

Markovian approach to time transition inference on bayesian networks

Adamo L. Santana, Diego L. Cardoso,
João C. W. Costa and Carlos R. L. Francês
Federal University of Pará
Brazil

1. Introduction

Also known in literature as belief networks, causal networks or probabilistic networks, Bayesian networks (BN) can be seen as models that codify the probabilistic relationships between the variables that represent a given domain (Chen, 2001); being one of the most prominent when considering the easiness of knowledge interpretation achieved. These models possess as components a qualitative (representing the dependencies between the nodes) and a quantitative (conditional probability tables of these nodes) structure, evaluating, in probabilistic terms, these dependencies (Pearl, 1988). Together, these components provide an efficient representation of the joint probability distribution of the variables of a given domain (Russel and Norvig, 2003).

An abundance of papers in literature study BNs and the many aspects and characteristics of their inherent architecture. The development of these studies have led the BNs to be known in many areas out of their original scope, and their application and capabilities are still being passed on to many other areas and domains. BNs are more known and popularised by the name of Bayesian networks.

The wide study and evolution of BNs has led not only to a spread in their usability, but also, perhaps most importantly, to their development and improvement. Their features (e.g. graphical modelling, representation, inference, analysis, diagnosis, etc.) have been carefully studied, and provided us with both enhanced quality and performance. The studies heretofore are, however, still a fraction of what can still be accomplished; for, as it holds true similarly as in many other models, there is plenty of room for improvements, whether it is on particular aspects, discovering new applications, creating new hybrid systems or models with its theoretical principles, etc.

The fact remains that BNs are now widely used in the most varied areas of study, and their use has spread such that nowadays they are not only limited to researchers, but also used by regular users, perhaps even unaware of the theory and mathematics behind them. Free and

commercial versions of programs implementing their algorithms are easily available and accessible.

This paper is mainly focused on the inference and representation of BNs. The main objectives are as follows: (i) present a time analysis approach for BNs based on Markovian models by using a graphical representation to model the networks' attributes and transitions; (ii) allow to directly model the effect of inferences in all the attributes of the network within their state space and instances of time; and (iii) to make possible for analyses of inferences considering the order that they are applied.

In section 2, some concepts of probabilistic networks are presented. The theoretical model proposed here is presented in section 3. The description of the model is target of section 4. Section 5 shows a case study application. In section 6, the final remarks are presented.

2. Probabilistic networks

A probabilistic network is composed of several nodes, with each node representing a variable (i.e. an attribute of the domain); arcs connecting them and whose direction implies in the relation of dependency between the variables; and probability tables for each node.

One of the major advantages of BNs is their semantics, which facilitates, given the inherent causal representation of these networks, the understanding and the decision making process for the users of these models [2]. This is basically because the relations between the variables of the domain can be visualised graphically, besides providing an inference mechanism that allows quantifying, in probabilistic terms, the effect of these relations.

We will consider here the notation for the probability of an event b given the evidence of a as $P(b|a)$, where $a \in A$ and $b \in B$, and A, B are variables of the BN. To calculate the posterior probability, the Bayes' Rule (1) is used.

$$P(b|a) = \frac{P(a|b)P(b)}{\sum_{b'} P(a|b')P(b')} \quad (1)$$

The analysis of BN presented here excludes the initial activity of creation of the BN graphical structure, assuming it has been previously made. This step is, however, of extreme importance, being when the independence relations are discovered (whether automatically or with the help of a domain expert).

The learning of the network's model is also complemented by the learning its parameters (i.e. the associated probabilities of the attributes), thus creating the structure representation (qualitative and quantitative). We will abstain to further detail this aspect here, but there are many papers in literature that study the learning of graphical representation of the PN and its details, among them (Cooper and Herskovitz, 1992), (Li et al., 2004), (Santana et al., 2007), (Spirtes et al., 1994) and (Zheng and Kwoh, 2004).

It can, then, be seen that BN represent a time variant model, representing the relations between the variables of a domain. Such relations are thus modelled in an architecture composed of nodes and directed arcs, and the direction of these arcs represent a relation of cause and effect; which, by definition implies on a relation of time, however brief it might be.

3. Background and theoretical model

In most works presented in literature, time analysis is made by using time series models. However, techniques such as dynamic Bayesian networks (DBN) (Murphy, 2002), hidden Markov models (HMM) (Rabiner and Juang, 1986) or Kalman filters (Kamllman, 1960) are more appropriate when there is a need to study the dependencies between variables, adding also a probabilistic reasoning. Hidden Markov models and Kalman filters can also be considered as particular cases of dynamic Bayesian networks (Nilsson, 1998).

