

An Overview of Bayesian Network Applications in Uncertain Domains

Khalid Iqbal, Xu-Cheng Yin, Hong-Wei Hao, Qazi Mudassar Ilyas, and Hazrat Ali

Abstract—Uncertainty is a major barrier in knowledge discovery from complex problem domains. Knowledge discovery in such domains requires qualitative rather than quantitative analysis. Therefore, the quantitative measures can be used to represent uncertainty with the integration of various models. The Bayesian Network (BN) is a widely applied technique for characterization and analysis of uncertainty in real world domains. Thus, the real application of BN can be observed in a broad range of domains such as image processing, decision making, system reliability estimation and PPDM (Privacy Preserving in Data Mining) in association rule mining and medical domain analysis. BN techniques can be used in these domains for prediction and decision support. In this article, a discussion on general BN representation, draw inferences, learning and prediction is followed by applications of BN in some specific domains. Domain specific BN representation, inferences and learning process are also presented. Building upon the knowledge presented, some future research directions are also highlighted.

Index Terms—Uncertainty, knowledge discovery, Bayesian network, image processing, decision making, privacy preservation, system reliability estimation.

I. INTRODUCTION

Uncertainty is a commonly faced problem in real world applications. Uncertainty can be described as an inadequate amount of information [1]. Nevertheless, uncertainty may also exist in situations that have enough amount of information [2]. Furthermore, uncertainty may be alleviated or eliminated with the addition of new information. Addition of more information in complex processes may lead to mining of limited knowledge. Uncertainty can be computed

mathematically with probability theory. In uncertain situations, there is an involvement of possibility of states of attributes. Consequently, the models established on probabilistic inferences have the capability to assign a probabilistic value according to a defined principle. Accordingly, the prediction with large number of states in a model is accomplished. The question rises "how prediction is realized in the presence of large number of states in a model?" An answer to this question is the employment of Bayesian Network (BN) with several variables [3]-[5]. BNs, also known as belief networks, belong to the family of probabilistic graphical models. These graphical structures correspond to knowledge about an uncertain domain. More specifically, each node in the graphical structure represents a random variable, while the edges/arcs between the nodes represent conditional dependencies among nodes. These conditional dependencies are estimated by using acknowledged statistical and computational methods. Consequently, BNs incorporate concepts from graph and probability theory, computer science, and statistics.

Since last two decades, BN is recognized as an important tool for a number of expert systems especially in domains involving uncertainty [6]. This recognition of BN has several reasons behind it. First, BN encodes the dependencies of entire set of variables transparently, even if some data is missing. Second, causal links can be shown in a graphical form with variables which can easily be comprehended by humans. Third, BN can be used to build a bond between data and the prior knowledge in order to draw observations based conclusions. Lastly, BN, in conjunction with statistical methods, can be used to overcome data over fitting problem. In this way, the mathematical precision with probabilistic inferences is guaranteed with the help of BN in domains involving uncertainty. Several domains are identified by various researchers in terms of their application, learning or integration with BN and some of them include Image Processing [7]-[18] System Reliability Analysis and PPDM (Privacy Preserving in Data Mining) [19]-[24], Medical Diagnosis [25]-[32] and Decision Support [33]-[38].

In this study, we have explored applications of BN in a variety of domains having uncertainty. We have presented literature on BN based approaches in these domains. The objective is to highlight the importance of BN with detailed literature in different research domains. Moreover, the strengths and limitations of these approaches have also been reviewed.

The remainder of the paper is organized as follows. In Section II, the theoretical concepts and mathematical modeling of Bayesian Network is presented to support a better understanding of subsequent sections. In Section III to Section VII, an overview of application of BN in a select set

Manuscript received May 21, 2014; revised September 15, 2014.

Khalid Iqbal is with the Department of Computer Science and Technology, School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, P.R. China. He is also with the Department of Computer Science, COMSATS Institute of Information Technology, Attock Campus, Pakistan (e-mail: khalidqbal@ciit-attock.edu.pk).

Xu-Cheng Yin is with the Department of Computer Science and Technology, School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, P.R. China (corresponding author; tel.: +8613810223098, fax: +8610-62332873, e-mail: xuchengyin@ustb.edu.cn).

Hong-Wei Hao is with Institute of Automation, Chinese Academy of Sciences, Beijing 100190, P.R. China (e-mail: hongwei.hao@ia.ac.cn).

Qazi Mudassar Ilyas is with College of Computer Sciences and Information Technology, King Faisal University, Saudi Arabia (e-mail: qilyas@kfu.edu.sa).

Hazrat Ali is with the Department of Communication Engineering, School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, P. R. China. He is also with the Machine Learning Group, Department of Computer Science, City University London, United Kingdom (e-mail: engr.hazratali@yahoo.com).

of domains is presented. The emphasis of discussion in each domain is on BN representation, drawing inferences and learning mechanism. In Section VIII, future research directions are provided for researchers interested in this domain. Finally, the paper is concluded in Section IX.

II. BAYESIAN NETWORK

A. Historical Background

Historically, the term "Bayesian" refers to Thomas Bayes and is believed to be introduced for the first time in 1950s. A generalized version of the Bayesian theorem was independently demonstrated by Pierre-Simon, Marquis de Laplace. Besides this, BN was presented in [39] by emphasizing on three traits. The first two traits of BN are: subjective nature of input information and updating information. The last trait is the distinction between the connection and evidence reckoning mode which was stressed by Thomas Bayes as in [40]. Later in [5] and [41], the benefits of BN were emphasized to set up BN as a field of learning. Consequently, we can notice several research fields with the use of BN such as image processing, decision making, system reliability analysis and PPDM, medicine and data analysis, and further supplementary knowledge domains. The following sections present several methodologies based on BN, applied in the aforementioned fields of research.

B. Definition

A probabilistic graphical model based on a set of random variables (*r.v.*) and their conditional probabilities is known as Bayesian Network (BN) [3]. BN has two components, a Graphical Model (*G*) and the set of parameters Θ . *G* can be constructed from *r.v.* like $X_1, X_2, X_3, \dots, X_n$. Θ contains the states of each *r.v.* given the parents set π_i in *G*. A BN can be used to find the joint probability distribution over *r.v.* notations used in this work are presented in Table I.

TABLE I: NOTATIONS AND MEANINGS

Notations	Meanings
β	Bayesian network
<i>G</i>	Graphical Model
Θ	Set of Parameters
<i>X</i>	A Node
<i>V</i>	Set of Nodes
π_i	Set of Parents
x_i	State or Value of <i>X</i>

C. Features of Bayesian Network

The most significant feature of Bayesian Network is to put forward an effective mathematical structure. The fundamental need of such a structure is to model the complicated relationships among random variables in a straightforward way. Additionally, these relationships can also be visualized. Furthermore, the communication power of this structural representation permits to compare the domain expert's knowledge with expert system's knowledge in terms of probability as shown in Fig. 1.

The left most figure shows the conditional dependency of

B on *C* and *S* is conditionally independent of *B* and *C*. The middle figure represents the conditional dependency of *B* on *C* and *S*. The right most figure represents the conditional dependency of *C* on *S* and *B* on both *C* and *S*. Consequently, BN can be interpreted graphically over complex domains with the several advantages.

- 1) Bayesian theory can be used to assess risks involved in domains to formulate a decision
- 2) With the use of probability theory, uncertain inferences can be worked out in an unswerving and explicit way
- 3) Expert domain knowledge can be handled as prior distribution over statistical data in a practical way
- 4) BN can easily be interpreted semantically using domain expert knowledge
- 5) Both continuous and discrete variables can be handled after transformation

BN can handle missing data theoretically after marginalizing over all possibilities.

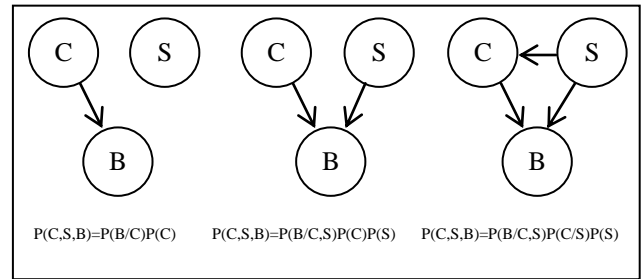


Fig. 1. Probabilistic relationship amongst variables.

