Temporal Functional Dependencies and Temporal Nodes Bayesian Networks

WEI YI LIU¹, NING SONG² and HONG YAO³

¹School of Information Science and Engineering, Yunnan University, Kunming 650091, PR China
²Department of Metallurgy, Kunming University of Science and Technology, Kunming 650093, PR China
³Department of Computer Science, University of Regina, SK, Canada, S4S 0A2
Email: liuweiyi2000@yahoo.com.cn, songning@public.km.yn.cn, yao2hong@cs.uregina.ca

A temporal nodes Bayesian network (TNBN) is an alternative graphical representation of temporal probabilistic information. The temporal functional dependency is an important data dependency in temporal relational databases. In this paper temporal data dependencies are extended into delayed conditions, called asynchronous dependencies. The purpose of this paper is to set up a connection between these asynchronous data dependencies and TNBNs. We show that we can obtain the maximum conditional independencies implied by given temporal functional dependencies. Furthermore, we propose an approach to construct a TNBN from a set of temporal functional dependencies.

Received 4 September 2003; revised 29 July 2004

1. INTRODUCTION

Bayesian networks (BNs) have been successfully established as a framework for managing uncertainty using probability [1]. They have been applied in practice to a wide variety of applications involving uncertainty such as knowledge representation, forecasting and reasoning [1, 2, 3, 4]. In order to represent temporal probabilistic information in some real world applications such as medical and industrial diagnosis dynamic Bayesian networks (DBNs) [5, 6, 7, 8] are used to represent system states using ‘time slices’. A DBN is graphically represented by a directed acyclic graph (DAG) together with a set of conditional probability tables (CPTs) for each node. However, the structure of DBNs becomes too complex in some cases when only few system states change in a temporal range. Arroyo-Figueroa et al. [9, 10] present an alternative representation called temporal nodes Bayesian network (TNBN) to simplify the representation of DBN. Each node of TNBN represents a temporal attribute, and the arc between nodes corresponds to a causal-temporal relation.

In this paper we have proposed an approach to construct TNBNs by exploiting temporal data dependencies used in temporal databases [11, 12, 13]. The proposed approach is novel in that it can obtain the structure of the TNBN from a set of temporal data dependencies without extensive computational cost incurred by learning TNBN directly from data, because temporal data dependencies can be obtained directly from real-world semantics instead of learning from a large amount of data.

In order to exploit temporal data dependencies to build a TNBN first the theoretical relationship between BNs and temporal relational databases is analyzed. Three temporal data dependencies: probabilistic temporal functional dependency (PTFD), probabilistic temporal multivalued dependency (PTMD) and probabilistic temporal acyclic join dependency (PTAJD) are discussed in this paper by extending the probabilistic relational model [14, 15, 16] into a probabilistic temporal relational model. Next, the desired DAG of the TNBN is constructed by factorizing the joint probabilistic distribution following the order of variables which is obtained from given temporal data dependencies. Finally the CPT for each node of a TNBN is obtained from temporal relational databases.

Our contributions to this paper are: first we extend temporal data dependencies to consider delayed conditions, which are called asynchronous dependencies, and present a probabilistic temporal relational model. Second a joint probabilistic distribution can be factorized using obtained temporal data dependencies, the tentative structure of a TNBN can be constructed according to factorization. Third a method is proposed to collect the maximum conditional independencies information implied by the given set of temporal functional dependencies (TFDs). The collected information can be used to simplify the structure of TNBNs. Next we show that simplified TNBN is suited to data since it is a maximal preserved PTAJD of a given set of PTMDs. A method to find a maximal preserved PTAJD of a given set of PTMDs is also proposed. Finally, though our attention is focused on the structure of TNBNs, we also give a
method to learn CPTs of TNBNs from temporal relational databases.

This paper is organized as follows. The related work is described in Section 2. In Section 3 we briefly review TNBN. The probabilistic temporal relational model is introduced in Section 4. In Section 5 we discuss the temporal data dependencies with delayed time on the probabilistic temporal relational model. In Section 6 an approach to construct a TNBN from temporal data dependencies is presented. A detailed example to illustrate the proposed approach is given in Section 7. Conclusions are given in Section 8.

2. RELATED WORK

BNs have been successfully used in real world applications [1, 2, 3, 4], however, in order to further deal with temporal variables, DBNs are required. A lot of researchers have proposed many methods for constructing DBNs [5, 6, 7, 8]. Friedman et al. propose the structural expectation maximization (SEM) algorithm to learn structures of BN from incomplete data by extending the standard expectation maximization (EM) algorithm [6, 7]. They adopt a stochastic search method to avoid local maximums. Wang et al. [8] uses evolutionary algorithms for learning DBNs.

Although the DBN is an alternative representation for temporal probabilistic causality, it becomes a complex representation in some cases when only few system states change in a temporal range. Thus Arroyo-Figueroa et al. [9, 10] present TNBN.

It is a quite natural way to extend the methods for constructing BNs to construct TNBNs. There are two major approaches to construct BNs [17, 18, 19, 20, 21, 22, 23, 24]. The first one is a knowledge-based approach. It uses first-order probability logic to represent the structure of a BN since first-order probability logic is a desirable way to represent probabilistic information [17, 19, 21, 23]. The drawback of this approach is that it focuses on the formal reasoning in the knowledge base but does not provide practical algorithms to construct BNs. The second is a statistics approach. Heckerman et al. [20] presented an algorithm for learning BNs based on the statistical retrieved method. However, the statistics approach is time consuming.

A recent approach for constructing BNs is to associate data dependencies in databases with conditional independencies in probabilistic networks. Wong [15] provided a probabilistic relational model to represent conditional probability. Butz et al. [14, 16] presented three kinds of probabilistic data dependencies (probabilistic functional dependency, probabilistic multivalued dependency and probabilistic acyclic join dependency (PAJD)). They indicated that a probabilistic functional dependency is a sufficient condition of conditional independency, a probabilistic multivalued dependency is a necessary condition of conditional independency and a probabilistic acyclic join dependency can be viewed as a BN factorization. To apply this approach to construct a TNBN the data dependencies should be further extended to temporal data dependencies.

3. TNBN MODEL

In this section the TNBN is reviewed.