The model presented here differs from the application of temporal or dynamic Bayesian networks, in which the time constraints are seen differently. While we observe each directed arc as the representation of a given instance of time t ; in a DBN, the full network structure is considered, remaining unchanged for each t , which is held separately.

The data model for a time series can be represented as a structure formed by a time scale with a number of k cases, where $k = 1, 2, \dots, t$; a number of j attributes $j = 1, 2, \dots, p$, usually divided into i discrete objects (or time intervals) which repeat throughout the studied period of time. Figure 1 presents the time series model according to the data cube representation (Dillon and Goldstein, 1984).

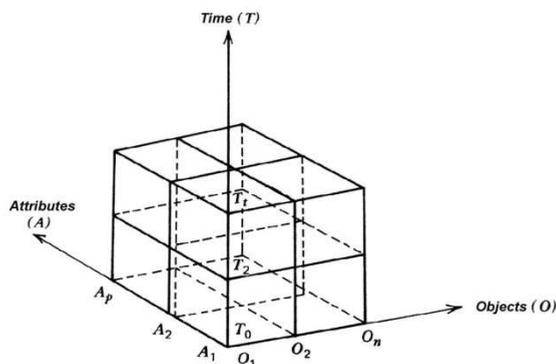


Fig. 1. Data cube structure.

A classical initial problem when working with BNs in the time would be the existing need to build conditional probability tables for each discrete unit of time analysed. Thus, a stationary random process is often assumed.

In the work described here, the time analysis and transition preceding from the BN are modeled into a discrete time Markov chain. Providing with means to compute, for example, the effect of a given inference after n units of time or how many units of time would take to achieve desirable probabilistic states for the attributes.

The approach presented uses the qualitative and quantitative data of the BN by modelling, for a given variable, a Markovian time transition matrix according to a first-order process; but also intrinsically considering the other variables of the domain, which might also influence in the behaviour of this attribute. This is because a BN can be seen as an array of attributes that might influence on one another over time.

To exemplify the model, a simple example of a BN can be considered, composed by only two variables: *Grade* and *Study*; where the grade obtained on a given test depends on the amount of study applied. It is also assumed that the tests are taken on a monthly time scale. It is considered as possible values for the attributes the following: Study (Hard, Medium, Little); and Grade (Excellent, Good, Regular).

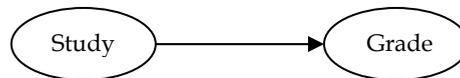


Fig. 2. Bayesian network for variables Grade and Study.

In this sense, the BN would also present the values of initial and conditional (for *Grade* only, given that it is the only attribute that possesses a parent attribute, that is, a dependence relation of the *Grade* given the *Study*) probabilities. The dependency model and the probability tables would represent all the data the BN could offer us.

Following the Markovian modeling, what we are seeking to obtain is the time instant that, given an inference, a determined probability configuration of an attribute would happen (e.g. considering our example, given that we study *Hard*, when we would obtain a grade *Excellent* with probability of 70%, *Good* with 25% and *Regular* with 5%).

Given that what we seek is in fact the new configuration of a determined attribute, what we end up needing is to set up the Markovian transition matrix of this attribute. This is done by mapping the transition probabilities for the states of the attribute onto the matrix, based on the conditional probabilities that it possesses given its dependencies with the other attributes (e.g. also considering the example, we must map the transition probabilities of *Grade* for: *Excellent* and pass to *Good*, *Excellent* to *Regular*, *Excellent* and achieving *Excellent* again etc). That is, we would have to compute the transition probabilities for the states of a given variable, which Markovianly speaking we can anagously see as the transition probability to achieve a state N_{t+1} based on N_t . Hence we seek to find the probability

$P(N_{t+1} = s_y \mid N_t = s_x) = p_{xy}$; thus creating a Markov transition matrix, according to the model in Table 1.

Grade\Grade	<i>Excellent</i>	<i>Good</i>	<i>Regular</i>
<i>Excellent</i>	P_{EE}	P_{EG}	P_{ER}
<i>Good</i>	P_{GE}	P_{GG}	P_{GR}
<i>Regular</i>	P_{RE}	P_{RG}	P_{RR}

Table 1. Model of the Markov transition matrix to be mounted

However, considering only the factor of study in relation to the grade is not enough to verify the relation of the variable *Grade* with itself and to make the transition between its states, as the Markov transition matrix would immediately converge to the stationary state. So, we must also consider the value of the attribute *Grade* at a previous point of time, acting together with the variable *Study* and thus obtaining the transition relations for the variable *Grade*.