D. Bayesian Theory

Let $\beta = (G, \Theta)$ be a BN. In *G*, let a *r.v.* $V = \{X_1, X_2, X_3, \dots, X_n\}$ with joint probability distribution with their values or states $x_1, x_2, x_3, \dots, x_n$ of *V*. The probabilities of these variables can be denoted as $P(X_1=x_1, X_2=x_2, X_3=x_3, \dots, X_n=x_n) = P(x_1, x_2, x_3, \dots, x_n)$. A BN corresponds to graphical model *G* which is a Direct Acyclic Graph (DAG). The structure of DAG is defined as vertices and directed edges. The vertices *v* is represented as the set of nodes in *V*; and the edges represent the relationship among the vertices. As a result, each vertex in graphical structure against *V* has its own Conditional Probability Distribution (CPD), which can be termed as $P(x_i / \pi_i)$. Hence, the Joint Probability Distribution (JPD) of BN is the product of CPDs.

$$P(x_1, x_2, x_3, \dots, x_n) = \prod_{i=1}^n P(x_i / \pi_i),$$

$$x_i \subseteq x_1, x_2, x_3, \dots, x_{n-1}$$

Therefore,

$$P(x_1, x_2, x_3, \dots, x_n) = \prod_{i=1}^n P(x_i / x_{i-1}, x_{i-2}, x_{i-3}, \dots, x_1) = \prod_{i=1}^n P(x_i / \pi_i).$$

Additionally, BN assumes independence assumption of nodes from the entire set of its predecessors except the direct parental nodes.

E. Bayesian Network Modeling

BN are ideally suited to domains that have inherent

uncertainty and require decision support. Random variables and their state space including the causal relationships are used to build a BN. The causal effects can be determined manually to construct BN from root node and influential nodes. Several BN tools are readily available to build BN as summarized in Table II.

With the use of these tools, the realization of x_i of X_i given π_i can be presented in a graphical form. For illustration of BN, consider an example as in [42] similar to [5], suppose a subject is suffering from backache due to some injury caused by improper posture during a sport activity. It has been found that uncomfortable chairs are installed in the subject's office. Some of the coworkers are reportedly having a similar backache. From this example, we can identify random variables like B, C, S, W and A . Each variable has the binary states 1 or 0. Thus, the compact factorization of JPDs is represented by BN as shown in graphical form in Fig. 2 along with the probability distributions of each node. In addition to example in Fig. 2, C and S are the parents of B while A is the child of B .

TABLE II: TOOLS FOR BUILDING BAYESIAN NETWORK

Tool Name	Source
Bayesian Builder	www.snn.ru.nl
Weka-3-7-10	www.cs.waikato.ac.nz
Netica for MS Windows	www.norsys.com
Banjo	www.cs.duke.edu
Bayes Server 4.0	www.bayesserver.com

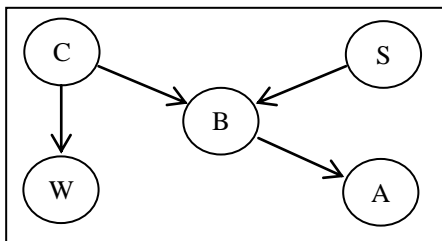


Fig. 2. Example of Bayesian network.

Similarly, W is the child of C . Furthermore, several independence statements can be observed amongst variables in BN shown in Fig. 2. For instance, B creates conditionally dependency; otherwise C and S are independent from each other. Likewise, W and B become conditionally independent when C is given. Also, A is conditionally independent from its ancestors C and S if B is given. Thus by chain rule, BN provides the compact factorization of JPDs of entire set of variables. This factorization can mathematically be presented as $P(C,S,W,B,A)=P(C)P(S|C)P(W|S,C)P(B|W,S,C)P(A|B,W,S,C)$.

The exclusive JPD of BN in a factored form can be defined as $P(C,S,W,B,A)=P(C)P(S)P(W|C)P(B|S,C)P(A|B)$.

Moreover, BN reduces the number of parameters from 31 to 10 in this case using 2^{n-1} where n represents the number of variables. As a result, the reduction of parameters provides benefits in terms of learning, inference and reduces the required computational resources.

F. Flexibility of Bayesian Network

The flexibility of BN can be evaluated in terms of its representation, inference and learning. Therefore, it is demanding to present BN-based domains according to

adaptation, use, integration, prediction and classification of information. Thus, we describe BN structure learning in a general way. The procedure of learning BN from prior knowledge is known as learning. The prior knowledge of a domain with uncertainty involves massive amount of statistical data. Moreover, the prior knowledge of a problem domain also contains network fragments, probabilities of nodes and their causal relationships. Based on such prior information, BN can be constructed with less effort as compared to manual construction. Thus, resultant network can be more precise than manually constructed BN structure over a sufficient amount of data. To learn BN structure, a prior knowledge is used in computation of conditional dependencies among variables in a particular domain under consideration. In this learning process, parameters of problem domain are taken into account to establish a structure from cases of a large database. Moreover, a JPD of each node in a structure is computed for the purpose of drawing inferences. In this way, no stakeholder can criticize because of the quantification of reasonable use of domain knowledge according to their rights. However, the question raises how to quantify a domain prior knowledge and predict the future knowledge? The answer to this question is simple because BN structure learning and inferences can be obtained with the use of readily available software programs such as Hugin¹ and Netica². These softwares can be used to classify and analyze data of a particular uncertain domain.

III. APPLICATION OF BAYESIAN NETWORKS IN IMAGE PROCESSING

The effectiveness of image processing techniques is severely hampered by the level of uncertainty in the subject task. This uncertainty issue can be resolved through BN. In this section, we present literature on application of BN in image processing in a concise and simple way. The subsequent sub-sections provide details on BN representation, its learning and derivation of inferences in image processing.

Rathod *et al.* [9] compared Image Segmentation Techniques (IST) to detect flaws in weldments. However, this comparative study was confined to region-based growing segmentation and less focused on BN approach. Watershed technique is considered to be the best IST despite its limitations such as insensitivity to noise and poor detection of thin boundaries. Similarly, Zhang *et al.* [43] focused on automatic and interactive mode for Image Segmentation (IS) supported by BN. The focal point of this work is to encode the relationships among regions, edges, vertices and angles. Thus, authors have provided a systematic way to model IS probabilistically by integrating new knowledge with user intervention to predict the next best spot for selection. In this way, a user is forced to select the predicted spot only. Wachsmuth *et al.* [14] integrated speech and images in order to associate spatial and information classes to establish relationship between similar objects resulting in a robust integration of speech and image. The efficiency is limited due

¹ <http://www.hugin.com/>

² <http://www.norsys.com/netica.html>

to a large number of instances in speech and image integration process. Luo *et al.* [16] proposed a unified framework to analyze street images with the use of BN. The purpose of the proposed framework is to build an inference engine for domain knowledge and training data to solve uncertain vision problem. The flexibility, interpretability, simplicity and statistically better performance are the key aspects of the proposed framework. Velikova *et al.* [13] came up with a unified probabilistic framework for information fusion to boost decision making in determining the characteristics of regions in complimentary image views. In this framework, accuracy is ensured in uncertain domain knowledge without considering the image resolution. Serrano *et al.* [17] mined low-level features from indoor and outdoor images to classify *Sky* and *grass* in efficient way. From low-level features, semantic features are used to predict and classify multiple classes of indoor/outdoor images in a simple and flexible manner. The semantic features of classified images are useful for a remote client because of features transmission rather than signal transmission. However, due to the limited number of low-level features, the proposed technique is unstable for correct classification decision as well as to detect entire set of semantic features.

Wang *et al.* [44] investigated BN for diagnosis of breast cancer by integrating image and non-image features sets. The image features are acquired from mammographic findings, while non-image features are obtained from physical examination and patients' clinical history (non-mammographic). These feature sets are pre-processed before diagnosing the breast cancer. BN classifier is trained upon the pre-processed features in determining the cancerous elements. Hongjun *et al.* [8] and Junejo [18] classified images and trajectories by proposing novel methodologies based on semantic and BN, and scene modeling to perform video surveillance respectively. The former method is requiring better technology than the currently available, while the later one intuitively handles small number of trajectories. Dynamic BN based MRI analysis is proposed by Chen *et al.* [7] to find interacting brain regions. This analysis has two stages: (1) Dataset preparation, (2) Two-slice Temporal BN (2-TBN) structural parameters. The objective of these stages is to analyze dissimilarities of interacting brain regions volume-change rate in normal-aging group and serene cognitive impairment group. This analysis is suitable for discrete time series data by modeling it explicitly. The approach becomes unstable and overfits due to the limited number of longitudinal morphometrics. Barat *et al.* [12] presented a descriptor combination method with an aim to enhance the recognition rate considerably. In this method, they used shape measures and shape descriptors in a more robust, scalable and less complicated way. As the method is dependent upon zernike descriptor which has inherent limitations of computing precision. Moreover, Generalized Fourier Descriptor (GFD) - a shape descriptor [45] is reliant on both. Jeon *et al.* [11] proposed video deinterlacing method with BN, a flexible approach to reliably measure weight support in a de-interlacing system. This method integrates pragmatic relationships into causal framework with inadequate number of parameters. Therefore, undersized dataset is used for the whole probability distribution of a BN.