DEFINITION 3.1. [9, 10]. A TNBN consists of the following: a set of temporal nodes and a set of directed edges between temporal nodes. Temporal nodes represent temporal attributes. Each value of temporal attribute A is called a temporal atom, which is an ordered pair (σ, µ), where σ is a value from the domain of A and µ is the time interval associated with σ. Temporal nodes together with directed edges form a DAG. To each temporal node A with parents B1, . . . , Bn there is attached a conditional probability table. If A has no parents then the table reduces to unconditional probabilities P(A · µk) (k = 1, 2, . . . ), where (A · µk) denotes the value of A at time interval µk.

The following example of an accident defines a TNBN.

EXAMPLE 1. In a car accident there are three non-temporal variables: collision C, head injury H, and internal bleeding I; and two temporal variables: pupils dilated P and unstable sign V. P has an initial state (0, 5) and two temporal intervals (0, 3) and (3, 5), which represent that a head injury occurs, the swelling causes the pupils to dilate within 3 min if patient is checked or 5 min without checking. V has an initial state (0, 60) and three different time intervals (0, 10), (10, 30) and (30, 60), where (0, 10) represents that a head injury tends to destabilize vital signs within 0 to 10 min; (10, 30) represents that the bleeding is gross and it will take 10–30 min to destabilize vital signs; (30, 60) represents that the bleeding is slight and it will take 30–60 min to destabilize vital signs.

The corresponding TNBN is shown in Figure 1.

FIGURE 1. A TNBN of accident example.

Jensen et al. [11], Wijsen [12] and Wang et al. [13] proposed the concepts of the TFD and the temporal multi-valued dependency (TMD) for temporal variables within the context of identical time interval, called synchronous data dependencies. In some real world applications temporal data dependencies often involve delayed concepts, called asynchronous dependencies. Thus it is important that synchronous data dependencies should be extended into asynchronous data dependencies.
A temporal relational model is defined as follows:

**Definition 4.1.** [25]. \( R = (A_1, \ldots, A_n) \) is a temporal relational scheme with \( n \) temporal attributes \( A_i \). A relation of \( R \) is a set of tuples, \( \langle a_{i1}, \ldots, a_{in} \rangle \), where \( a_{ij} \) is a temporal atom or a set of temporal atoms.

Figure 2 shows a relation over temporal relational scheme \( R = (K, A, B, C, D) \).

The temporal relational model is extended to a probabilistic temporal relational model (PTRM) by adding a probabilistic attribute \( p \) to the temporal relational model. More precisely let \( R(U) \) be a temporal relational scheme and \( p \) be the probabilistic attribute, then \( R(U, p) \) is a probabilistic temporal relational scheme by adding \( p \).

A probabilistic temporal relation is shown in Figure 3 where \( a_{ij} \) \((j = 1, \ldots, n)\) are the temporal atoms and \( p(t_i) = p(A,B) \) is the joint probabilistic distribution of tuple \((a_{i1}, \ldots, a_{in})\). If tuples \( t_1 = t_2 \), then remove \( t_2 \), and change

\[
p(t_1) = \sum_{i=1}^k p(t_i).
\]

Figure 3 shows a probabilistic temporal relation.

**Example 2.** Recalling Figure 2, \( r' = \prod_{X}^{p} A(B|r) \) is shown in Figure 4. \( r' = \prod_{X}^{p} A(B,8,9)(r) \) is obtained in Figure 5.

(ii) Join on two relations \( r(X + p_X) \) and \( r'(Y + p_Y) \), denoted by \( r \bowtie r' \), is defined as follows:

(a) Compute the natural join, \( r \bowtie r' \).

(b) Add a new attribute \( p_{XY} \) to \( r \bowtie r' \), \( p_{XY}(t) = p_{X}(t_i) * p_{Y}(t_j) \) where \( t_i \in r \) and \( t_j \in r' \) and \( t = t_i \bowtie t_j \). \( r \bowtie r' = \prod_{X}^{p} A(B,8,9)(r) \) is obtained in Figure 5.

**Example 3.** Let \( R(X + p_X) \) and \( R(Y + p_Y) \) be two probabilistic temporal schemes, where \( X = \{A(3,5), D(6,8)\} \) and \( Y = \{A(3,5), B(9,10)\} \).
The inverse relation \( r \) of \( X \rightarrow P \) is defined by setting:

\[
P(t_i) = p^{-1}(t_i) \quad \text{if } p(t_i) = 0.
\]

We say that the temporal functional dependency \( t \rightarrow \mu X \) and \( t \rightarrow \mu Y \) holds on a temporal relation \( R(X, Y, Z) \) if, for all pairs of \( t_i \) and \( t_j \) in \( R \) such that \( t_i[X] = t_j[X] \) and \( t_i[Y] = t_j[Y] \), there exists another tuple \( t_{ij} \) in \( R \) such that:

\[
\mu_{ij} = (e, f) \quad \text{for all } e \in \mu_X, \text{ and } f \in \mu_Y.
\]

Example 4. Recalling \( r \) in Figure 6, \( r^{-1} \) is shown in Figure 9.

5. TEMPORAL DATA DEPENDENCIES IN PTRMS

In this section three kinds of temporal data dependencies in PTRMs are introduced. They are the PTFD, PTMD and the PTAJD.

5.1. PTFD

Definition 5.1. Let \( r \) be a relation over temporal scheme \( R(U) \) and \( X, Y \subseteq U \). Let \( t_i, t_j \) be any two tuples of \( r \). If \( t_i[X \cdot \mu_X] = t_j[X \cdot \mu_X] \) implies \( t_i[Y \cdot \mu_Y] = t_j[Y \cdot \mu_Y] \), we say that the temporal functional dependency \( X \cdot \mu_X \rightarrow d Y \cdot \mu_Y \) is true in \( r \), where \( t_i[X \cdot \mu_X] \) represents the value of \( t_i[X] \) in time interval \( \mu_X \), \( d \) represents the delayed time between \( \mu_X \) and \( \mu_Y \) (i.e. if \( \mu_X = [a, b], \mu_Y = [e, f] \), then \( d = e - a \)). Of course, if \( \mu_X \) and \( \mu_Y \) are synchronous, then \( d = 0 \).