For such, the first record in the existing historical database is ignored so that we can insert in the analysis, analogously to a 1st order Markovian process, the Previous Grade obtained. Tables 2 and 3 present the marginal and conditional (*Study*, *Grade* and the *Grade* in the previous period) probabilities of the Current *Grade* considering the *Study* and the Previous *Grade* (*Grade-1*).

Study		Grade	Grade	Grade-1
Hard (Ha)	0,133	Excellent (E)	0,210	0,333
Medium (Me)	0,534	Good (G)	0,467	0,333
Little (Li)	0,333	Regular (R)	0,323	0,333

Table 2. Initial probabilities of the Bayesian network.

Study ∩ G-1 \ Grade	E	G	R
Ha ∩ E	0,934	0,033	0,033
Ha ∩ G	0,333	0,333	0,333
Ha ∩ R	0,333	0,333	0,333
Me ∩ E	0,491	0,491	0,018
Me ∩ G	0,033	0,934	0,033
Me ∩ R	0,018	0,491	0,491
Li ∩ E	0,333	0,333	0,333
Li ∩ G	0,018	0,491	0,491
Li ∩ R	0,033	0,033	0,934

Table 3. Conditional probabilities of the Bayesian network - P(Grade | Study ∩ Grade-1).

The calculations for the Markov transition matrix would follow:

$$p_{EG} = P(E) \times [P(G | Ha \cap E)P(Ha) + P(G | Me \cap E)P(Me) + P(G | Li \cap E)P(Li)] \quad (2)$$

Generalizing, the Markovian transition matrix (Table 4) will be computed by mapping the transition probabilities of the states of a given variable; that is, the transition probability to achieve N_{t+1} based on N_t , being $P(N_{t+1} = s_y | N_t = s_x) = p_{xy}$.

$$\begin{matrix}
 & x_1^{t+1} & x_2^{t+1} & \dots & x_n^{t+1} \\
 x_1^t & \left[\begin{array}{cccc}
 p_{x_1} & p_{x_1x_2} & \dots & p_{x_1x_n} \\
 p_{x_2x_1} & p_{x_2} & \dots & p_{x_2x_n} \\
 \vdots & \vdots & \ddots & \vdots \\
 p_{x_nx_1} & p_{x_nx_2} & \dots & p_{x_n}
 \end{array} \right. \\
 x_2^t & \\
 \vdots & \\
 x_n^t &
 \end{matrix}$$

Table 4. General model of the Markov transition matrix.

The probabilities p_{xy} for the transition matrix are calculated according to:

$$p_{xy} = \frac{\sum_{i=1}^n P(s_y | s_x \cap Pa_i) \times P(Pa_i)}{\sum_{j=1}^m \sum_{k=1}^n P(s_j | s_x \cap Pa_k) \times P(Pa_k)} \quad (3)$$

where s represents the observed variable and its respective states; Pa is the variable that represents the parents of variable s ; m is the number of states the attribute can assume; and n is the number of possible states and/or combinations that the parents of this attribute can assume.

Consisting the denominator of the equation only as a normalized function (α), we have:

$$p_{xy} = \alpha \sum_i^n P(s_y | s_x \cap Pa_i) \times P(Pa_i) \quad (4)$$

Calculating from (4), we obtained the Markov transition matrix (represented by the letter P), presenting the transition probabilities for the states of the variable studied. For the considered example, we would have (Table 5):

Grade\Grade	<i>Excellent</i>	<i>Good</i>	<i>Regular</i>
<i>Excellent</i>	0.497	0.378	0.125
<i>Good</i>	0.068	0.707	0.225
<i>Regular</i>	0.065	0.318	0.618

Table 5. Markov transition matrix obtained.

Furthermore, to find the probability vector at a given time n , we need only to calculate the n th power of the probability matrix $P^{(n)}$, as described by the Equations of Chapman - Kolmogorov (Bolch et al., 1988).

$$P^{(n)} = P^{(m)} \times P^{(m-n)} \quad (5)$$

where $P^{(n)}$ is the transition matrix in the step n ; and thus $P^{(n)} = P^n$.