Finally, traffic incident detection and novel BN representation for signature verification are proposed by Zhang *et al.* [10] and Xiao *et al.* [15] respectively. The former approach incorporates traffic incident detection into Automated Incident Detection (AID) algorithm to manage traffic knowledge with strong evidential reasoning in real time for making better decisions. The limitation of the method is the use of BN and dynamic BN without discretization process for quantified outcome. The latter approach attempts to fulfill the demands of personal signature identification and verification as a public facility. But, the mutually exclusive components, scalability and image rotation at certain angle are not justified. Table III summarized the application of BN in image processing for further research directions.

A. Representation of Bayesian Network in Image Processing

BN is used in various image processing applications such as in image segmentation [9], [43], development of unified framework [13], [16], classification image based information [8], [17], [18], interacted regions analysis of an image [7] with enhanced recognition rate [12], traffic incident detection [10] and signature verification [15]. BN is supportive in establishing the relationships among associated objects [14].

Let I be an image with p pixels states. In image segmentation, an original natural image can be referred to as an observed and segmented image with a limited number of regions. Original image is denoted as B_p and an observed image is denoted as A_p . Random fields of space can be represented as $B_p = \{b_1, b_2, b_3, \dots, b_n\}$ and $A_p = \{a_1, a_2, a_3, \dots, a_n\}$. To represent BN in this context, the probability of an original image, $P(B=b)$, can be referred to as priori probability. According to Bayes formula, $P(B=b/A=a)$ can be referred to posterior probability. This observable fact can be expressed as follows:

$$P(B_p = b / A_p = a, \psi, \omega) = (P(A_p = a / B_p = b, \psi, \omega) / Y) P(B_p = b) = \frac{\sum_{z=1}^n z^n P(A_p = a / B_p = z, \omega)}{\sum_{p=1}^n P(B_p = b, \psi)}$$

where, ψ and ω are hyper-parameters and $z = \{z_1, z_2, z_3, \dots, z_n\}$.

B. Drawing Inferences in Image Processing

In general, a direct application of BN is not promising as inferences in BN is a NP-complete problem [46]. NP-completeness means an exponential increase in complexity with number of variables n . However, exponential computational complexity can be minimized with the help of quality measures like K2 scoring function [47]. A lower complexity of inference results in a simpler algorithm.

According to Cox [48] and Luttrell *et al.* [49], the proposition of inference directs to use probabilities to get consistency in inferences. Inference direction approach in [48], [49] has three vital stages.

- Select a state space,

- Assign a joint probability density functions,
- Perform inferences by computing conditional probability density functions,

The above stages are referred to as Bayesian method. Bayesian method splits state space into two or more sub-spaces to discriminate between data and model parameters. Joint probability density function is used to make predictions over past data, future data and model parameters after computing the conditional probability density function of future given historical data. In this context, image segmentation [43] is performed with the use of image regions, edges, vertices and angles to extract new knowledge. The main objective of [43] is to predict the best possible spot with user intervention. In addition, relationships are identified in [14] by integrating speech and images to associate spatial and information classes. Besides, a unified image understanding framework with the use of BN in [16] addresses the problem of uncertainty in vision problem. The goal is to build inference engine for domain knowledge and training of data of few case studies. In [13], the characteristics of regions in complimentary images are determined probabilistically for boosting decision making. Low-level image of indoor and outdoor scenes are classified by predicting semantic features. In addition, video surveillance for scene modeling and novel approach based on semantics is performed to classify trajectories and images with the use of BN in [18] and [17]. Also, a dynamic BN and MRI analysis is carried out in [7] to find interacted regions. Moreover, a descriptor combination method based shape measures and shape descriptors are proposed in [12] to enhance recognition rate. BN-based video de-interlacing method integrated relationships practically into causal framework with limited number of parameters. Traffic incident detection [10] and signature verification [15] are inferred for traffic management with strong evidence-based reasoning and quick reflection in signature verification with the help of BN.

C. Bayesian Network Learning in Image Processing

In image analysis, the fundamental impediment is the non-uniform image dimensions. An image can either be segmented with the use of BN as in [50] or use of quality of measures such as scoring function. Another possibility is dimension reduction of an image by using measures like Principal Component Analysis (PCA). However, BN are used to segment image by learning visual features and identifying image regions followed by linking of the identified image regions. There is uncertainty involved in instances of an image corresponding to image regions. Hence, uncertainty of image instances can be associated with the use of BN in image segmentation process. In short, parameters such as features and labels of features of an image must be identified for learning a BN. This is required to obtain prior knowledge for predicting labels of an image by maximizing posterior conditional probability. The goal of maximizing posterior conditional probability of image features is to find joint probability distribution. The joint probability distribution provides information to make decisions about image regions, verification of information, incident management and classification of images and

trajectories as already discussed.

IV. BAYESIAN NETWORKS APPLICATION IN PPDM AND SYSTEM RELIABILITY

System Reliability Estimation is the calculation of performance of a system to achieve a specified task with the help of probability under stated constraints. Therefore, probability of system behavior is estimated with the use of BN to develop an effective PPDM model [23], [51]. PPDM (Privacy Preserving Data Mining) is a challenging research area to minimize the disclosure risks of XML association rules discovered by data mining techniques. The following sub-section gives application of PPDM as a reliable hiding strategy in preserving XML association rules with minimum side effects. BN-based techniques are reviewed to estimate the system behavior and a hidden strategy for XML rules with reliability.

Doguc *et al.* [19] proposed a BN model for SOE (System Operation Effectiveness) to substantiate relationships between system components and their developments over time. The objective is to extensively monitor a system for detection of abnormalities in an iterative way. Microsoft MSBNX software is used to build model with complex computations. Reliability of an entire system is estimated through BN by Doguc *et al.* [23]. Lower probability values of components are considered to be better while estimating the system behavior using BN. Luo *et al.* [21] surveyed the privacy preserving algorithms and classified them into three main categories; Heuristic-based, Data Reconstruction-based Association Rules and Cryptographic-based techniques. However, this survey does not present any privacy preserving rule mining technique based on BN specifically. Iqbal *et al.* [24] proposed to preserve XML Association Rules (XARs) with BN. In [24], sensitive node(s) are identified through BN using mode and maximum probabilistic node(s). These identified sensitive node(s) are used to perturb the largest sized transaction(s) in the original data source. The purpose of modifying the largest size transaction(s) is to keep the minimum perturbation to avoid generating new rules, ghost rules and lost rules with reliability. Despite this, sensitive XARs are also revealed. Wright *et al.* [20] presented a privacy-preserving protocol by modifying K2 scoring function to secure the distributed heterogeneous data by confining accessibility heuristically. Despite this heuristic approach, partial information is still revealed in perspective of relative ordering. In contrary to [20], Samet *et al.* [22] presented an Object-Oriented Bayesian Network (OOBN) protocol for homogeneously partitioned data to secure it over public channels against colluding attacks in an improved way for real world applications.

A. Bayesian Network Representation

In system behavior, probabilistic relationships among system components are assessed with the help of Conditional Probability Table (CPT). The components CPT can be obtained from $P(X/S)$; where S is the set of X 's parents. Each member in the set of parents in CPT of X 's is instantiated as either "Success" or "Failure". PPDM is introduced in [47], [52] which refers to protecting fragments of the original data.

This is a BN-based NP-hard problem [53] as argued in [54]. A variety of PPDM algorithms are proposed for association rules [55], [56] and XML Association rules [24], [56], clustering [57], naive Bayes classifiers [58], [59], statistical analysis [60], [61], and discovering common elements [30, 62], [63]. Amongst such a variety of PPDM research areas, we focused on application of BN for system reliability analysis and estimation [19], [23], and PPDM in XML association rule mining [24].

B. Inference Using Bayesian Network in PPDM and System Reliability

To demonstrate an application of BN of system behavior, a BN of five components interacting in a system is shown in Fig. 3. In this network, a relationship among components can be observed. Moreover, degrees of these relationships for each component are expressed as scores of entire possible candidate parents set in [23].