Example 5. Figure 10 shows that the current price of materials \( M \) influences price of products \( P \) two months later. Since \( t_1[M, (5, 7)] = t_2[M, (5, 7)] = 5 \) and \( t_1[P, (7, 9)] = t_2[P, (7, 9)] = 10 \), \( t_i[M, (7, 9)] = t_i[P, (7, 9)] = 6 \) and \( t_i[P, (9, 11)] = t_i[P, (9, 11)] = 12 \), then \( r \) satisfies the asynchronous dependency \( M \rightarrow d P \), where \( d = 2 \).

5.2. PTMD

Before introducing PTMD a TMD is first extended to consider delayed conditions.

Definition 5.2. Let \( R(U) \) be a temporal relational scheme and \( r \) be a relation over \( R \). Let \( X, Y, Z \subseteq U \). The TMD \( X \rightarrow Y \) holds on \( r \), if for all pairs of \( t_i \) and \( t_j \) in \( r \) such that \( t_i[X \cdot \mu_X] = t_j[X \cdot \mu_X] \), there exists another tuple \( t_{ij} \) in \( r \) such that \( t_i[X \cdot \mu_X] = t_j[X \cdot \mu_X] = t_i[Y \cdot \mu_Y] = t_j[Y \cdot \mu_Y] \).\( t_i[Z \cdot \mu_Z] = t_j[Z \cdot \mu_Z] \).

Note that there is a delayed time between \( \mu_X \) and \( \mu_Y \) or \( \mu_X \) and \( \mu_Z \). There are two important properties of TMDs.

(i) Any TMD can be regarded as a lossless decomposition. More precisely a relation \( r \) satisfies the TMD \( X \rightarrow Y \) if \( r \) is losslessly decomposed into two schemes \( R_1 = XY \) and \( R_2 = XZ \).

(ii) The TMD \( X \rightarrow \cdots Y/Z \) indicates that \( Y \) and \( Z \) are conditionally independent given \( X \).

Now, we define PTMD.

Definition 5.3. Let \( R(U, p) \) be a probabilistic temporal scheme and \( r \) be a relation over \( R \). Let \( X, Y, Z \subseteq U \) and \( Z = U \). The TMD \( X \cdot \mu_X \rightarrow Y \cdot \mu_Y \) satisfies PTMD, if \( r = \prod_{X \cdot \mu_X} p_{(X \cdot \mu_X)} \cdot \prod_{X \cdot \mu_Y} p_{(X \cdot \mu_Y)} \cdot (\prod_{X \cdot \mu_z} p_{(X \cdot \mu_z)})^{-1} \). For
simplicity $\prod^p_{\mu_1, Y \mu_2}(r) \times \prod^p_{X \mu_3, Z \mu_4}(r) \times (\prod^p_{X \mu_5}(r))^{-1}$ would be written as $\prod^p_{X \mu_1, Y \mu_2, Z \mu_4}(r) \otimes \prod^p_{X \mu_3, Z \mu_4}(r)$.

Example 6. Consider the $r$ in Figure 10. Let $r_1 = \prod^p_{M(5,7),P(7,9)}(r)$ shown in Figure 11, $r_2 = \prod^p_{M(5,7),N(9,10)}(r)$ shown in Figure 12 and $r_3 = (\prod^p_{M(5,7)}(r))^{-1}$ shown in Figure 13. By definition of PTMD $r_1 \otimes r_2$ is shown in Figure 14.

5.3. PTAJD

By the theory of relational databases if a relation $r$ satisfies the multi-valued dependency $X \rightarrow Y$ then $R_1 \cap R_2 \rightarrow \cdots \rightarrow R_1$ is a special case of join dependency $*[R_1, R_2]$. It is desirable to extend PTMD to a PTAJD because a PTAJD can be viewed as a BN factorization [14, 15, 16].

Now, we further extend PAJD to PTAJD.

Definition 5.4. Let $R(R_1, \ldots, R_n, p)$ be a probabilistic temporal scheme, and $r$ be a relation over $R$. If there is a permutation $S_1, \ldots, S_n$ of $[R_1, R_n]$ such that $\langle s_1 \Rightarrow s_2 \rangle \Rightarrow \cdots \Rightarrow s_n$ has increasing join property, where $s_i = \prod^p_{S_i \mu_i}(r)$, then $r$ satisfies PTAJD $\otimes [S_1, \ldots, S_n]$.

There are important properties of temporal acyclic join dependencies (TAJDs).

Let $*[R_1, \ldots, R_n, p]$ be a TAJD over $R$, and $r$ be a relation of $R$. Then there is a permutation $S_1, \ldots, S_n$ of $R_1, \ldots, R_n$ such that each join operation `$\Rightarrow$' on $\langle s_1 \Rightarrow s_2 \rangle \Rightarrow s_n$ is lossless.

By definition we have:

(i) A relation $r$ satisfies MVD $X \cdot \mu_x \rightarrow Y \cdot \mu_y$, if $r$ satisfies PTMD $X \cdot \mu_x \rightarrow \cdots \rightarrow Y \cdot \mu_y$.

(ii) A relation $r$ satisfies TAJD $*[R_1 \cdot \mu_1, \ldots, R_n \cdot \mu_n]$, if $r$ satisfies PTAJD $\otimes [R_1 \cdot \mu_1, \ldots, R_n \cdot \mu_n]$.

The next example shows the relationship between PTAJD and the joint probability distribution (JPD).