Thus, following on the example, if the unit of time is discretized in months and if we wanted to obtain the probabilities for the grades occurrence three months from now, we would have to find the power P^3 of the matrix (Table 6).

Grade\Grade	<i>Excellent</i>	<i>Good</i>	<i>Regular</i>
<i>Excellent</i>	0.1878	0.5274	0.2851
<i>Good</i>	0.1085	0.5561	0.3359
<i>Regular</i>	0.1071	0.4976	0.3974

Table 6. States transition matrix in the step $n = 3$.

The analysis presented (Tables 5 and 6), considered the behavior of the domain, given the available data, in time without any inference being made. Such analysis, however, can also be made, thus providing make the analysis in time given the evidence of a determined state of a variable, being able, as well, to consider its impact in a given time step. As example, considering as fact that the level of *Study* applied to make the test was *Medium*, we would have (Table 7):

Grade\Grade	<i>Excellent</i>	<i>Good</i>	<i>Regular</i>
<i>Excellent</i>	0.491	0.491	0.018
<i>Good</i>	0.033	0.934	0.033
<i>Regular</i>	0.018	0.491	0.491

Table 7. Transition matrix considering the inference made - Study: Medium.

Thus, considering the inference made, we would have in a step $n = 3$ the following matrix (Table 8).

Grade\Grade	<i>Excellent</i>	<i>Good</i>	<i>Regular</i>
<i>Excellent</i>	0.150	0.805	0.045
<i>Good</i>	0.054	0.892	0.054
<i>Regular</i>	0.045	0.805	0.150

Table 8. Transition matrix in the step $n = 3$ considering the inference made - Study: Medium.

To go back from the Markovian transition matrix to the marginal probabilities of the variable we apply (6).

$$P(s_x^t) = \sum_i^n p_{ix} \times P(s_i^{t-1}) \quad (6)$$

From (6), the probabilities for each state of the attribute *Grade* in a time period $n = 3$ given the inference of *Medium Study* can be found. The probabilities for the attribute *Grade* considering the example given here are as follow: *Excellent* 0.083, *Good* 0.834 and *Regular* 0.083.

4. Model description

In order to keep track of the whole network, and allow to directly model the effect of inferences in all the attributes of the network - and not just one at a time, as it was initially specified in [4] - we will first ascertain the diagram representation, to which we will map the BN.

We take a simple example of a BN (Fig. 3), for sake of simplicity, from which we will explain and build our model. The BN consists of six binary variables $X = [A, B, C, D, E, F]$.

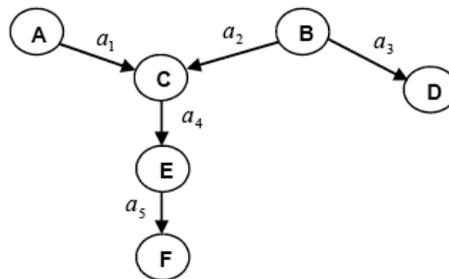


Fig. 3. Bayesian network example.

In Fig. 3 we see the existence of five arcs (a_1 to a_5) connecting the six variables of the BN, considering τ as the set of all r arcs in a BN, whereas $\tau = [a_1, a_2, \dots, a_r]$ and each arcs connects two nodes of the network. Notably, each arc of τ can represent a different time instant in the domain's transition timeline, from which an event (cause) inferred in the network will take to present an impact (effects) in the node directly connected to it.

We insert here the definition of *eras*. While this concept might be familiar to some, and has been applied in the literature of quantum networks (Tucci, 1998), we use it here with some different considerations, pending toward the analysis of each node.

The set of eras E , where $E = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, could be specified by removing successive layers of nodes [16], either internally (starting from the root nodes) or externally (starting from the leaf nodes). Considering the network in Fig. 3, by removing each layer of root nodes one after the other, we would have the eras as depicted in Fig. 4a. Similarly, by removing the layers of external nodes, the schemata would be as shown in Fig. 4b.

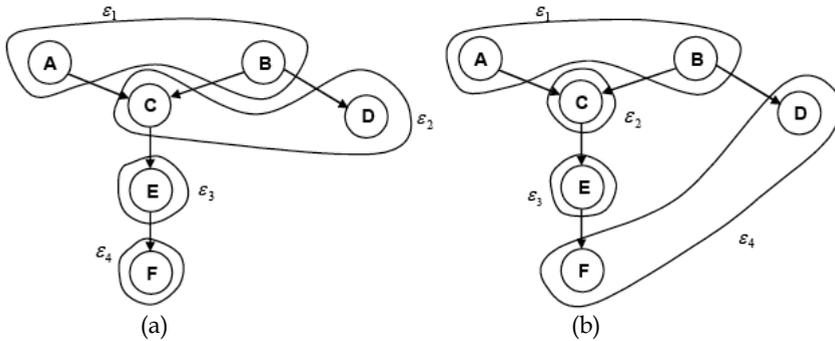


Fig. 4. (a) BN separated by eras considering root nodes removal; (b) BN separated by eras considering external nodes removal.