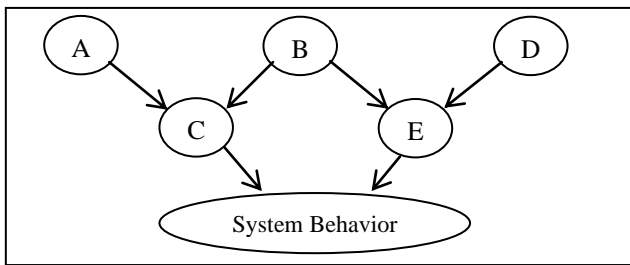


Fig. 3. Bayesian network example.

The top most nodes/components A , B and D are conditionally independent. However, the overall system behavior is dependent because of outgoing edges in the network from A , B and D to C and E . Therefore, the prior probabilities of these nodes/components must be known in advance based on historical data or with the help of domain expert. Moreover, components/nodes C , E are conditionally dependent and their probabilities should be calculated using Bayes theorem. The overall system reliability can be computed by including the components with 'Success' probability to get system behavior. BN is used in PPDM for hiding sensitive XML Association Rules as in [24]. The main objective of using BN in PPDM is to predict sensitive node(s) in large datasets like system behavior node. The purpose of finding sensitive node is to hide sensitive XML Association Rules. Consequently, BN supports and provides quite reasonable results in minimizing the disclosure risk [24].

C. Bayesian Networks Learning for PPDM with Reliability

In Fig. 3, BN presents a joint probability distribution over a finite set of components $\gamma = \{A, B, C, D, E\}$ of discrete random variables $\{A, B, C, D, E\}$. Formally, a BN is a pair of G and Θ i.e. $\beta = (G, \Theta)$. The learning process of BN can be described over γ . To construct BN over components of γ , the scoring function of K2 algorithm can be used to find a relationship. The process of scoring function is to find a relationship by maximizing the scores between a component and its parents set (predecessor components). As a result, the overall system behavior based on the given components of a system can be aggregated to make a reliable decision. Also,

K2 algorithm based BN is used in PPDM for locating the most frequent (contributing/sensitive) component/node [62], [63] to preserve association rules in XML domain [24]. The purpose of locating the most frequent component(s) is to perturb the original database with minimum effect. The apriori algorithm based XML association rules are compared with the mined XML rules after perturbation of original database. XML association rules mined from the perturbed database are matched with the original database mined rules according to hiding effect, side effects (lost and ghost/new XML Association rules).

V. BAYESIAN NETWORKS APPLICATION IN MEDICAL DOMAIN

Many researchers have used BN in medical domain because of inherent uncertainty in this domain and the ability of BN to handle uncertainty in the diagnosis process of diseases with improved patient caring facilities. In context of clinical tests for a diagnosis of a disease, BN can be helpful in selecting better alternative tests to ensure reliability. Literature review on BN-based techniques in medical domain is presented that support to resolve uncertainty issues in the treatment of patients.

Lee *et al.* [25] used BN for knowledge discovery from nursing database. Nevertheless, effectiveness cannot be judged without validated outcome. Aoki *et al.* [26] developed a BN-based methodology for efficient patrolling of nurses. The purpose of nurse patrolling is to improve the dependability of calling a nurse according to patient's needs by analyzing data based on demand level of nurse calls (time zone of previous nurse call, interval of previous nurse call, reason of previous call, and degrees of freedom in life) and patient's condition (sex, age, hospital department, and degrees of freedom in life). It may help in decision support process according to the condition of a patient and the level of need to call a nurse for an on-time attention. However, this approach has an inherent shortcoming associated with it; i.e. the need of frequent interruption by the supervisory staff. In support to clinical side, Bruland *et al.* [30] integrated Case-Based Reasoning (CBR) and Bayesian Network (BN) for making decision under uncertainty. Although the approach lacks an implementation, the authors have attempted to prove its effectiveness through intuition. Moreover, Flores *et al.* [64] integrated expert knowledge to examine different approaches to devise novel method for assessment of experimental parameters and transitional results. However, the method has limited application for all real variables, despite the fact of reduction of mining burden. Fenton *et al.* [31] assessed the risks of alternative medical diagnosis with BN. The objective is to support decision making with a variety of alternatives in medical diagnosis using BN. The assumption based decision making has the trade-off between the region accuracy and storage space. Chattopadhyay *et al.* [29] employed BN to examine dental pain. The aim is to build Bayesian classifier to analyze frequent dental diseases with an acceptable accuracy. The performance of Bayesian classifier is questionable because of halting tendency. Fuente *et al.* [28] presented interconnection

between biological abnormalities in Borderline Personality Disorder (BPD) by learning BN to address uncertainty using K2 algorithm. The objective is to evaluate the literature based collected variables for 20 BPD patients and developed a documented criterion. A limited number of patient's data is measured using BN-based interactive degree of robust conditional dependency among limited number of variables related to BPD for hypothesis testing. Furthermore, explicit diagnostic system on Mild Cognitive Impairment (MCI) is presented by Sun *et al.* [27]. This diagnostic system is capable of handling incomplete data for prediction and assist in improving precision in terms of execution time, lowest average and variance. Nonetheless, variation effect cannot be overlooked.

A. Bayesian Networks Representation in Medical Domain

BN have demonstrated a superb performance for predictions in medical domain. Therefore, BN techniques have been applied effectively in diagnosis like examining dental pain [29], BPD [28], risk analysis in alternative medical diagnosis [31] and efficient patrolling of nurses [26]. The advantage of BN in such domains is the handling of incomplete or missing data for prediction with precision. For a detailed BN representation, we consider a simple example as presented in Fig. 4. This BN has four nodes/variables in a clinical situation. Consider nodes A, B, C, D in Fig. 4 for Coughing, Spasm, High Pressure and Alarm respectively. The root nodes {A, B} for {Coughing, Spasm} respectively are related to a prior probability distribution and non-root nodes {C, D} for {High Pressure, Alarm} have local conditional probability distribution. The purpose of finding priori and local probability distribution is to quantify the parent-child relationship.

The joint probability distribution of each node can be computed by taking the product prior or conditional probabilities assigned to each node in BN according to chain rule as in [25]. The common topology of Conditional Probability Table of the given BN is shown in Table III which can be articulated in a joint probability as $P(A,B,C,D) = P(A)P(B)P(C/A,B)P(D/C)$.

B. Bayesian Networks Inference

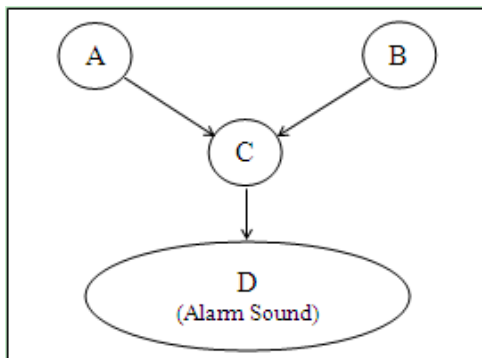


Fig. 4. Bayesian network example in medical domain.

Inference in BN refers to the computation of posterior probabilities conditioned on occurrence of evidence. Thus, BN provide a flexible way to select any node as a target node for inferences. BN are used in diverse problems like

diagnostics, causal inter-causal and mixed inferences. Diagnostics inferences are accomplished for justifying cause and effect relationship. Causal inferences are used for reasoning of causes to effects. Inter-causal inferences are used to identify the causes of a common effect. Mixed inferences referred to the combination of two or more of the aforementioned inferences. Consequently, the inferences of BN in Fig. 4 as reported by [25], have been described with the help of Table III and IV.

TABLE III: CONDITIONAL PROBABILITY TABLE OF FIG. 4

A	B	P(C/A,B)	
		Yes	No
Yes	Yes	0.9	0.1
Yes	No	0.95	0.05
No	Yes	0.7	0.3
No	No	0.1	0.9

Consequently, joint probability distribution table of BN in Fig. 4 can be represented in Table IV.

From the structure of BN, the joint probability distribution can be computed. For instance, the probability of event $D='Yes'$ can be computed in a situation with $A='No'$, $B='No'$ and $C='No'$. The following computations shows the outcome of $P(D='Yes', C='No', A='No', B='No') = P(D='Yes' / C='No') P(C='No' / A='No', B='No') P(A='No') P(B='No') = 0.5 \times 0.9 \times 0.2 \times 0.8 = 0.072$.