Example 7. Let $R(U, p)$ be a probabilistic scheme, where $U = \{A_1, \ldots, A_6\}$. Let $R = \{R_1, R_2, R_3, R_4\}$ be a decomposition of $R$, where $R_1 = \{A_1, A_2, A_3\}$, $R_2 = \{A_1, A_2, A_4\}$, $R_3 = \{A_2, A_3, A_5\}$ and $R_4 = \{A_5, A_6\}$. Let $r$ be a relation satisfying PTAJD $\otimes [R_1 \cdot \mu_1, R_2 \cdot \mu_2, R_3 \cdot \mu_3, R_4 \cdot \mu_4]$. Then for any tuple $t \in r$ with value $(a_1, a_2, a_3, a_4, a_5, a_6) = p(a_1, a_2, a_3, a_4, a_5, a_6)$, since $r$ satisfies $\otimes [R_1 \cdot \mu_1, R_2 \cdot \mu_2, R_3 \cdot \mu_3, R_4 \cdot \mu_4]$, by definition $p(a_1, a_2, a_3, a_4, a_5, a_6)$ can be factorized as follows:

$$p(a_1, a_2, a_3, a_4, a_5, a_6) = \frac{p(a_1, a_1, a_2, a_2, a_3, a_3) p(a_1, a_1, a_2, a_2, a_4, a_4) p(a_1, a_1, a_2, a_2, a_5, a_5) \times p(a_2, a_2, a_3, a_3, a_5, a_5) p(a_2, a_2, a_3, a_3, a_6, a_6)}{p(a_1, a_1, a_2, a_2, a_3, a_3) p(a_2, a_2, a_3, a_3, a_5, a_5) p(a_2, a_2, a_3, a_3, a_6, a_6)}$$

Equation 22 indicates that for any tuple $t \in r$, a JPD $p(t)$ can be decomposed by a given PTAJD $\otimes R$.

5.4. The relationship among PTFDs, PTMDs and PTAJDs

In this section we discuss the relationship among three dependencies. The main results we obtained are: a set of PTFDs $F$ implies a set of PTMDs $M$, written as $F \models M$ and $M$ implies a PTAJD $J$, written as $M \models J$.

5.4.1. The relationship between PTFDs and PTMDs

Theorem 5.1. Let $R(U, p)$ be a probabilistic temporal scheme and $\beta = (R_1(U_1, p_1), R_2(U_2, p_2))$ be a decomposition of $R$. Let $F$ be a set of temporal functional dependencies over $R$. Then $\beta$ has a lossless join if $(U_1 \cup U_2) \rightarrow (U_1 - U_2) \in F'$ or $(U_1 \cup U_2) \rightarrow (U_2 - U_1) \in F'$, where $F'$ is the closure of $F$.

Proof. Let $XY \subseteq U$, $Z = U - XY$ and $U_1 = XY$, $U_2 = XZ$. Let $r$ be a relation over $R$ satisfying $F$.

(i) Let $r' = \prod u(r)$. By the relationship between TFD and TMD, $X \cdot \mu_x \rightarrow Y \cdot \mu_y$ logically implies $X \cdot \mu_x \rightarrow \cdots \rightarrow Y \cdot \mu_y$.

(ii) Since $r$ satisfies $X \cdot \mu_x \rightarrow Y \cdot \mu_y$, for any $t_1, t_2 \in r$, if $t_1[X \cdot \mu_x] = t_2[X \cdot \mu_y]$ then $t_1[Y \cdot \mu_y] = t_2[Y \cdot \mu_y]$.

Thus, for any tuple $(x \cdot \mu_x, y \cdot \mu_y, z \cdot \mu_z)$ in $r$, $p(x \cdot \mu_x) = p(x \cdot \mu_x, y \cdot \mu_y)$. Since $p(x \cdot \mu_x) = p(x \cdot \mu_x, y \cdot \mu_y)$, $p(x \cdot \mu_x) = p(x \cdot \mu_x, y \cdot \mu_y, z \cdot \mu_z)$. Thus, $p(x \cdot \mu_x, y \cdot \mu_y, z \cdot \mu_z) = p(x \cdot \mu_x, y \cdot \mu_y) p(y \cdot \mu_y, z \cdot \mu_z) p(x \cdot \mu_x)$. From (i) and (ii) we know that if $r$ satisfies $X \rightarrow Y$, then $r$ satisfies $X \rightarrow \cdots \rightarrow Y[Z]$. As a result if $r$ satisfies $X \cdot \mu_x \rightarrow Y \cdot \mu_y$, then $r$ satisfies $X \cdot \mu_x \rightarrow \cdots \rightarrow Y \cdot \mu_y$.

Example 8. Let $F = \{ X \cdot \mu_x \rightarrow A_1 \cdot \mu_{A_1}, \ldots, X \cdot \mu_x \rightarrow A_n \cdot \mu_{A_n}, Y \cdot \mu_y \rightarrow B_1 \cdot \mu_{B_1}, \ldots, Y \cdot \mu_y \rightarrow B_k \cdot \mu_{B_k}, Z \cdot \mu_z \rightarrow C_1 \cdot \mu_{C_1}, \ldots, Z \cdot \mu_z \rightarrow C_m \cdot \mu_{C_m} \}$ be a minimal set of PTFDs over $R(U, p)$. By Theorem 5.1 the corresponding set of PTMDs from $F$ is: $M = \{ X \cdot \mu_x \rightarrow A_1 \cdot \mu_{A_1}, \ldots, X \cdot \mu_x \rightarrow A_n \cdot \mu_{A_n}, Y \cdot \mu_y \rightarrow B_1 \cdot \mu_{B_1}, \ldots, Y \cdot \mu_y \rightarrow B_k \cdot \mu_{B_k}, Z \cdot \mu_z \rightarrow C_1 \cdot \mu_{C_1}, \ldots, Z \cdot \mu_z \rightarrow C_m \cdot \mu_{C_m} \}$.

5.4.2. The relationship between PTMDs and PTAJDs

From Example 8 we see that the right side of each PTMD is just a single attribute, because all PTMDs are obtained from a minimal set of PTFDs. By Definition 5.4 for a given PTAJD, we have a set of PTMDs to be equivalent to it. However, for a given set of PTMDs $M$, there may not exist a PTAJD. The next example shows this case.

Example 9. Let $R(U, p)$ be a probabilistic scheme and $U = \{A_1, A_2, C, D, E, F\}$. Let $M = \{ A \cdot \mu_A \rightarrow B \cdot \mu_B, D \cdot \mu_D \rightarrow C \cdot \mu_C \Rightarrow E \cdot \mu_E, E \cdot \mu_E F \cdot \mu_F \Rightarrow D \cdot \mu_D, B \cdot \mu_B C \cdot \mu_C \Rightarrow E \cdot \mu_E \}$ be a set of PTMDs over $R$. Since $D \cdot \mu_D \rightarrow \cdots \rightarrow E \cdot \mu_E \in M$, and $E \cdot \mu_E F \cdot \mu_F \Rightarrow D \cdot \mu_D \in M$, we know that $M$ includes a cycle. Therefore, there is not any PTAJD equivalent to $M$.

