if there exists an independent relationship between U and V, the statistic $2NM_D(U, V)$ approximates to a distribution $\chi^2(l)$ with $l = (c_U - 1)(c_V - 1)$ degrees of freedom.

We can then learn the CP-net model from the following theoretical judgement.

Corollary 2: The dependent relationship between U and V can be obtained from the preference database P in Definition 3.

To begin with, the conditional independence assertion [i.e., $I(U, W, V)$] is modeled as the null hypothesis. Suppose that we use the likelihood-ratio χ^2 test and the χ^2 statistic is calculated by:

$$\chi^2 = \sum \frac{(D-E)^2}{E} = \sum_{i=1}^{k} \frac{(D_i - E_i)^2}{E_i} \quad (2)$$

where,D_i is the current observed statistical data from database P, and E_i is expected to represent the statistics.

Suppose that the possible instantiations of the variables U, V, and W is u, v, and w respectively, and that they follow a distribution χ^2 with $(u-1)\times(v-1)\times w$ degree of freedom. The value is achieved by applying the computed χ^2 against the distribution, which shows the least level of significance for which the given data leads to the rejection of the null hypothesis. If the value is less than a predefined cutoff value α, then the test shows strong evidence to reject the hypothesis; otherwise, the hypothesis cannot be rejected.

In general, when calculating χ^2, we take all samples into consideration. However, some information is useless, and such data would be filtered out as one of the first steps in our work. To explain further, we provide some concepts [32]. And those notations are all summarized in TABLE 3.

(1) For each $(u, u') \in dom(U) \times dom(U)$, $u \neq u'$, and $O_{uu'}$ is described as the subset of pairs $(o, o') \in P$, such that $o[U] = u \wedge o'[U] = u'$ or $o[U] = u' \wedge o'[U] = u$;

(2)We present unconditional evidence as $E((u, u'), P) = \frac{|O_{u,u'}|}{|P|}$;

(3) For each $v \in dom(V)$, $E_{v|(u,u')}$ is described as the subset of $O_{uu'}$ including the pairs (o, o'), such that $o[V] = o'[V] = v$;

(4) We present conditional evidence as $E((v|(u, u'), P) = \frac{|E_{v|(u,u')}|}{|\bigcup_{v' \in dom(v)} E_{v'|(u,u')}|}$

With the above-mentioned definitions, the following algorithm 1 is proposed in TABLE 4.

After running Algorithm 1, we achieve a new database P' which filters out the noise data and then construct the CP-net graph mode l based on it. In this algorithm, $\alpha 1$ and $\alpha 2$ are flexible for different applications and purposes.

TABLE 3. Notations and definitions.

Notations	Description
CP-net	Conditional Preference Network
G	The directed graph
V	A vertex set in the directed graph
T	Conditional preference table(*CPT*)
o, o'	Two outcomes in a CP-net
$\succ$	A strict partial order
$R(1, 2, \ldots, n)$	A preference schema with n attributes
$Dom(X)$	A set of domain values regarding attribute X
$Pare(X)$	A set of direct parent nodes of attribute X
$I(U,W,V)$	The conditional independence assertion of any two given nodes U, V and a conditioning set W.
P	A preference database
α_1 , α_2	The thresholds in the preference database filtering process
$E((u,u'),P)$	Unconditional evidence
$E((v\,\vert\,(u,u'),P)$	Conditional evidence
$E_{v\vert(u,u')}$	the subset of $O_{uu'}$ including the pairs (o,o')
$O_{uu'}$	the subset of pairs $(o,o')\in P$
$Degree(U,V\,\vert\,P')$	The degree of Dependence of (U, V) with respect to P'
Similarity	The similarity of edge sets

TABLE 4. The algorithmforpre-filter of the preference database.

Algorithm 1. The filter of the preference database P

Input: a preference database P; the thresholds α_1 and α_2 .

Output: a new preference database P' by filter.

For each pair $(u,u')\in dom(U)\times dom(U)$, $u\neq u'$ do
  If $E((u,u'),P)<\alpha_1$
    Delete (u, u')
  Else compute $E((v\,\vert\,(u,u'),P)$
    If $E((v\,\vert\,(u,u'),P)<\alpha_2$
  Delete $v\vert$ (u, u')
Return P'

Algorithm 2 works on the obtained database in the TABLE 5.