TABLE IV: JOINT PROBABILITY DISTRIBUTION OF FIG. 4

A	B	D=Yes		D=No	
		C=Yes	C=No	C=Yes	C=No
A=Yes	B=Yes	0.1505	0.0040	0.0015	0.0040
	B=No	0.5702	0.0320	0.0058	0.0320
A=No	B=Yes	0.0238	0.0080	0.00024	0.0080
	B=No	0.0158	0.0720	0.00016	0.072

C. Bayesian Networks Learning

Once the nodes and their states to be modeled have been identified, a BN can be constructed in two ways. First, manual construction of BN is completely dependent upon expert knowledge. In this structure learning process, causal influence with conditional independence of nodes is considered as shown in Fig. 4. Second, automatic or semi-automatic BNs are learnt through algorithms in knowledge discovery procedure in large datasets. These algorithms determine the independence or dependence relationship discovery among nodes. The potential outcome of learning BN structure refers to the investigation of hidden structure that conventional statistical methods fail to identify. A number of readily available tools can be utilized to learn BN structure using search-and-score-based algorithm and constrained-based-algorithms. A scoring metric is used in learning the best model structure with the use of *Bayesian Knowledge Discoverer (BKD)* [65] and *Bayesian Lab*³. On the other hand, a constraint-based algorithms finds the best possible BN structure by discovering overall conditional dependencies and independencies with the support of *HUGGIN* [66], *BN Power Constructor*, *Power Predictor* [7], [67], [68], or *TETRAD* [69]. The constraint-based algorithms

³ <http://www.bayesia.com/en/index.php> , Access_Date: 27-01-2014

use domain knowledge in specifying relationships among variables.

VI. DECISION MAKING AND BAYESIAN NETWORKS

BN support decision making by finding high-quality strategies with the use of most probable configurations of attributes. The following literature presents several BN applications for better decision making.

Lauria *et al.* [35] presented BN for IT implementation to support decision making by integrating BN into Decision Support System (DSS). This method has emerged as classification tool to make decision in uncertain situations. The classification tool has limited number of settings for illustrating relationships among variables of a real-world dataset on IT implementation. Another BN model, proposed by Pendharkar *et al.* [33], integrates risks to predict software development effort. This model establishes probabilistic bounds for managers to use it as a benchmark. However, the proposed BN-based model is constructed over readily available data with small number of instances. Additionally, this model ignores complex variables and claims to address uncertainty with assumptions. Besides this, Jaworska *et al.* [70] suggested BN approach for Integrated Testing Strategies (ITS) to recognize constraints. The main focus of identifying constraints in ITS realization is to provide a guiding resource for making decisions on management risks and a chemical hazard. Nevertheless, BN-based ITS realization can be used to track quality with constraint based decisions about risks and hazards. Therefore, ITS cannot be declared as a guiding resource as a ready-made managerial decisions. Watthayu *et al.* [36] presented Multi-Criteria Decision Making (MCDM) framework which focused uncertainty in making decisions based on BN and Influence Diagram (ID) (an extended version of BN). This approach has the explicit probabilistic independence in the entire set of elements. However, the explicit independent elements in MCDM framework can produce biased or invalid results and increases computational complexity. Nadkarni *et al.* [37] transformed causal maps into BN to provide a systematic procedure, named as Causal BN, by combining exploratory and confirmatory methods. Causal BN quantified wide-ranging relationships between variables with no prior assumption with strong probabilistic inference using Netica tool. However, causal BN approach is restricted to normal decision with the employment of ID framework. Baesens *et al.* [34] presented a BN classifier for discovering the slope of life cycle of loyal customers. The key purpose is to predict the increase or decrease in expenditure of a new customer. Consequently, unrestricted BN classifier showed better performance. But, Markov Blanket concept is used without stressing the way of selecting effective attributes. Zorilla *et al.* [38] used BN in participatory water resource management for integrating understanding, significance and standards in decision making. In this way, public involvement is boosted up over a complex dilemma, but BN generalization is confined in terms of updated verdict generation.

A. Bayesian Networks Based Decision Making

BN can be used in making decision in variety of

applications as discussed in previous sub-section. Based on this discussion, we have considered causal maps for representing BN for making decision. The procedure for constructing Bayesian causal maps can be divided into two stages namely: qualitative and probabilistic. The qualitative stage eliminates the limitations of modeling and makes causal maps compatible with the BN. This procedure has four major modeling issues. First, conditional independencies mean that a network can either be Dependence Map (D-Map) or Independence Map (I-Map). I-map is a graph in which the entire set unconnected of nodes corresponds to independent variables. On the other hand, D-map is a graph that represents the entire set of connected nodes corresponding to dependent variables. A model that follows both I-maps and D-maps is known as a perfect map. Further details of dependency model with necessary and sufficient conditions on perfect maps are given in [5]. Second, reasoning underlying cause-effect relations can either be abductive or deductive [71], [72]. Therefore, a reason is a rational ground or motive for an action or decision. Thus, abductive reasoning is a process that justifies a rational ground from effects to causes. The deductive reasoning is a process that justifies a rational ground from causes to effects. The difference between both reasoning processes is shown in Fig. 5(a) [37]. Third, distinguishing between direct and indirect relationships means to identify the conditional independencies as depicted in Fig. 5(b). The purpose is to draw attention to the conditional independencies assumption affects in causal maps [37]. Also, causal maps lack of distinguishing in direct and indirect relationship.

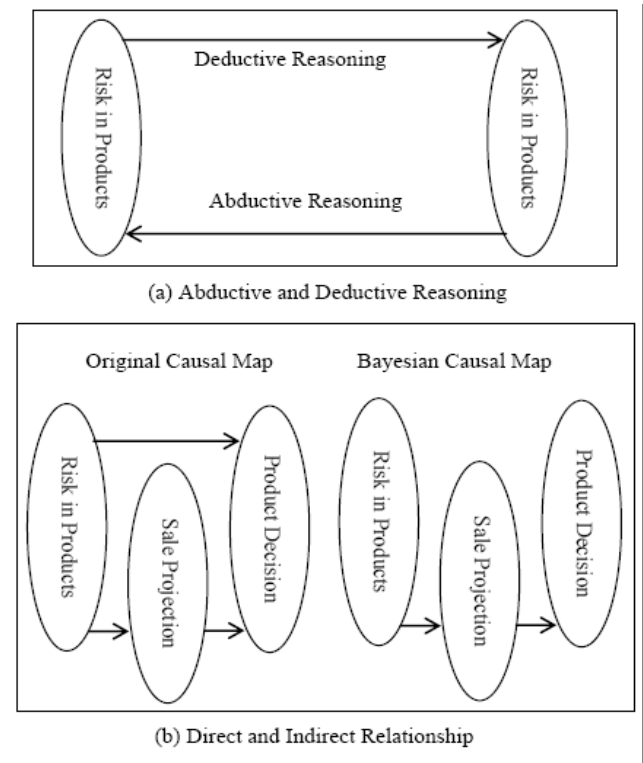


Fig. 5. Bayesian network in decision making.

B. Inference for Decision Making

For inferencing in decision making, uncertain associated nodes with background noise are considered. In causal maps,

features like uncertain nodes, background noise and interactive effects of multiple causal variables/nodes are represented except their qualitative analysis and validity. On the other hand, BN represents the precise knowledge about uncertain variables and background noise. Moreover, interactive multiple causal variables effects are represented through conditional probabilities with quantification using sensitivity analysis. Hence, Bayesian causal maps are the combination of causal maps and BN. The limitations of causal maps are eliminated with the use of BN. Furthermore, qualitative representation of causal maps is transformed into quantitative representation with more precision. The main advantage of combining qualitative and quantitative representation of causal maps with support of BN is to have features (uncertain variables) with their numerical knowledge. Thus, Bayesian causal maps can extract knowledge, helping in comparative analysis while validating qualitative information in decision making. Consequently, Bayesian causal maps characterize abstract information as figurative facts. The figurative facts are obtained through *Netica* tool for making inferences.

C. Bayesian Networks Learning for Decision Making

A Bayesian causal map is constructed in [37] based on a case study of product development decision. The Bayesian causal map development procedure is carried out in two stages. First, qualitative stage gives details on follow-up interviews to deal with four modeling issues. However, original causal maps in competitive strategy lead to Sales Projection and Product Decision. In modified causal maps, Sales Projection or Product Decision does not have direct effect. Furthermore, variables effects with their conditional independence on each other can be seen in [37].

VII. BAYESIAN NETWORKS APPLICATIONS MISCELLANEOUS DOMAINS

In this section, additional but dissimilar BN applied domains are reviewed. The reviewed literature provides an overview on a variety of BN applied domains.