The above example raises a question: for a given set of PTMDs $M$, does there exist a PTAJD equivalent to $M$ or not? The next Lemma answers this question.
LEMMA 5.1. Let \( R(U, p) \) be a probabilistic scheme, \( XY \subseteq S \subseteq U \) and \( r \) be relations over \( S \) and \( r' = \prod_S^2(r) \) be a relation over \( R(S, p) \). If \( r \) satisfies \( X \cdot \mu_X \rightarrow Y \cdot \mu_Y \), then \( r' \) satisfies \( X \cdot \mu_X \rightarrow Y \cdot \mu_Y \) too.

Proof. Can we immediately obtain the result by the definition of PTMDs.

\[ \square \]

DEFINITION 5.5. Let \( R(U, p) \) be a probabilistic scheme and \( M \) be a set of PTMDs over \( R \). A decomposition tree \( T \) for \( R \) given \( M \) is a binary tree with following properties:

(i) The nodes in \( T \) are labeled as \( R_i(U_i, p) \), where \( U_i \subseteq U \).

(ii) The root of \( T \) is labeled as \( R \).

(iii) If a node \( R_k \) has two children \( R_i \) and \( R_j \) then \( (R_i, R_j) \) is a decomposition of \( R_k \) such that \( R_k = R_i \otimes R_j \in M \).

By the classical theory of relational databases we know that a multi-valued dependency can be looked upon as the binary join dependency. In probabilistic temporal relational models the conclusion still holds. In other words a PTMD \( R_1 \cap R_2 \rightarrow \cdots \rightarrow R_1 \) is equivalent to a PTAJD \( \otimes [R_1, R_2] \). It means that there is at least one PTAJD equal to \( M'(M' \subseteq M) \). We define such \( M' \) as follows:

DEFINITION 5.6. Let \( R(U, p) \) be a probabilistic scheme and \( M \) be a set of PTMDs over \( R \). Let \( \otimes R \) be a PTAJD equivalent to \( M'(M' \subseteq M) \). \( \otimes R \) is the maximal preserved PTAJD of \( M \), if there is no other PTAJD equivalent to \( M' \subseteq M \) and \( |M'| > |M'| \), where \( |M'| \) is the cardinality of \( M' \).

6. FROM TFDs TO TNBNS

Since we have shown PTFD implies PTMD, PTMD implies PTJD and PTJD is equivalent to a BN in Section 5, it motivates us to construct a TNBN from TFDs.

There are two problems that need to be solved when a TNBN is constructed from TFDs:

(i) How to obtain conditional independencies implied by \( F \)?

(ii) How to find a variable order to obtain the maximum conditional independencies implied by \( F \)?

It is important that some dependencies of \( F \) may not be applied to construct BNs. We illustrate this using the following example.

EXAMPLE 10. Consider the dependency graph obtained from \( F \) in Figure 15. There are two cycles \( A \rightarrow B \rightarrow C \rightarrow A \) and \( A \rightarrow B \rightarrow D \rightarrow A \). When deleting edges \( B \rightarrow C \) and \( D \rightarrow A \) an acyclic graph is obtained. When deleting edge \( A \rightarrow B \), we obtain another acyclic graph. The two acyclic graphs encode different conditional independency. The former loses two conditional independencies, the latter loses one conditional independency. It indicates that the structure of a TNBN is sensitive to the variable order. It is necessary to obtain a suitable variable order to obtain the maximum conditional independencies implied by \( F \).
FIGURE 16. A TNBN constructed from $F = \{A_1 \cdot \mu_1, A_2 \cdot \mu_2 \rightarrow A_3, A_1 \cdot \mu_1 \rightarrow A_4 \cdot \mu_4, A_3 \cdot \mu_3, A_4 \cdot \mu_4 \rightarrow A_2 \cdot \mu_2, A_4 \cdot \mu_4 \rightarrow A_6 \cdot \mu_6, A_4 \cdot \mu_4 \rightarrow A_5 \cdot \mu_5\}$.

is simplified to $p(A_5 \cdot \mu_5|A_4 \cdot \mu_4)$. Since $S_6$ is $p(A_6 \cdot \mu_6|A_1 \cdot \mu_1, A_2 \cdot \mu_2, A_3 \cdot \mu_3, A_4 \cdot \mu_4, A_5 \cdot \mu_5)$ and $F \models A_4 \cdot \mu_4 \rightarrow A_6 \cdot \mu_6$, $S_6$ is simplified to $p(A_6 \cdot \mu_6|A_4 \cdot \mu_4)$. Thus, the join distribution can be simplified as: $P(A_1 \cdot \mu_1, A_2 \cdot \mu_2, A_3 \cdot \mu_3, A_4 \cdot \mu_4, A_5 \cdot \mu_5, A_6 \cdot \mu_6) = p(A_2 \cdot \mu_2)p(A_3 \cdot \mu_3|A_2 \cdot \mu_2)p(A_1 \cdot \mu_1|A_2 \cdot \mu_2, A_3 \cdot \mu_3)p(A_4 \cdot \mu_4|A_1 \cdot \mu_1)p(A_5 \cdot \mu_5|A_4 \cdot \mu_4)p(A_6 \cdot \mu_6|A_4 \cdot \mu_4)$. The corresponding time slice of the TNBN is shown in Figure 16.

Furthermore, if for any time slice, $F = \{A_1A_2 \rightarrow A_4, A_1 \rightarrow A_4, A_3A_4 \rightarrow A_5, A_4 \rightarrow A_5, A_4 \rightarrow A_6\}$, we obtain the TNBN shown in Figure 17. In Figure 17, the causal relationship between $A_1$ and $A_4$ has delayed time $d_1$.

It is known that for a given joint distribution, it can be represented by several BNs. So PTFD is not a necessary condition for conditional independencies, because PTFD implies some conditional independencies. Thus, the obtained time slices are partial. In other words from TFDs, we can obtain an initial structure of a TNBN without complex calculations. Thus, we can further learn the complete structure of the TNBN by other methods [8] based on learned initial structure of the TNBN.