V. CP-nets LEARNING

The dependent relation among attributes in Section 4 can be regarded as an analysis-based or constraint-based learningapproach. These constraints are usually conditionally independent statements. The constraint-based approach finds a network that can explain the dependent and independent relationships in the data. It works on the assumption that Conditional Preference networks depict conditional dependence relationships among the variables. Hence, this approach attempts to construct a conditional preference network using dependency information obtained from the data. Typically,

TABLE 5. The algorithm for degree of dependence of pairwise attributes.

Algorithm 2. The degree of dependence of pairwise attributes

Input: P' : a preference database; (U, V): a pairwise attributes

Output: The degree of Dependence of (U, V) with respect to P'

For each pair $(v,v')\in dom(V)\times dom(V)$, $v\neq v'$ and (v,v') is comparable do
For each $u\in dom(U)$, u is a cause for (v,v') being comparable do
Let $f_1(E_u\vert_{(v,v')})=\max\{N,1-N\}$, where $N=$ $\dfrac{\left\vert\{(o,o')\in E_u\vert_{(v,v')}: o>o'\wedge(o[V]=v\wedge o'[V]=v')\}\right\vert}{\left\vert E_u\vert_{(v,v')}\right\vert}$
Let $f_2(O_{(v,v')})=\max\{f_1(E_u\vert_{(v,v')}):u\in dom(U)\}$
Let $Degree(U,V\,\vert\,P')=\max\{f_2(O_{(v,v')}):(v,v')\in dom(V)\times dom(V)\}$, $v\neq v'$, (v,v') is comparable
Return $Degree(U,V\,\vert\,P')$

TABLE 6. The algorithm for output structure of CP-net.

Algorithm 3. The Output structure of CP-net

Input: the filtered preference database P'; cutoff α of χ^2

Output: G of learned CP-net with respect to P'

For each pair of attributes (A_i, A_j) . **do**
  Calculate the degree of attribute dependence between the pair (A_i, A_j) .
  Let T be the outcome set of these calculations, the form of expression is $(A_i \xrightarrow{degree} A_j)$.
  Select from T elements whose p>2.71, where df=1, α =0.10
  Order the element $(A_i \xrightarrow{degree} A_j)$ in T in decreasing order on the basis of their p.
  Start the graph G of the CP-net with a node for each attribute.
  For each element $(A_i \xrightarrow{degree} A_j)$ ∈ order set T **do**
    When the insertion edge (A_i, A_j) does not form a loop, it can be inserted into the graph G.
Return G.

the existence of a perfect graph assumes that there exists a conditional preference network that expresses all of the conditional dependence relationships induced by database P. Consequently, this approach constructs a network G by testing the conditionally preferential independence expressed by statement $I(U, W, V)$. It can access the evidence from P, and it follows that U should be dependent on V by W in G; otherwise, U is independent of V by W.

Although the independent basis constraints can be utilized in the structure, in some cases the latent variables exist. However, we are concerned only with no latent variables or missing values. The constraint-based algorithm with filtering at first in the above mentioned has certain robustness which means that the result of the output fluctuates lowly.

In the algorithm 3 of TABLE 6, df is the degree of freedom. The formula is $df = \vert dom(U)\vert - 1$. The level of significance is the probability to estimate that the overall parameters fall within a certain interval may be wrong, expressed in α. Chi-square test uses a small probability principle, which allows the small probability of the standard, known as the level of significance. α is not a fixed number, i.e., the greater α, the more likely that the original hypothesis is rejected. The significance level depends on the nature of the study and the requirements for the accuracy of the conclusions. In this paper, the α value selected is 0.10.

Theorem 2 (Correctness of Algorithm): Based on the principle of algorithm learning to get the CP-nets structure must be acyclic.

Proof: In algorithm 3, the corresponding CP-nets structure is based on the principle of algorithm and topological ordering. That is to say, the parent of each vertex is selected before the dependence vertex. For example, Let $V = \{X1, X2, \ldots, Xn\}$, if we select X_1 as the current vertex, the corresponding candidate parent setscomposing of any subset of $Pare(X1) = \{X2, X3, \ldots, Xn\}$. Immediately following, we select one of the biggest weight value as the parent of X_1, such as $\{X2, X3\}$, then $\{X1, X2, X3\}$ becomes the current vertex, and the corresponding parent set is selected only by any subset of $Pare(\{X2, X3\}) = \{X4, X5, \ldots, Xn\}$, such as $\{X4, X7, X8\}$. We keepthe above-mentioned steps to learn until all vertices are identified. In the learning process, the vertices have been visited will not be searched again. Thus, the learning method does not appear to be the case that the vertices are mutually parent to each other. In summary, we can prove the correctness of algorithm 3, that is to say, the CP-nets structure obtained from the algorithm principle is acyclic.