In the model proposed here, the eras can also be built starting from either the root or leaf nodes. For each era, separated space instants are drawn for each of the nodes held therein. For each of these spaces, a sub-network is placed, consisting of the node and its parents. From the BN graphical structure in Fig. 3, a temporal structure for the theoretical model presented here is built, as presented in Fig. 5.

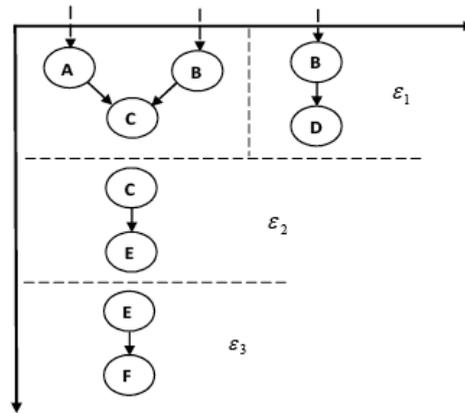


Fig. 5. Diagram structure for the BN.

In the diagram shown in Fig. 5 we separate the nodes of the network according to their dependencies and order of transitions. To visualize the network in this manner is useful when we consider that, for the Markovian time analysis that is induced, the transition matrix is calculated based on the individual node and the other nodes it is directly correlated.

The criteria for the subnets differs, however, for the one of a Markov blanket (Lauritzen , 1996), which, for any node in a BN, represents the set of nodes comprising the parents, the children, and the parents of the children of the node of interest (Chang et al., 2000); consisting here of a given node and its directly related parents as root nodes. To account for the diagnosis type of evidence analysis, however, which involves the backward flow in the active trail of the inference, the consideration of all correlated nodes involved would be necessary for the calculation of the transition matrix (thus making use of the entire Markov blanket space).

As described previously, the model presented here focuses on the analysis and inference method of the BN. The former is applied here by building a corresponding time specific model characterized by the transition of successive eras, from which the latter will take place by following classical search parameters (Pearl, 1988) for the inference calculation processing, defined by (4).

According to the probability rules of which BNs are based, and together with their graphical structure, it can be easily seen the reason for which an order of occurrence for multiple evidences in a BN cannot be defined. We can, however, make assumptions here to account the impact of such ordering, and to consider the evidences as being simultaneous or successive.

As it was stated, we can see the more than intrinsic relation between both, that is, that causal influence determines temporal relationship - as cause regularly precedes its contiguous

effect (Hume, 1975). Thus the temporal order is determined by the causal order (Carrier, 2003). Again, considering every arc in τ as a time transition instant.

The idea that the order of occurrence for multiple evidences cannot be defined in BNs presents as a faculty toward simultaneity, in which mutual evidences in the system are tied together. Such need for ordering might be irrelevant though, as the single probabilistic effect of each evidence can be differentiated by the path of causal arcs and the probability values.

So, while their evidence is made simultaneously, their effect can be seen as successive. The exception might be for the inferences made on the parent nodes of the given variable. In fact, unless the events are bound to the same space, their simultaneity in a given frame do not imply a simultaneity in another. Hence their grouping in subnets, for states that are simultaneous in experience stand in interaction with one another and are mutually tied together (Kant, 1787).

But what if, disregarding their causality order of the BN structure, a temporal analysis: as to an evidence e_a is made, and only after n units of time an evidence e_b is considered?

Taking the example of Fig. 5, and the evidence $P(E = e | C = c \wedge B = b)$, being e_a and e_b as $(C = c)$ and $(B = b)$, respectively, and $n = 2$. In this example, the influence of evidence e_a over C is continuous long before e_b is applied. A successive application of (4) would account on such matters of differently temporal instantiation of inferences (7).

$$P(a | e_1, \mathbf{K}, e_n) = \alpha \sum_t^n P(e_t | a) P(e_t) \quad (7)$$

It is important to notice that such assumption is only possible on some domain analysis. In the sense that, considering a model in which e_a would bring to an ultimate absorbing state (e.g. death, destruction, a deadlock in the system, etc.), no size of n or evidence e_b would cause any change of state.