Though a simplified joint distribution can be learned from Algorithm 1, the learned dependency structure of a TNBN is sensitive to the variable order. Next we will show that for a fixed $F$, we can find a suitable variable order to construct a desired TNBN.

Let $M$ be a set of PTMDs implied by $F$ and $J$ be a PTAJD implied by $M$. We say that the variable order is suitable for $F$, if corresponding TNBN can obtain the maximum conditional independencies implied by $F$. More precisely, we have to find a way to learn an ordering of variables such that TNBN is a maximal preserved PTAJD of $M$.

Before we propose an algorithm to find a maximal preserved PTAJD of $M$, some related notations are given.

DEFINITION 6.1. Let $R(U, p)$ be a probabilistic temporal scheme and $M$ be a set of PTMDs over $R$. The connected graph of $M$, denoted by $G_m$, is a directed graph $(N, E)$ where $N$ is a set of nodes and $E$ is a set of edges. Every attribute in $M$ is a node of the $G_m$. If $A \cdot \mu_A, B \cdot \mu_B \rightarrow C \cdot \mu_C \in M$, then $(A, B) \in E$, $(B, C) \in E$. Both edges are labeled as $ab$.

EXAMPLE 12. Let $M = \{A \cdot \mu_A \rightarrow \cdots \rightarrow C \cdot \mu_C, D \cdot \mu_D \rightarrow \cdots \rightarrow E \cdot \mu_E, E \cdot \mu_E, F \cdot \mu_F \rightarrow \cdots \rightarrow D \cdot \mu_D, B \cdot \mu_B, C \cdot \mu_C \rightarrow \cdots \rightarrow E \cdot \mu_E\}$. The corresponding connected graph of $M$ is show in Figure 18.

DEFINITION 6.2. Let $R(U, p)$ be a probabilistic temporal scheme and $M$ be a set of PTMDs over $R$. Attribute $A$ is called an inverse key of $R$, denoted by $k_A^I$. if for each $X \cdot \mu_X \rightarrow \cdots \rightarrow Y \cdot \mu_Y \in M$, $A \notin X$.

The indegree of a node $A$, denoted by $In|A|$, is the number of incoming edges with different labels of this node. For example in Figure 18 $In|A| = 0$, $In|D| = 1$ and $In|E| = 2$.

We denote $RI(X)$ as a set of PTMDs such that the right side of each dependency in the variable $X$. For example in Figure 18, $RI(E)$ includes two dependencies: $B \cdot \mu_B C \cdot \mu_C \rightarrow \cdots \rightarrow E \cdot \mu_E$ and $D \cdot \mu_D \rightarrow \cdots \rightarrow E \cdot \mu_E$.

DEFINITION 6.3. A circuit in $G_m$ is a path $(e_1, \ldots, e_k)$ such that the end node of edge $e_k$ is equal to the start node of edge $e_1$.

There may be several circuits in a $G_m$. If $e_i$ is the common node of $j$ different circuits, then the weight of $e_i$ is $j$, denoted by $w(e_i) = j$. 

![Figure 17. A TNBN constructed from $F = \{A1A2 \rightarrow d_1, A3A1 \rightarrow d_2, A4 \rightarrow d_3, A5 \rightarrow d_4, A6 \rightarrow d_5\}$](image1)

![Figure 18. A connected graph of $M$.](image2)
An algorithm to find a maximal preserved PTAJD of $M$ is given as follows.

Algorithm 2
Input: a probabilistic temporal scheme $R(U, p)$, and a set of PTMDs $M$ over $R$.
Output: a queue $S$ to store an ordered set of PTMD.
Begin
1. Construct the connected graph $G_m$ of $M$.
2. Label the weight of each edge in $G_m$.
3. Calculate the indegree of each node of $G_m$.
4. Construct a minimal spanning tree $T$ of $G_m$ on the increasing order of weights and indegrees.
5. Construct an ordered set $S$ of PTMDs as follows.
   While the set of edges $E$ in $T$ is not empty Do
      1. Find an inverse key $k$.
      2. Add an edge $\{X \rightarrow \cdots \rightarrow k\} \in E$ to $S$.
      3. $E = E - RI(k_i^{-1})$.
   End While
6. Pop each element in $S$.
End

We use the following example to illustrate Algorithm 2.

Example 13. Let $R(U, p)$ be a probabilistic temporal scheme, $U = \{A, B, C, D, E, F, G, H\}$ and $M = \{H \cdot \mu_H, A \cdot \mu_A, D \cdot \mu_D \rightarrow \cdots \rightarrow A \cdot \mu_A, B \cdot \mu_B \rightarrow \cdots \rightarrow D \cdot \mu_D, G \cdot \mu_G \rightarrow \cdots \rightarrow E \cdot \mu_E, F \cdot \mu_F \rightarrow \cdots \rightarrow E \cdot \mu_E\}$ be two sets of PTMDs over $R$.

Step 1: Construct the connected graph $G_m$ of $M$, which is shown in Figure 19.

Step 2 and Step 3: Label the weight of each edge and indegree of each node in $G_m$, which is shown in Figure 20.

Step 4: Construct a minimal spanning tree $T$ of $G_m$, which is shown in Figure 21. Note that the weights of $\cdots \rightarrow E \cdot \mu_E, E \cdot \mu_E \rightarrow \cdots \rightarrow F \cdot \mu_F$ and $F \cdot \mu_F \rightarrow \cdots \rightarrow D \cdot \mu_D$ are equal. Since the entry degrees of $F \cdot \mu_F \rightarrow \cdots \rightarrow D \cdot \mu_D$ and $E \cdot \mu_E \rightarrow \cdots \rightarrow F \cdot \mu_F$ are smaller than $D \cdot \mu_D \rightarrow \cdots \rightarrow E \cdot \mu_E$, we select $F \cdot \mu_F \rightarrow \cdots \rightarrow D \cdot \mu_D$ and $E \cdot \mu_E \rightarrow \cdots \rightarrow F \cdot \mu_F$.