Theorem 3 (Completeness of Algorithm): Based on algorithm 3, the learning method of obtaining the acyclic CP-nets structure is a local optimal solution.

Proof: According to algorithm 3, when executing algorithmat each step of the vertex parent selection process, the biggest weight value is selected as the parent of the current vertex. This operation is performed continuously until the parent set of all vertices is determined. Because of the learning process of acyclic CP-nets, the maximum value is determined for each step of the parent's determination. In other words, the current optimal value is the parent of the current vertex. Therefore, the calculation method based on the algorithm can obtain a local optimal solution.

Theorem 4 (Complexity of Algorithm): The computing complexity of algorithms is $O(n^2 \cdot m)$.

Proof: We analyze the complexity of algorithms as explained below.

(i) Scan a bituple to save the D (set of attributes that have different values in the two tuples) and the I (set of attributes that have equal values in the two tuples) are $O(n)$, where n is the number of attributes;

(ii) The computation of the degree of dependence between the pairs of attributes is $O(n^2)$;

(iii) The computation of the topology is $O(n^2 \cdot m)$, where m is the number of edges of the CP-nets. Therefore, the total complexity for building the model in the algorithm is $O(n^2 \cdot m)$.

However, over time, the algorithm will consider only a few edges in its processing, differently from the algorithm which always considers all possible edges. Thus, in the practice, as the more attributes the preference data has, the greater the advantage of the algorithms.

VI. EXPERIMENTS AND DISCUSSION

We found that CP-nets learning requires consistent training samples. In order to verify our model and algorithm's correctness and accuracy, we set out to put the algorithm into practice on real data. In experiment, real users' preferences collected by GroupLens Research are used to generate the learning samples.

A. EXPERIMENTAL RESULTS ON THE MOVIELENS DATASET

In order to evaluate our approach over real-world datasets, we consider data containing preferences related to movies collected by GroupLens Research [33] from the MovieLens [34] web site. These movies include all 18 types of movie genre: Action, Adventure, Animation, Children's, Comedy, Crime, Documentary, Drama, Fantasy, Film-Noir, Horror, Musical, Mystery, Romance, Sci-Fi, Thriller, War and Western. Each genre of movie can be considered as an attribute in the CP-net.

The following FIGURE 2 and FIGURE 3 show the structured learning result of random user U_1 and U_2 by using a Chi-square test approach compared to the our approach proposed in our work. In the both FIGURE 2 and FIGURE 3, ①, ②, ③ and etc. denote one attribute of a movie genre.

1) THE PERFORMANCE OF OUR METHOD

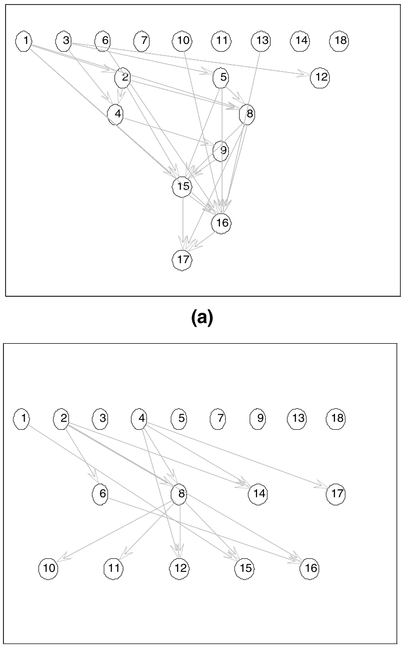
We find the consistency law of the certain user's preference [36]. Therefore, we classify the data of a certain user, instead of a few users or all users, into a training set and a test set. To prove the performance of the proposed learning method, we will compare the acyclic CP-nets generated by the random selection of the user's training data in the first part to the acyclic CP-nets generated by the latter part of the test data.