The model presented here can also address such matter of evidence ordering; using (3) for building the transition matrix and thereon applying inferences according to the order of evidences e_t and their given time instant t . The transition model for our current example (Fig. 3 and 5) can be seen as described in Fig. 6.

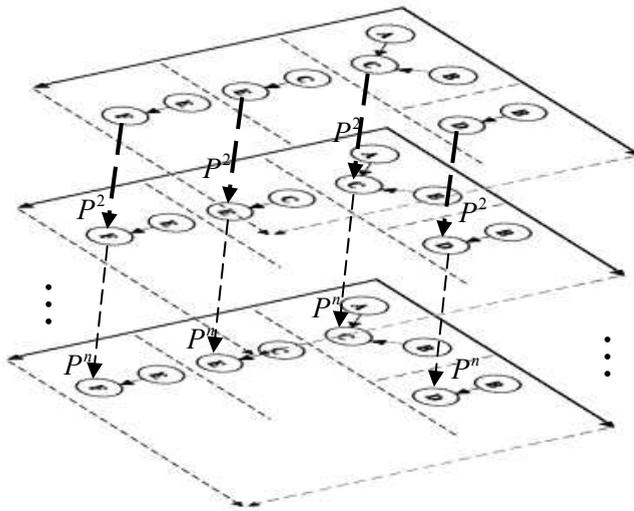


Fig. 6. Time transition model.

We are thus able to calculate the probabilities of the attributes in their time transitions, visualising the impact of the inferences as they progress. It is also elastic enough to allow the insertion of a new evidence at any time frame, visualising the influence of the evidences according to the order that they occur.

5. Case study application

An example of application of the proposed model in a case study in the area of power systems is presented next, ratifying the applicability of the method.

The analysis presented is part of a study made in (Rocha et al, 2006), to establish projections for the consumption of energy in a given region. One of the most desired aspects for power suppliers is the acquisition/sale of energy for a future demand. However, power consumption forecast is characterized not only by the variables of the power system itself, but also related to socio-economic and climatic factors.

Since the methods for load forecast use only the consumption data, it was necessary to offer means to analyze the correlations. Hence the use of Bayesian networks to codify the probabilistic relations of the variables and to make inferences on the conditions of the power system from the historical consumption and its correlation with the climatic and socio-economic data.

We present an application of the model for the power suppliers to project and correlate these parameters, studying the progression of their behaviour through time.

The data used in the work referred to a study of correlations for the consumption of energy of the city of Oriximiná - Pará and the climatic factors, established in a monthly time scale.

The database is composed by ten variables, with eleven arcs connecting them in the BN (Figure 7). The attributes denote the observed types of power consumption (residential, commercial, industrial and public) and climatic factors (temperature, relative humidity and pluviometric rate).

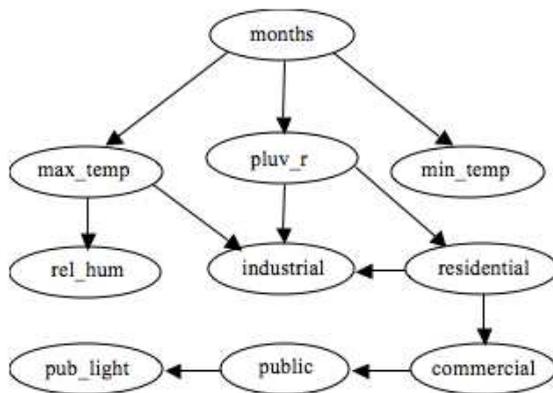


Fig. 7. Bayesian network correlating the Power consumption and the climatic factors

The analysis considered for example intends to study the changes occurred in the probabilities of the variable of commercial consumption (commercial), given an inference in the increase of the pluviometric rate, assuming this constant increment in a period of six months. The attribute of pluviometric rate (pluv_r), used to infer in the BN model, is a continuous variable by nature; its values, however, are presented as discretized in five states, according to the frequency of their values, which vary from a value of 1.479 to a maximum of 315.292 mm; the variable commercial, which represents the power consumption (in MW) in the commercial sector, had its values discretized in five states as well, varying from 126,918 to 219,649. The discretized states are displayed in Table 9.