In ontology enhancement, Larik *et al.* [73] proposed a blend of ontology and BN approaches in order to classify various methods according to their application. Moreover, the strengths and weaknesses of these methods are identified comprehensively. The open source tool used for ontologies and BN, such as Protégé and UnBBayes, are utilized for validation purposes. Similarly, Matlab Toolbox [32] is used by Phillippot *et al.* [74] to construct BN. The aim of constructing BN structure was to classify the manually filled form fields by digital pen. For this purpose, Maximum Weight Spanning Tree (MWST) algorithm is more efficient, which diverted the focus from BN to MWST as well as more dependency on manual intervention. Milns *et al.* [75] and Bartlett *et al.* [76] presented their BN based ecological network and probabilistically cache analysis respectively. The prior method focused to analyze spatial data with massive data preparation requirement and the latter one automates cache use on program traces. Lakka *et al.* [77] used BN modeling approach for cross media analysis. The

objective is to enhance the efficient cross media by smooth adaptation of explicit knowledge. However, few media types such as text and images are ignored. An adaptive tutoring system is developed by Millan *et al.* [78] to endow with a background to education practitioners. The adaptive tutoring system as a student engineering model, target variable lost its generality with the use of knowledge feature.

VIII. DISCUSSION AND FUTURE DIRECTIONS

In this work, we reviewed BN-based techniques for image processing, medical domain, system reliability and estimation and PPDM, and in miscellaneous domains. We have summarized the literature on each domain and present the way of representing BN, drawing inferences and the learning of BN with an aim to quantify uncertainty issue. The study has also sought to know whether BN use in the aforementioned reviewed research fields lead us to improve results in challenging research area currently. In this perspective, we present application of BN in image processing for detecting text in scene images, system engineering with reliability estimation and PPDM, medical domain, decision making and miscellaneous BN-based approaches.

BN-based MRI, image segmentation, shape descriptor, semantic features, video de-interlacing, and flaws in weldments, traffic knowledge management and Signature verification techniques are presented. Besides these BN-based techniques in image processing, text localization is a major challenge for researchers to recognize text in scene images. In recent years, features of image descriptor regions for text detection have emerged as a powerful tool in matching and recognition [79], [80]. The most popular feature descriptors are scale-invariant feature transform (SIFT) [81] and maximally stable extremal region (MSER) [82], [83]. SIFT locates points in linear scale-space. However, SIFT is not affine-invariant [84]. MSER is an alternative to SIFT that extracts stable regions of an image which are connected and characterized by the uniform intensity. For this purpose, features of descriptor regions are extracted with the use of maximally stable extremal region (MSER) [82], [83]. MSER is an efficient method to extract features in distorted images for text localization. However, features of an image descriptor region extracted by MSER are limited. Based on above BN-based techniques, BN can be computationally inefficient. Therefore, a combination of features and BN of a descriptor region extracted through MSER can help in improving accuracy to localize text in scene images. The advantage of accuracy can help visually impaired persons with head mounted device on sign boards in streets [85]. In future, BN in conjunction with features of MSER-based extracted descriptor region can also be applied to detect text in scene image video sequences [86], postal address box localization [87], automatic forms reading [88], document image analysis [89].

Sensitivity analysis is an essential issue in system components which are working together to achieve a given target. To avoid abnormalities in a system, every component of a system and its operation is monitored to establish a

relationship to make a reliable decision. The effective sensitivity analysis of system components according to their operation can provide a support to estimate the overall behavior of a system. In future, sensitivity analysis and abnormality detection Minimum Spanning Resource Tree (MSRT) can be analyzed by BN. Similarly, association rules mined from a supermarket database using data mining techniques can disclose sensitive information to business competitors. To restrict sensitive information, supermarket database should be perturbed for BN-based identified sensitive attributes in a highest degree transaction. The advantage of perturbing supermarket database, in this way, is to keep the minimum effect to original database with the maximum benefit with no side effects.

Admitted patients in hospitals for treatment is also a sensitive research area to select the best medicine according to their clinical tests under an intensive care from nurses. In case of critical conditions, nurses look after by patrolling is not helpful to make a quick decision to choose the right medicine. The correct medicine under intensive care patient can be selected without any delay to wait for physicians. Usually, physicians make decisions on the historical data of a patient to give medicine. To make better decisions over time, BN can be incorporated to make predictions of medicine depending upon the current symptoms and historical data of a patient to eliminate delayed risks. Also, the best suitable clinical tests can be decided amongst a wide range of alternatives. Clinical tests are still a challenge especially to diagnose dental pain and exposure to X-ray radiation. As a result, a BN-based OPD tool is a critical need for better health care of people. Additionally, BN incorporation into Decision Support System (DSS) can automate law resources to predict the best possible decisions depending upon problem and previous judgments to support lawyers.

BN can play an important role in decision making because of a significant involvement of uncertainty in this domain. The competitive businesses make decisions upon their own customer data increase profits by targeting new potential customers. Based on expenditure of customers, a major challenge is to select an appropriate attributes to predict new potential customers with minimum risk. Therefore, integration of BN into Decision Support Systems (DSS) can help in making decisions with multi-criteria decision making (MCDM) in a quick and consistent manner. In addition, water resources are diminishing and have been recycled in developed countries to fulfill the needs of their citizens. In future, water resource share management can be a challenging research problem in which every stakeholder is satisfied. BN incorporation to fair management of water resource share can involve all the active stakeholders. Currently, EU water framework is a challenging research problem. Consequently, all stakeholders around the world can lead us to save, recycle and distribute equal share of water based on EU water framework. BN represents a flexible framework that can be extended to combining domain knowledge from spatial data. BN can be made more effective by adding the temporal factor to analyze and assess earthquake hazards with accurate prediction. In addition, natural disasters such as flood, typhoon and landslide with risks involvement can be assessed based on historical data. Besides this, BN-based student engineering model is reviewed that support people in academia. Thus, BN allows

determining the perception style of a student to discover his learning style over internet-based courses. To discover learning style of a student, BN can be applied over the features and semantic features to assess students' knowledge. Finally, accuracy and scalability of BN-based cross media analysis can be a major concern in future works.

IX. CONCLUSION

Bayesian Network is integrated in a number of practical researched domains with the existence of uncertainty. In this paper, we have not only summarized the BN integrated methodologies in domains such as Image Processing, Decision Making, System Reliability Analysis and PPDM, Medicine field and further supplementary research fields, but also presented the critical literature by stressing their strengths and limitations. In addition, we have also explained the BN representation, BN inferences and its learning in respective domains. In this way, importance of BN in handling uncertainty is highlighted. Consequently, review of BN-based methods is supportive in numerous real-world problems, especially in uncertain conditions. In this perspective, BN-based approaches have logical justification for quantification of uncertainty and innovation of improved knowledge in the considered domains. However, quantification of uncertainty in several domains with BN has put forward more challenges to deal with. Therefore, we have discussed future directions to support new researchers in their research fields. Hence, the significance of this work is clear in terms of its critically evaluated literature in various uncertain domains based on BN, the way of BN learning in respective domains specifically as well as generally and directing the academia to meet further challenges in research.

REFERENCES

- [1] S. O. Funtowicz and J. R. Ravetz, *Uncertainty and Quality in Science for Policy*, vol. 15, Springer, 1990.
- [2] M. B. van Asselt and J. Rotmans, "Uncertainty in integrated assessment modelling," *Climatic Change*, vol. 54, pp. 75-105, 2002.
- [3] E. Castillo *et al.*, "Sensitivity analysis in discrete Bayesian networks," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 27, pp. 412-423, 1997.
- [4] F. V. Jensen, *An Introduction to Bayesian Networks*, vol. 210, London: UCL Press, 1996.
- [5] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.
- [6] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2009.
- [7] R. Chen *et al.*, "Dynamic Bayesian network modeling for longitudinal brain morphometry," *Neuro Image*, vol. 59, pp. 2330-2338, 2012.
- [8] C. Hongjun and Z. Junfeng, "Classification of image based on semantic features and Bayesian networks," in *Proc. International Conference on Consumer Electronics, Communications and Networks (CECNet)*, 2011, pp. 4858-4861.
- [9] V. R. Rathod and R. Anand, "Comparison of different segmentation techniques in detection of flaws in weldments," in *Proc. The 2nd International Conference on Computer and Automation Engineering (ICCAE)*, 2010, pp. 673-677.
- [10] K. Zhang and M. A. Taylor, "Effective arterial road incident detection: a Bayesian network based algorithm," *Transportation Research Part C: Emerging Technologies*, vol. 14, pp. 403-417, 2006.
- [11] G. Jeon *et al.*, "Application of Bayesian belief network in reliable analysis for video deinterlacing," *IEEE Transactions on Consumer Electronics*, vol. 54, pp. 123-130, 2008.
- [12] S. Barrat and S. Tabbone, "A Bayesian network for combining descriptors: application to symbol recognition," *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 13, pp. 65-75, 2010.