In this paper, we, focusing solely on the structure learning of CP-nets, choose similarity as the evaluating parameter of learning results instead of agreement [11], [12].

Because the CP-nets structure is a graph, it is difficult to judge the similarity between graphs. Therefore, the similarity of structure can be transformed into the similarity of edge sets. The similarity of edge sets is used to measure the similarity of structure. The higher the similarity is, the better level of the performance of the algorithm do. The similarity formula of edge set is defined as follows:

$$similarity = \frac{num(agree\_edge)}{num(total\_edge)} \tag{3}$$

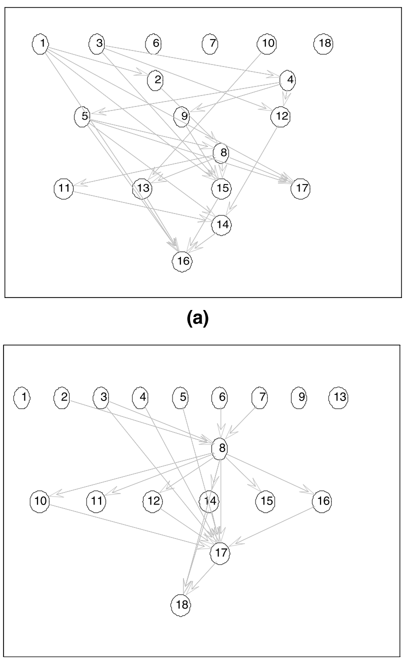
We also evaluate the performance of our approach through comparing with Chi-square approach. The comparison results



(a)

(b)

FIGURE 2. Comparison of Chi-square test approachin figure (a) toour approachin figure (b) of User U_1 (a) A CP-net with Chi-square test approach. (b) A CP-net with our approach.

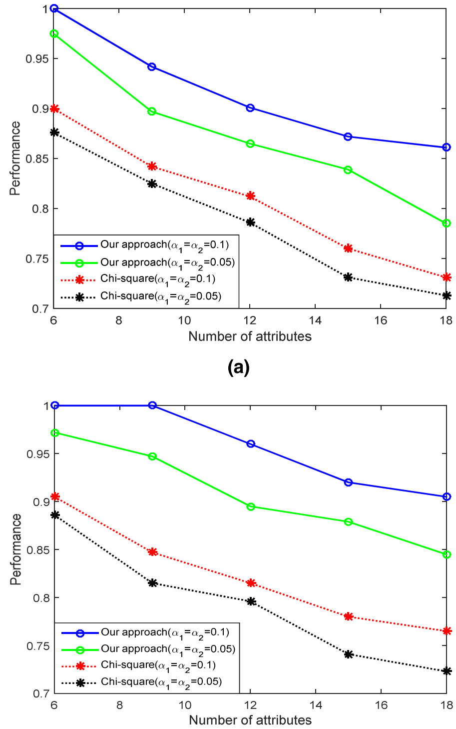


(a)

(b)

FIGURE 3. Comparison of Chi-square test approachin figure (a) toour approachin figure (b) of User U_2. (a) A CP-net with Chi-square test approach. (b) A CP-net with our approach.

are shown in FIGURE 4. The numbers in the following diagram on the horizontal axis represent the number of movie attributes. For example, the 6 numbers on the horizontal axis



(a)

(b)

FIGURE 4. Performance of user U_1 in figure (a) and U_2 in figure (b) by comparing the Chi-square test approach to our approach.

denotes the six attributes randomly selected from the total of 18 attributes, similarly, the number 12 on the horizontal axis denotes twelve attributes randomly taken out of the total of 18 attributes, and so on. FIGURE 4 illustrates the similarity contrast betweenour approach and the Chi-square algorithms evolved through the number of attributes derived from 200 pairwise datasets taken from the preference database. Our approach demonstrates more significant performance than the Chi-square algorithms.

2) THE RUNNING TIME OF OUR METHOD

Finally, we analyze the time required to execute our approach. FIGURE 5 shows the time measured in seconds required by our approach (versus the Chi-square algorithms) to generate the model at each refresh point. Notice that the time required by our approach to refresh is very low—a matter of milliseconds. Therefore, FIGURE 5 shows that the time taken by our approach to generate the results is little difference with the Chi-square algorithms.