Pluv_r	Commercial
[1.497 → 32.408)	[126,918 → 148,047)
[32.408 → 43.422)	[148,047 → 160,840)
[43.422 → 88.154)	[160,840 → 174,684)
[88.154 → 161.583)	[174,684 → 195,908)
[161.583 → 315.292]	[195,908 → 219,649]

Table 9. Discretized states of the variables pluv_r and commercial

The progression of the commercial consumption given the established hypothesis is computed according with the Equation (4), thus obtaining the Markovian transition matrix for the observed variable, as presented in Table 10. The discretized states (range of values), pointed in Table 9, are, for simplification, represented by labels from C_1 to C_5 , according to the increasing values of its states.

$$P = \begin{matrix} & & C_1 & C_2 & C_3 & C_4 & C_5 \\ \begin{matrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \end{matrix} & \left[\begin{array}{ccccc} 0.371 & 0.371 & 0.086 & 0.086 & 0.086 \\ 0.319 & 0.191 & 0.391 & 0.049 & 0.049 \\ 0.049 & 0.238 & 0.427 & 0.143 & 0.143 \\ 0.078 & 0.078 & 0.205 & 0.360 & 0.278 \\ 0.116 & 0.116 & 0.116 & 0.301 & 0.351 \end{array} \right] \end{matrix}$$

Table 10. Markovian transition matrix for the variable of power consumption

Its equivalent obtained after the sixth iteration, that is, the Markovian matrix representing the transition probabilities after a six months period, is presented in the following table.

$$P^6 = \begin{matrix} & & C_1 & C_2 & C_3 & C_4 & C_5 \\ \begin{matrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \end{matrix} & \left[\begin{array}{ccccc} 0.181 & 0.205 & 0.266 & 0.176 & 0.171 \\ 0.179 & 0.203 & 0.265 & 0.177 & 0.172 \\ 0.177 & 0.201 & 0.264 & 0.181 & 0.175 \\ 0.175 & 0.199 & 0.262 & 0.184 & 0.177 \\ 0.176 & 0.199 & 0.262 & 0.183 & 0.177 \end{array} \right] \end{matrix}$$

Table 11. Markovian transition matrix after the transition of six time units

Applying Equation (6), the marginal probabilities for the given analysis can be obtained again, identifying the following distributions for the commercial variable: $C_1 = 0.1776$; $C_2 = 0.2014$; $C_3 = 0.2638$; $C_4 = 0.1802$; and $C_5 = 0.1744$. Resulting in an update in the probabilities of the events and a presents the evidence of a higher consumption in the intermediate state, which ranges the values from 160,840 to 174,684 MW.

6. Remarks on the presented work

This work described a Markovian approach to represent the variables in a probabilistic network and their behavior throughout time, providing with a method for visualising and capturing their correlations.

The use of a Markovian model introduces advantages from its mathematical basis: the assumption that the present state depends only of its previous state and, adding to it, the fact that the Markovian models possess relatively simple solutions compared to its computational effort and to the mathematical complexity involved; which stimulates and facilitates its use.

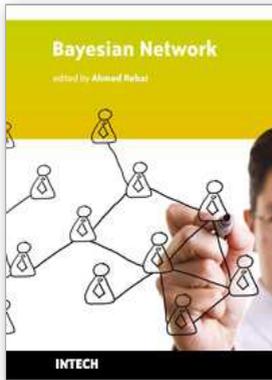
A Markovian approach for time transition is shown, highlighting the use of the network's structure, that alone expresses the relations of dependence and causality among the variables; and the probabilities associated to them, which serve as a basis for the creation of the Markovian transition matrix. Thus providing means for the study of the probabilistic transitions of the observed events, considering or not inferences, throughout the time.

The model also provides for the analysis of inferences considering the order in time that that they are applied in the network. This fact allows extending the interpretability of the probabilistic networks and adjusting them even further for applications of the real world.