Step 5: Obtain $S$ as follows.
$I = 1 : k = A, s_1 = C \cdot \mu_C \rightarrow \cdots \rightarrow A \cdot \mu_A, E_1 = \{G \cdot \mu_G \rightarrow \cdots \rightarrow E \cdot \mu_E, E \cdot \mu_E \rightarrow \cdots \rightarrow F \cdot \mu_F, F \cdot \mu_F \rightarrow \cdots \rightarrow D \cdot \mu_D, B \cdot \mu_B \rightarrow \cdots \rightarrow C \cdot \mu_C\}$
$I = 2 : k = C, s_2 = B \cdot \mu_B \rightarrow \cdots \rightarrow C \cdot \mu_C, E_2 = \{G \cdot \mu_G \rightarrow \cdots \rightarrow E \cdot \mu_E, E \cdot \mu_E \rightarrow \cdots \rightarrow F \cdot \mu_F, F \cdot \mu_F \rightarrow \cdots \rightarrow D \cdot \mu_D, B \cdot \mu_B \rightarrow \cdots \rightarrow D \cdot \mu_D\}$
$I = 3 : k = D, s_3 = B \cdot \mu_B \rightarrow \cdots \rightarrow D \cdot \mu_D, E_3 = \{G \cdot \mu_G \rightarrow \cdots \rightarrow E \cdot \mu_E, E \cdot \mu_E \rightarrow \cdots \rightarrow F \cdot \mu_F\}$
$I = 4 : k = F, s_4 = E \cdot \mu_E \rightarrow \cdots \rightarrow F \cdot \mu_F, E_4 = \{G \cdot \mu_G \rightarrow \cdots \rightarrow E \cdot \mu_E\}$
$I = 5 : k = E, s_5 = G \cdot \mu_G \rightarrow \cdots \rightarrow E \cdot \mu_E, E_5 = \{\}$.

Notice that $S = (s_3, \ldots, s_5)$ is not unique. Selections on the edges with equal weights or indegrees will result in different PTAJDS.

The corresponding decomposition tree of $S$ [26, 27, 28] is shown in Figure 22.

Before proving the correctness of Algorithm 2, three lemmas are introduced.
LEMMA 6.1. Let R(U, p) be a probabilistic temporal scheme and M be a set of PTMDs over R. Let M*, a subset of M, be equivalent to a PTAJD. If X · μ_x → → A · μ_A is a PTMD in M*, then R(I(A)) = {X · μ_x → → A · μ_A}.

Lemma 6.1 indicates that there exists a unique non-trivial dependency in M* and its right side is A.

LEMMA 6.2. Let R(U, p) be a probabilistic temporal scheme and M be a set of PTMDs over R. If M*, a subset of M, is equivalent to a PTAJD, then there is no circuit in G_m^*.

From the structure of the decomposition tree, we know that Lemma 6.1 and Lemma 6.2 are correct.

LEMMA 6.3. Let G' be a minimal spanning tree of G_m. The number of edges in G' is biggest among all possible spanning trees of G_m.

By graph theory we know that Lemma 6.3 is correct.

Next we prove the correctness of Algorithm 2 by the following theorem.

THEOREM 6.1. Let R(U, p) be a probabilistic temporal scheme and M be a set of PTMDs over R. Let S = \{s_1, \ldots, s_n\} ⊆ M be the result of Algorithm 2, then S = \{s_1, \ldots, s_n\} is then a maximal preserved PTAJD of M.

Proof. Let M* ⊆ M be a PTAJD. Since M* is a PTAJD, by Lemma 6.1, there do not exist two non-trivial PTMDs with the same right side. Let the number of RI in M be m. By RI's definition, m is the largest possible number of different labeled edges in G_m^*. Let E_m^* = \{e_1, \ldots, e_m\} be a set of edges in G_m and E' = E - E_m^* = \{e'_1, \ldots, e'_{n-m}\}, where E is a set of edges in G_m and n is the number of edges in E. By Lemma 6.2, there does not exist any circuit in G_m. If G_m has some circuits we can transform G_m into an acyclic graph G_m by removing some edges of these circuits. The deleted edges are denoted by Q = \{f_1, \ldots, f_k\} (if i ≠ j then f_i ≠ f_j). The number of edges in Q is denoted as |Q|. Let Q' = Q ∩ E'. Thus we have |E_m^*| = |\{E - E'\}| - |Q| + |Q'|. |E_m^*| is expected to be maximum when |Q| - |Q'| is minimum. Since |E - E'| = m is constant, |E_m^*| varies only in accordance with |Q| - |Q'|.

Since T is a minimal spanning tree by Lemma 6.3, the number of edges of T is the biggest among the spanning trees of G_m. Therefore |Q| is minimum. Thus in algorithm 2, once the edge with the smallest indegree is selected, |Q'| is maximum.

In the rest of this section we show how to learn CPTs of BN from a temporal relational database.

Let R(U, p) be a probabilistic temporal scheme and XY ⊆ S ⊆ U. Let \( r = \prod E(r) \) be a relation over R'. By the definition of PTMD we have the following conclusion: if r satisfies X · μ_x → → Y · μ_y, then r' satisfies X · μ_x → → Y · μ_y.

DEFINITION 6.5. [26, 27]. Let R(U, p) be a probabilistic temporal scheme and M be a set of PTMDs over R. A decomposition tree T for R over M is a rooted binary tree with the following properties:

(i) The nodes in T are labeled Ri(Ui, p), where Ui ⊆ U.
(ii) The root of T is labeled R.
(iii) If a node labeled Ri has two children labeled R_i and R_j, then (R_i, R_j) is a decomposition of R_i and R_j.

EXAMPLE 14. Let R(U, p) be a probabilistic temporal scheme, \( U = \{A, B, C, D, E, F\} \). Let M = \{B · μ_B · \mu_C \rightarrow \rightarrow D · μ_D, E · μ_E \rightarrow \rightarrow F · μ_F, A · μ_A · B · μ_B \rightarrow \rightarrow C · μ_C\} be a set of PTMDs over R.

The decomposition tree T for R over M is shown in Figure 23.

Since E · μ_E \rightarrow \rightarrow F · μ_F holds on R implies that E · μ_E \rightarrow \rightarrow F · μ_F holds on R_2, we can decompose R into a tree T. In this example each multi-valued dependency in M is decomposed into two nodes of T, then T is a lossless decomposition tree of R over M.