3) COMPARISON TO PREVIOUS WORK

Ultimately, we compare our work in this paper with other studies. The method proposed by Dimopoulos *et al.* [17] can't deal with noise data. In other words, if the experimental data are obtained in the course of observation, the Dimopoulo's method will lose its function, and it can't learn the corresponding CP-nets structure.

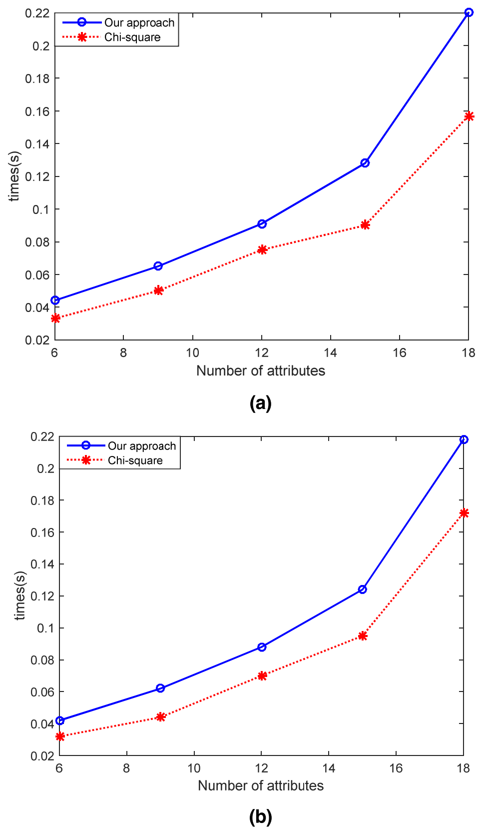


FIGURE 5. Run time of user U_1 in figure (a) and U_2 in figure (b) by comparing the Chi-square test approach to our approach.

TABLE 7. Comparison to previous work.

Method of different reference	Noise data	Inconsistent data	Number of candidate parents	Number of exact parents	Obtain CP-nets
The method of [17]	No	No	1	1	No
The method of [11]	Yes	No	1	1	No
The method of [12]	Yes	Yes	1	1	No
Our method	Yes	Yes	n-1	<n-1	Yes

Liu *et al.* [11], [12] propose the two method considering the X_i's candidate parent set $Pare(X_i)$ in the case of only one. That is only related to V-X_i but not all of other attributes are considered. At the same time, the Chi-square test method cannot ensure the learning of the CP-nets structure of the acyclicity. However, Liu's work focuses on noise data and inconsistent data for the first time, and proposes a specific algorithm.

As shown in the TABLE 7, the comparison between our calculation method and other learning methods is given, which can deal with noise data, inconsistent data. The approach we proposed also increase the number of candidate parent sets and the number of parents.

VII. CONCLUSIONS AND FUTURE WORK

With great attentions paid on artificial intelligence, research of preference learning highlights value in practice. As a compact and powerful model, CP-nets reduce significantly the difficulty and complexity of preference learning. In this paper, we present the dependent degree of attributes to learn structure of CP-nets from preference database even including noise data. We propose theoretical support to a conditional independent test for learning CP-nets. We also show concrete structure of 18 attributes in real-world dataset.

In future work, we will pay more attention to learn from large-scale preference database with convolutional neural networks [35], dynamic data streams [36] and PCP-nets [37]. Such problems would bring us closer to working with applications encountered in real-time e-commerce and multi-sourced sensor data.

ACKNOWLEDGMENT

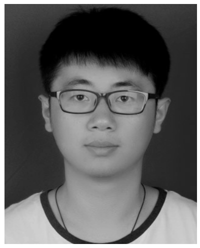
The authors would like to thank the anonymous reviewers for their constructive comments. They would also like to thank for the SMILE team of Yantai University for their helpful discussions.