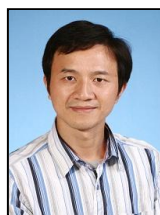
- [13] M. Velikova *et al.*, "A probabilistic framework for image information fusion with an application to mammographic analysis," *Medical Image Analysis*, vol. 16, pp. 865-875, 2012.
- [14] S. Wachsmuth and G. Sagerer, "Bayesian networks for speech and image integration," *AAAI/IAAI*, pp. 300-306, 2002.
- [15] X. Xiao and G. Leedham, "Signature verification using a modified Bayesian network," *Pattern Recognition*, vol. 35, pp. 983-995, 2002.
- [16] J. Luo *et al.*, "A Bayesian network-based framework for semantic image understanding," *Pattern Recognition*, vol. 38, pp. 919-934, 2005.
- [17] N. Serrano *et al.*, "Improved scene classification using efficient low-level features and semantic cues," *Pattern Recognition*, vol. 37, pp. 1773-1784, 2004.
- [18] I. N. Junejo, "Using dynamic Bayesian network for scene modeling and anomaly detection," *Signal, Image and Video Processing*, vol. 4, pp. 1-10, 2010.
- [19] O. Doguc and W. Jiang, "A Bayesian Network (BN) Model for System Operational Effectiveness Assessment & Diagnosis," in *Proc. 26th ASEM National Conference Proceedings*, 2005.
- [20] R. Wright and Z. Yang, "Privacy-preserving Bayesian network structure computation on distributed heterogeneous data," in *Proc. the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2004, pp. 713-718.
- [21] Y. Luo *et al.*, "A survey on the privacy preserving algorithm of association rule mining," in *Proc. ISECS'09. Second International Symposium on Electronic Commerce and Security*, 2009, pp. 241-245.
- [22] S. Samet and A. Miri, "Privacy-preserving bayesian network for horizontally partitioned data," in *Proc. International Conference on Computational Science and Engineering*, 2009, pp. 9-16.
- [23] O. Doguc and J. E. Ramirez-Marquez, "A generic method for estimating system reliability using Bayesian networks," *Reliability Engineering & System Safety*, vol. 94, pp. 542-550, 2009.
- [24] K. Iqbal *et al.*, "Hiding Sensitive XML Association rules with supervised learning technique," *Intelligent Information Management*, vol. 3, pp. 219-229, 2011.
- [25] S.-M. Lee and P. A. Abbott, "Bayesian networks for knowledge discovery in large datasets: basics for nurse researchers," *Journal of Biomedical Informatics*, vol. 36, pp. 389-399, 2003.
- [26] S. Aoki *et al.*, "Nurse call data analysis using Bayesian network modeling," in *Proc. Aware Computing (ISAC), 2010 2nd International Symposium on*, 2010, pp. 272-277.
- [27] Y. Sun *et al.*, "Diagnose the mild cognitive impairment by constructing Bayesian network with missing data," *Expert Systems with Applications*, vol. 38, pp. 442-449, 2011.
- [28] D. L. Fuente *et al.*, "Interconnection between biological abnormalities in borderline personality disorder: Use of the Bayesian networks model," *Psychiatry Research*, vol. 186, pp. 315-319, 2011.
- [29] S. Chattopadhyay *et al.*, "Application of Bayesian classifier for the diagnosis of dental pain," *Journal of Medical Systems*, vol. 36, pp. 1425-1439, 2012.
- [30] T. Bruland *et al.*, "Architectures integrating case-based reasoning and bayesian networks for clinical decision support," *Intelligent Information Processing V*, Springer, 2010, pp. 82-91.
- [31] N. Fenton and M. Neil, "Comparing risks of alternative medical diagnosis using Bayesian arguments," *Journal of Biomedical Informatics*, vol. 43, pp. 485-495, 2010.
- [32] K. Murphy, "The bayes net toolbox for matlab," *Computing Science and Statistics*, vol. 33, pp. 1024-1034, 2001.
- [33] P. C. Pendharkar *et al.*, "A probabilistic model for predicting software development effort," *IEEE Transactions on Software Engineering*, vol. 31, pp. 615-624, 2005.
- [34] B. Baesens *et al.*, "Bayesian network classifiers for identifying the slope of the customer lifecycle of long-life customers," *European Journal of Operational Research*, vol. 156, pp. 508-523, 2004.
- [35] E. J. Laur á and P. J. Duchessi, "A Bayesian belief network for IT implementation decision support," *Decision Support Systems*, vol. 42, pp. 1573-1588, 2006.
- [36] W. Wathayu and Y. Peng, "A Bayesian network based framework for multi-criteria decision making," in *Proc. the 17th International Conference on Multiple Criteria Decision Analysis*, 2004, pp. 6-11.
- [37] S. Nadkarni and P. P. Shenoy, "A causal mapping approach to constructing Bayesian networks," *Decision Support Systems*, vol. 38, pp. 259-281, 2004.
- [38] P. Zorrilla *et al.*, "Evaluation of Bayesian networks in participatory water resources management, Upper Guadiana Basin, Spain," *Ecology and Society*, vol. 15, 2010.
- [39] J. Pearl, *Bayesian Networks: A Model CF 'Self-Activated Memory for Evidential Reasoning*, 1985.
- [40] T. Bayes, "An essay toward solving a problem in the doctrine of chances," *Philosophical Transactions of the Royal Society of London*, vol. 53, 1984.
- [41] R. E. Neapolitan, *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*, John Wiley & Sons, Inc., 1990.
- [42] I. Ben-Gal, *Bayesian Networks, Encyclopedia of Statistics in Quality and Reliability*, Wiley & Sons, 2007.
- [43] L. Zhang and Q. Ji, "A Bayesian network model for automatic and interactive image segmentation," *IEEE Transactions on Image Processing*, vol. 20, pp. 2582-2593, 2011.
- [44] X.-H. Wang *et al.*, "Computer-assisted diagnosis of breast cancer using a data-driven Bayesian belief network," *International Journal of Medical Informatics*, vol. 54, pp. 115-126, 1999.
- [45] H.-K. Kim *et al.*, "A modified Zernike moment shape descriptor invariant to translation, rotation and scale for similarity-based image retrieval," in *Proc. IEEE International Conference on Multimedia and Expo*, 2000, pp. 307-310.
- [46] D. M. Chickering, "Learning Bayesian networks is NP-complete," *Learning from Data*, Springer, 1996, pp. 121-130.
- [47] G. F. Cooper and E. Herskovits, "A Bayesian method for the induction of probabilistic networks from data," *Machine Learning*, vol. 9, pp. 309-347, 1992.
- [48] R. T. Cox, "Probability, frequency and reasonable expectation," *American Journal of Physics*, vol. 14, p. 1, 1946.
- [49] S. Luttrell, "An adaptive Bayesian network for low-level image processing," in *Proc. Third International Conference on Artificial Neural Networks*, 1993, pp. 61-65.
- [50] I. Gondra and F. I. Alam, "Learning spatial relations for object-specific segmentation using Bayesian network model," *Signal, Image and Video Processing*, pp. 1-10, 2012.
- [51] S. Qiu *et al.*, "A fusion of bayesian and fuzzy analysis for print faults diagnosis," *Computers and Their Applications*, pp. 229-232, 2001.
- [52] R. Agrawal and R. Srikant, "Privacy-preserving data mining," *ACM Sigmod Record*, vol. 29, pp. 439-450, 2000.
- [53] D. M. Chickering *et al.* (1994). Learning Bayesian networks is NP-hard. [Online]. Available: <http://research.microsoft.com/apps/pubs/default.aspx?id=69598>
- [54] M. Atallah, *et al.*, "Disclosure limitation of sensitive rules," in *Proc. Workshop on Knowledge and Data Engineering Exchange*, 1999, pp. 45-52.
- [55] M. Kantarcioglu and C. Clifton, "Privacy-preserving distributed mining of association rules on horizontally partitioned data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, pp. 1026-1037, 2004.
- [56] J. Vaidya and C. Clifton, "Privacy preserving association rule mining in vertically partitioned data," in *Proc. the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2002, pp. 639-644.
- [57] J. Vaidya and C. Clifton, "Privacy-preserving k-means clustering over vertically partitioned data," in *Proc. the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003, pp. 206-215.
- [58] J. Vaidya and C. Clifton, "Privacy preserving naive bayes classifier for vertically partitioned data," in *Proc. 2004 SIAM International Conference on Data Mining, Lake Buena Vista, Florida*, 2004, pp. 522-526.
- [59] M. Kantarcioglu and J. Vaidya, "Privacy preserving naive bayes classifier for horizontally partitioned data," in *Proc. IEEE ICDM Workshop on Privacy Preserving Data Mining*, 2003, pp. 3-9.
- [60] R. Canetti *et al.*, "Selective private function evaluation with applications to private statistics," *PODC*, 2001, pp. 293-304.
- [61] W. Du and M. J. Atallah, "Privacy-preserving cooperative statistical analysis," in *Proc. 17th Annual Computer Security Applications Conference*, 2001, pp. 102-110.
- [62] M. J. Freedman *et al.*, "Efficient private matching and set intersection," *Advances in Cryptology-EUROCRYPT 2004*, pp. 1-19, 2004.
- [63] R. Agrawal *et al.*, "Information sharing across private databases," in *Proc. ACM SIGMOD International Conference on Management of Data*, 2003, pp. 86-97.
- [64] M. J. Flores *et al.*, "Incorporating expert knowledge when learning Bayesian network structure: A medical case study," *Artificial Intelligence in Medicine*, vol. 53, pp. 181-204, 2011.
- [65] M. Ramoni and P. Sebastiani, "Learning Bayesian networks from incomplete databases," in *Proc. the Thirteenth Conference on Uncertainty in Artificial Intelligence*, 1997, pp. 401-408.
- [66] F. Jensen *et al.*, "Hugin-the tool for bayesian networks and influence diagrams," *Probabilistic Graphical Models*, pp. 212-221, 2002.
- [67] J. Cheng, *Belief Network (BN) Power Constructor*, 2001.