We give Algorithm 3 to learn CPTs from a temporal database.

Algorithm 3
Input: a set of PTDF F, relation r_1, \ldots, r_n, and a factorization of p(a_1, μ_1,\ldots,a_n, μ_n).
Output: the structure and CPTs of a TNBN.
Begin
Step 1. Construct TNBN by Algorithm 1.
Step 2. For each factor in the factorization, 1. examine the schemes of the database,
2. select schemes that are the subset of the factor.
3. construct a decomposition tree based on PTMDs consisting of selected schemes.
Step 3. join \( r_i \) (\( i = 1, \ldots, m \)) on increasing order based on the decompositions tree;
Step 4. Obtain CPTs from the join of relation.
End

7. A DETAILED EXAMPLE

Example 15. We give an example of product and part to illustrate our proposed method to learn a TNBN. A product consists of different parts and the price of a product is influenced by many factors such as prices of parts, the quality and the stock of the product. Usually these influences are asynchronous. Supposing a temporal relational database consists of the following attributes: quality of part \((A)\), price of part \((B)\), stock of parts \((C)\), quality of products \((D)\), stock of products \((E)\), by semantics of attributes, we have a set of PTFDs \( F = \{ AC \rightarrow d_1, B, A \rightarrow d_2, D, BDG \rightarrow d_3, E \} \). Applying the algorithm of 3NF decomposition [28], \( R_1 = \{ A, C, B \}, R_2 = \{ A, C, G \}, R_3 = \{ A, D \}, \) and \( R_4 = \{ B, D, E, G \} \). The corresponding relations \( r_1 \) to \( r_4 \) of \( R_1 \) to \( R_4 \) are shown in Figures 24–27.

Now, we apply the proposed method on this product temporal database.
Based on the set of TFDs \( F = \{ AC \rightarrow d_1, B, A \rightarrow d_2, D, BDG \rightarrow d_3, E \} \) and in order \( (A, C, B, D, G, E) \). We obtain the following factorization by the chain rule.

\[
p(ACBDGE) = p(A)p(C|A)p(B|A, C)p(D|A, B, C)p(G|A, B, C, D)p(E|A, B, C, D, G).
\]

Since \( AC \rightarrow d_1 \) implies \( AC \rightarrow \cdot \), we have \( AC \rightarrow \cdot \) to \( B \). Similarly, we have \( A \rightarrow \cdot \) to \( D \), \( BDG \rightarrow \cdot \) to \( E \). Thus we can simplify the above factorization to \( p(ACBDGE) = p(A)p(C|A)p(B|A, C)p(D|A, B, C)p(G|A, C)p(E|A, C, D) \), which represents a TNBN as shown in Figure 28.
It is important to realize that $\otimes(A, C, B, D, G, E)$ is a maximal preserved PTAJD of $M$.

Now, we obtain the CPTs of each node of TNBN as follows. Following the order $(A, C, B, D, G, E)$, the relational schema $R_1$ is unioned if $R_i$ contains a factor of factorization $p(ACBDGE)$.

Since $R_1 = \{ABC\}$ covers the attributes of $p(B|AC)$ and $R_1 \cup R_2 = \{ABCD\}$ covers the attributes of $p(D|AC)$, then $r_1 \otimes r_2$ satisfies $AC \rightarrow D \rightarrow B|G$. Since $R_1 \cup R_2 \cup R_3 = \{ABCDG\}$ covers the attributes of $p(G|AC)$, then $r_1 \otimes r_2 \otimes r_3$ satisfies $A \rightarrow D \rightarrow B|CG$. Since $R_1 \cup R_2 \cup R_3 \cup R_4 = \{ABCDG\}$ covers the attributes of $p(E|ACD)$, then $r_1 \otimes r_2 \otimes r_3 \otimes r_4$ satisfies $D \rightarrow B|CG \rightarrow AC$. We have the decomposition tree as shown in Figure 29.

From the bottom of the decomposition tree to its top, we can calculate the joint distribution $p(a_3(1, 10), c_3(4, 6), b_3(5, 7), d_2(1, 10), g_1(6, 7), e_2(7, 10))$ as follows.

By $r_1 \otimes r_2 \otimes r_3$, we get $p(e_2(7, 10)|c_3(4, 6), a_3(1, 10), d_2(1, 10)) = p(e_2(7, 10), c_3(4, 6), a_3(1, 10), d_2(1, 10))/p(c_3(4, 6), a_3(1, 10), d_2(1, 10)) = 1/5 \div 1/5 = 1$.

Finally we get: $p(a_3(1, 10), c_3(4, 6), b_3(5, 7), d_2(1, 10), g_1(6, 7), e_2(7, 10)) = 0.2 \times 1 \times 1 \times 1 \times 0.2 = 0.04$.

8. CONCLUSION

In this paper we proposed a new approach to construct a TNBN from a set of TFDs. Our approach is different from traditional statistical approaches since we utilize the relationship between the theory of temporal relational databases and dynamic probabilistic reasoning models.

The TFDs in temporal databases are important data dependencies. Since functional dependency $F$ implies multi-valued dependencies $M$, $M$ implies join dependency $J$, and $J$ is equivalent to a BN, TPFs encode some conditional independencies in data. In this paper we gave an approach to construct a TNBN based on given TPFs. Because TPFs can be often obtained from real-world semantics instead of being mined from a large amount of data, the computation cost for constructing a TNBN is reduced. Our method has some salient advantages. First we extended temporal data dependencies into delayed conditions and presented a probabilistic temporal relational model. Second a TNBN was constructed from given TPFs. Next we showed that the learned TNBN encoded maximum conditional independencies implied by TPFs. Finally we also gave a method to obtain CPTs of constructed TNBN from temporal relational databases.

ACKNOWLEDGEMENTS

This work was supported by grants NNSF 60263003 from the National Natural Science Foundation of China, YNSF 2002F0011M from the Yunnan Natural Science Foundation of China and IIP 2002-2 from the Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences.

REFERENCES


Int. Conf. on Uncertainty in Artificial Intelligence, Madison, July 24–26, pp. 139–147, Morgan Kaufmann Publishers, California.