REFERENCES

[1] M. Deshpande and G. Karypis, "Item-based top-n recommendation algorithms," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 143–177, 2004.
[2] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.
[3] R. Burke, "Hybrid recommender systems: Survey and experiments," *User Model. User-Adapted Interact.*, vol. 12, no. 4, pp. 331–370, 2002.
[4] J. Chomicki, "Preference formulas in relational queries," *ACM Trans. Database Syst.*, vol. 28, no. 4, pp. 427–466, 2003.
[5] Z.-P. Fan, J. Ma, and Q. Zhang, "An approach to multiple attribute decision making based on fuzzy preference information on alternatives," *Fuzzy Sets Syst.*, vol. 131, no. 1, pp. 101–106, 2002.
[6] Z. Xu, "Intuitionistic preference relations and their application in group decision making," *Inf. Sci.*, vol. 177, no. 11, pp. 2363–2379, 2007.
[7] C. Boutilier, R. I. Brafman, C. Domshlak, H. H. Hoos, and D. Poole, "CP-nets: A tool for representing and reasoning with conditional *ceteris paribus* preference statements," *J. Artif. Intell. Res.*, vol. 21, no. 1, pp. 135–191, 2004.
[8] J. Liu, "Research on CP-nets and its expressive power," *Acta Autom. Sin.*, vol. 37, no. 3, pp. 290–302, 2011.
[9] J. Goldsmith, J. Lang, M. Truszczynski, and N. Wilson, "The computational complexity of dominance and consistency in CP-nets," *J. Artif. Intell. Res.*, vol. 33, no. 1, pp. 403–432, 2014.
[10] O. Nez and A. Jonsson, "The complexity of planning problems with simple causal graphs," *J. Artif. Intell. Res.*, vol. 31, no. 1, pp. 319–351, 2011.
[11] J. Liu, Z. Yao, Y. Xiong, W. Liu, and C. Wu, "Learning conditional preference network from noisy samples using hypothesis testing," *Knowl.-Based Syst.*, vol. 40, no. 1, pp. 7–16, 2013.
[12] J. Liu, Y. Xiong, C. Wu, Z. Yao, and W. Liu, "Learning conditional preference networks from inconsistent examples," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 2, pp. 376–390, Feb. 2014.
[13] J. Fürnkranz and E. Hüllermeier, "Preference learning and ranking by pairwise comparison," in *Preference Learning*. Berlin, Germany: Springer, 2010, pp. 65–82.
[14] J. T. Guerin, T. E. Allen, and J. Goldsmith, "Learning CP-net preferences online from user queries," in *Proc. Int. Conf. Algorithmic Decis. Theory*, 2013, pp. 208–220.