- [68] J. Cheng, *BN Powerpredictor*, 2000.
- [69] P. Spirtes *et al.*, *Causation, Prediction, and Search*, vol. 81, The MIT Press, 2000.
- [70] J. Jaworska *et al.*, "Towards optimization of chemical testing under REACH: A Bayesian network approach to Integrated Testing Strategies," *Regulatory Toxicology and Pharmacology*, vol. 57, pp. 157-167, 2010.
- [71] P. C. Jackson. (2013). Introduction to artificial intelligence. [Online]. Available: <http://www.DoverPublications.com>
- [72] P. H. Winston, *Artificial Intelligence*, Addison-Wesley, Reading, MA, 1984.
- [73] A. S. Larik and S. Haider, "Efforts to blend ontology with Bayesian networks: an overview," in *Proc. 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, 2010, pp. V2-598-V2-602.
- [74] E. Philippot *et al.*, "Learning algorithms of form structure for Bayesian networks," in *Proc. 17th IEEE International Conference on Image Processing (ICIP)*, 2010, pp. 2149-2152.
- [75] I. Milns *et al.*, "Revealing ecological networks using Bayesian network inference algorithms," *Ecology*, vol. 91, pp. 1892-1899, 2010.
- [76] M. Bartlett *et al.*, "Probabilistic Instruction Cache Analysis using Bayesian Networks," in *Proc. IEEE 17th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)*, 2011, pp. 233-242.
- [77] C. Lakka *et al.*, "A Bayesian network modeling approach for cross media analysis," *Signal Processing: Image Communication*, vol. 26, pp. 175-193, 2011.
- [78] E. Millán *et al.*, "Bayesian networks for student model engineering," *Computers & Education*, vol. 55, pp. 1663-1683, 2010.
- [79] J. Sivic and A. Zisserman, "Video Google: A text retrieval approach to object matching in videos," in *Proc. Ninth IEEE International Conference on Computer Vision*, 2003, pp. 1470-1477.
- [80] C. RonKimmel *et al.*, *Affine Invariant Interesting Descriptors*, 2009.
- [81] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004.
- [82] G. Yu and J.-M. Morel, "A fully affine invariant image comparison method," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 1597-1600.
- [83] J. Matas *et al.*, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and vision computing*, vol. 22, pp. 761-767, 2004.
- [84] K. I. Kim *et al.*, "Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1631-1639, 2003.
- [85] C. Merino-Gracia *et al.*, "A head-mounted device for recognizing text in natural scenes," *Camera-Based Document Analysis and Recognition*, Springer, pp. 29-41, 2012.
- [86] C. Wolf *et al.*, "Text localization, enhancement and binarization in multimedia documents," in *Proc. 16th International Conference on Pattern Recognition*, 2002, pp. 1037-1040.
- [87] P. W. Palumbo *et al.*, "Postal address block location in real time," *Computer*, vol. 25, pp. 34-42, 1992.
- [88] E. Kavallieratou *et al.*, "Handwritten text localization in skewed documents," in *Proc. 2001 International Conference on Image Processing*, 2001, pp. 1102-1105.
- [89] Y. Xiao and H. Yan, "Text region extraction in a document image based on the Delaunay tessellation," *Pattern Recognition*, vol. 36, pp. 799-809, 2003.



Khalid Iqbal received his PhD degree in applied computer technology from University of Science and Technology Beijing in 2014, the B.Sc. and MS(CS) degrees from University of the Punjab, Lahore and SZABIST Karachi respectively. He was awarded a fully funded scholarship by Chinese Scholarship Council for the entire duration of his PhD studies. He also won the excellent researcher from University of Science and Technology Beijing. He is currently an assistant professor in the Department of Computer Science, COMSATS Institute of Information Technology, Attock Campus. He has worked on Bayesian network application for privacy preserving of XML association rules and text localization in scene images. His research work has been published in several international conference proceedings and journals. His research interests include pattern recognition, machine learning and data mining. He is the recipient of the CSC scholarship and QCRI/Boeing Travel

grant. He is also the member of IACSIT.



Xu-Cheng Yin received the Ph.D. degree in pattern recognition and intelligent systems from Institute of Automation, Chinese Academy of Sciences, China, in 2006, the B.Sc and M.Sc. degrees both in computer science from University of Science and Technology Beijing, China, in 1999 and 2002, respectively. He is currently an associate professor in the Department of Computer Science and Technology, School of Computer and Communication Engineering at the University of Science and Technology Beijing, China. From July 2002 to July 2006, he was a R&D engineer at the Research Center, Hanwang Technology Co. Ltd., Beijing, China. From August 2006 to June 2008, he was a scientific researcher at the Information Technology Lab., Fujitsu Research and Development Center. From Jan 2013 to Jan 2014, he was a visiting researcher in Center for Intelligent Information Retrieval, University of Massachusetts Amherst, USA.

His research interests include pattern recognition, machine learning, information retrieval, and document analysis and recognition. He has published more than 40 research papers.



Hong-Wei Hao is currently a professor in the Institute of Automation, Chinese Academy of Sciences, China. He received the Ph.D. degree in pattern recognition and intelligent systems from Institute of Automation, Chinese Academy of Sciences, China, in 1997. From 1999 to 2011, he worked as an associate professor and then a professor at the University of Science and Technology Beijing, China. From 2002 to 2003, he was a visiting researcher in the Central Research Laboratory, Hitachi Ltd., Tokyo, Japan.

His research interests cover large-scale semantic computing theory and technology, large-scale machine learning theory, and intelligent massive information processing.



Qazi Mudassir Ilyas is currently working as an assistant professor at King Faisal University, Saudi Arabia. He earned a bachelor degree in agricultural engineering from University of Agriculture, Pakistan with a gold medal, master degree in computer science from the same university and PhD degree in information and communication engineering from Huazhong University of Science and Technology, China with Excellent Student Award. His research interests include semantic web, knowledge engineering, machine learning and human computer interaction. He is a member of various scientific bodies including IEEE, Internet Society, Higher Education Commission of Pakistan and Pakistan Engineering Council.



Hazrat Ali received his bachelor and master degrees in electrical engineering from University of Engineering and Technology Peshawar Pakistan, in 2009 and 2012 respectively. He is currently pursuing his PhD degree in School of Computer and Communication Engineering, University of Science and Technology Beijing, China. He has been awarded a fully funded scholarship by the Higher Education Commission of Pakistan for the entire duration of his PhD studies. From 2010 to 2012, he worked at the Department of Electrical Engineering, University of Engineering and Technology Peshawar Pakistan, where he taught Control Systems, Electronic Circuits to undergraduate students. He has worked on corpus development and speech recognition tasks for Urdu and Pashto languages. His research work has been published in several international conference proceedings and journals. His current research interests include speech and image processing, speech recognition, machine learning and communication systems. He has acted as a review committee co-chair for City Informatics Phd Symposium 2014, a reviewer for International Journal on AI Tools and as a sub-reviewer for AAAI Conference on Weblogs and Social Media. He is the recipient of the HEC Scholarship, the TERENA/CISCO Travel grant, QCRI/Boeing Travel grant and the Erasmus Mundus STRoNGTiES research grant.