7. References

- Bolch, G. ; Greiner, S. ; Meer, H. and Trivedi ; K. S. (1998) *Queuing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. John Wiley & Sons, Inc, New York, USA.
- Carrier, M. (2003) How to Tell Causes from Effects: Kant's Causal Theory of Time and Modern Approaches, *Studies in History and Philosophy of Science*, 34, (2), 59-71.
- Chang, K. C. ; Fung, R. ; Lucas, A. ; Oliver, R. and Shikaloff, N. (2000) Bayesian networks applied to credit scoring, *IMA Journal of Management Mathematics*, 11 (1), 1-18.
- Chen, Z. (2001) *Data Mining and Uncertain Reasoning - an Integrated Approach*. John Wiley Professional.
- Cooper, G. and Herskovitz, E. (1992) A Bayesian Method for the Induction of Probabilistic Networks from Data, *Machine Learning*, 9, 309-347.
- Dillon, W. R. and Goldstein, M. (1984) *Multivariate Analysis - Methods and Applications*. John Wiley & Sons.
- Hume, D. (1975) *An Enquiry Concerning Human Understanding*. Oxford University Press.
- Kalman, R. E. (1960) A New Approach to Linear Filtering and Prediction Problems, *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, 35-45.
- Kant, I. (1787) *Critique of pure reason*. P. Guyer, & A. W. Wood - Ed. & Trans., Cambridge University Press.
- Lauritzen, S. L. (1996) *Graphical Models*. Oxford University Press.
- Li, X. ; Yuan, S. and He, X. (2004) Learning Bayesian networks structures based on extending evolutionary programming, *Machine Learning and Cybernetics, Proceedings of 2004 International Conference on*, 3, 1594-1598.
- Murphy, K. (2002) *Dynamic Bayesian Networks: Representation, Inference and Learning*, PhD Thesis, Computer Science Division, UC Berkeley.
- Nilsson, N. (1998) *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann Publishers.
- Pearl, J. (1988) *Probabilistic Reasoning in Intelligent System*. Morgan Kaufmann Publishers.
- Rabiner, L. R. and Juang, B. H. (1986) An introduction to Hidden Markov Models, *IEEE ASSP Magazine* 3 (1), 4-16.

- Rocha, C. ; Santana, A. L. ; Frances, C. R. ; Rego, L. ; Costa, J. ; Gato, V. and Tupiassu, A. Decision Support in Power Systems Based on Load Forecasting Models and Influence Analysis of Climatic and Socio-Economic Factors. SPIE, v. 6383, 2006.
- Russel, S. and Norvig, P. (2003) *Artificial Intelligence*. Prentice Hall.
- Santana, A. ; Frances C. and Costa, J. (2007) Algorithm for Graphical Bayesian Modeling Based on Multiple Regressions, *Lecture Notes in Computer Science*, 4827, 496-506.
- Santana, A. ; Frances C.; Rocha, C. ; Rego, L. ; Vijaykumar, N. ; Carvalho, S. and Costa, J. (2007) Strategies for Improving the Modeling and Interpretability of Bayesian Networks, *Data & Knowledge Engineering*, 63, 91-107.
- Spirtes, P. R. ; Shcheines, R. and Clark, G. (1994) *TETRAD II: Tools for Discovery*. Lawrence Erlbaum Associates, Hillsdale, NJ, USA.
- Tucci, R. (1998) How to compile a quantum Bayesian net, arXiv, quant-ph/9805016.
- Zheng, Y. and Kwoh, C. K. (2004) Improved mdl score for learning of Bayesian networks, *Proceedings of the 2nd International Conference on Artificial Intelligence in Science and Technology, AISAT*, 98-103.



Bayesian Network

Edited by Ahmed Rebai

ISBN 978-953-307-124-4

Hard cover, 432 pages

Publisher Sciyo

Published online 18, August, 2010

Published in print edition August, 2010

Bayesian networks are a very general and powerful tool that can be used for a large number of problems involving uncertainty: reasoning, learning, planning and perception. They provide a language that supports efficient algorithms for the automatic construction of expert systems in several different contexts. The range of applications of Bayesian networks currently extends over almost all fields including engineering, biology and medicine, information and communication technologies and finance. This book is a collection of original contributions to the methodology and applications of Bayesian networks. It contains recent developments in the field and illustrates, on a sample of applications, the power of Bayesian networks in dealing the modeling of complex systems. Readers that are not familiar with this tool, but have some technical background, will find in this book all necessary theoretical and practical information on how to use and implement Bayesian networks in their own work. There is no doubt that this book constitutes a valuable resource for engineers, researchers, students and all those who are interested in discovering and experiencing the potential of this major tool of the century.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Adamo Santana, Diego Cardoso, Carlos Renato Frances, and João Costa (2010). Markovian Approach to Time Transition Inference on Bayesian Networks, Bayesian Network, Ahmed Rebai (Ed.), ISBN: 978-953-307-124-4, InTech, Available from: <http://www.intechopen.com/books/bayesian-network/markovian-approach-to-time-transition-inference-on-bayesian-networks>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.