[15] F. Koriche and B. Zanuttini, "Learning conditional preference networks with queries," in *Proc. Int. Joint Conf. Artif. Intell.*, 2009, pp. 1930–1935.
[16] F. Koriche and B. Zanuttini, "Learning conditional preference networks," *Artif. Intell.*, vol. 174, no. 11, pp. 685–703, 2010.
[17] Y. Dimopoulos, L. Michael, and F. Athienitou, "Ceteris paribus preference elicitation with predictive guarantees," in *Proc. Int. Joint Conf. Artif. Intell.*, 2009, pp. 1890–1895.
[18] E. Alanazi, M. Mouhoub, and S. Zilles, "The complexity of learning acyclic CP-nets," in *Proc. Int. Joint Conf. Artif. Intell.*, 2016, pp. 1361–1367.
[19] C. Cornelio, U. Grandi, J. Goldsmith, N. Mattei, F. Rossi, and K. B. Venable, "Voting with CP-nets using a probabilistic preference structure," in *Proc. 5th Int. Workshop Comput. Soc. Choice*, 2014, pp. 12–26.
[20] O. Schulte, Z. Qian, A. E. Kirkpatrick, X. Yin, and Y. Sun, "Fast learning of relational dependency networks," *Mach. Learn.*, vol. 103, no. 3, pp. 377–406, 2016.
[21] C. Contaldi, F. Vafaee, and P. C. Nelson, "Bayesian network hybrid learning using an elite-guided genetic algorithm," *Artif. Intell. Rev.*, vol. 293, no. 1, pp. 1–28, 2018.
[22] I. Tsamardinos, L. E. Brown, and C. F. Aliferis, "The max-min hill-climbing Bayesian network structure learning algorithm," *Mach. Learn.*, vol. 65, no. 1, pp. 31–78, 2006.
[23] M. L. Wong, W. Lam, and K. S. Leung, "Using evolutionary programming and minimum description length principle for data mining of Bayesian networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 2, pp. 174–178, Feb. 1999.
[24] G. F. Cooper and E. Herskovits, "A Bayesian method for the induction of probabilistic networks from data," *Mach. Learn.*, vol. 9, no. 4, pp. 309–347, 1992.
[25] E. Alanazi and M. Mouhoub, "A framework to manage conditional constraints and qualitative preferences," in *Proc. Florida AI Res. Soc. Conf.*, 2013, pp. 176–179.
[26] W. W. Cohen, R. E. Schapire, and Y. Singer, "Learning to order things," *J. Artif. Intell. Res.*, vol. 10, no. 1, pp. 243–270, 1999.
[27] T. Kamishima, H. Kazawa, and S. Akaho, "Supervised ordering—An empirical survey," in *Proc. IEEE Int. Conf. Data Mining*, Nov. 2005, pp. 673–676.
[28] E. Fernandez, N. Cancela, and R. Olmedo, "Deriving a final ranking from fuzzy preferences: An approach compatible with the principle of correspondence," *Math. Comput. Model.*, vol. 47, nos. 1–2, pp. 218–234, 2008.
[29] J. Lang and J. Mengin, "The complexity of learning separable *ceteris paribus* preferences," in *Proc. Int. Joint Conf. Artif. Intell.*, 2009, pp. 848–853.
[30] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951.
[31] K. Pearson, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *Philos. Mag. J. Sci.*, vol. 50, no. 302, pp. 157–175, 1900.
[32] S. de Amo, M. S. Diallo, C. T. Diop, A. Giacometti, D. Li, and A. Soulet, "Contextual preference mining for user profile construction," *Inf. Syst.*, vol. 49, pp. 182–199, Apr. 2015.
[33] *MovieLens 100K Dataset*. Accessed: Sep. 20, 2017. [Online]. Available: https://grouplens.org/datasets/movielens/
[34] *Recommendations*. Accessed: Sep. 20, 2017. [Online]. Available: https://movielens.org/
[35] W. Zhang, Q. Chen, W. Zhang, and X. He, "Long-range terrain perception using convolutional neural networks," *Neurocomputing*, vol. 275, pp. 781–787, Jan. 2018.
[36] J. Lerner, N. Indlekofer, B. Nick, and U. Brandes, "Conditional independence in dynamic networks," *J. Math. Psychol.*, vol. 57, no. 6, pp. 275–283, 2013.
[37] C. Cornelio, U. Grandi, J. Goldsmith, N. Mattei, F. Rossi, and K. B. Venable, "Reasoning with PCP-nets in a multi-agent context," in *Proc. Int. Conf. Auton. Agents Multiagent Syst.*, 2015, pp. 969–977.

ZHAOWEI LIU was born in Yantai, China, in 1979. He received the B.S. and M.S. degrees from Shandong University, China, in 2002 and 2005, respectively, where he is currently pursuing the Ph.D. degree in control science and engineering. He is currently an Associate Professor with the School of Computer Science and Control Engineering, Yantai University. His research interests include intelligence control, multi-agents systems, and machine learning.



ZHAOLIN ZHONG was born in Qingdao, China, in 1993. He received the B.S. degree in automation from Yantai University, in 2016, where he is currently pursuing the M.S. degree in computer science. His main research interests include the learning of CP-nets in the field of artificial intelligence.

KE LI was born in Jinan, China, in 1979. He received the B.S. degree from Shandong Jianzhu University, Jinan, in 2002, and the Ph.D. degree (successive postgraduate and doctoral program) from Shandong University, Jinan, in 2007. He is currently an Associate Professor with the School of Control Science and Engineering and the Engineering Research Center of Power Electronics Energy-Saving Technology & Equipment, Ministry of Education, Shandong University. His research interests include renewable energy, energy storage technology, power electronics, and motion control.

CHENGHUI ZHANG was born in Dezhou, China, in 1963. He received the B.S. and M.S. degrees from the Shandong University of Technology, China, in 1985 and 1988, respectively, and the Ph.D. degree from Shandong University, China, in 2001. In 1988, he joined Shandong University, where he is currently a Full Professor with the School of Control Science and Engineering, and the Director of Research Center of Power Electronics Energy-Saving Technology & Equipment, Chinese Education Ministry. His current research interests include optimal control of engineering, electric vehicles, power electronics and motor drives, and energy-saving techniques. He was selected as a Changjiang Scholar of the Education Ministry and a Taishan Scholar of Shandong Province in 2009.