ENCYCLOPEDIA WITH SEMANTIC COMPUTING AND ROBOTIC INTELLIGENCE
Vol. 1, No. 1 (2017) 1630014 (13 pages)

DOI: 10.1142/S2425038416300147

Bayesian networks: Theory, applications and sensitivity issues

Ron S. Kenett
KPA Ltd., Raanana, Israel

University of Turin, Turin, Italy
ron@kpa-group.com

Accepted 11 November 2016; Published 17 March 2017

This chapter is about an important tool in the data science workbench, Bayesian networks (BNs). Data science is about generating information from a given data set using applications of statistical methods. The quality of the information derived from data analysis is dependent on various dimensions, including the communication of results, the ability to translate results into actionable tasks and the capability to integrate various data sources [R. S. Kenett and G. Shmueli, On information quality, *J. R. Stat. Soc. A* **177**(1), 3 (2014).] This paper demonstrates, with three examples, how the application of BNs provides a high level of information quality. It expands the treatment of BNs as a statistical tool and provides a wider scope of statistical analysis that matches current trends in data science. For more examples on deriving high information quality with BNs see [R. S. Kenett and G. Shmueli, *Information Quality: The Potential of Data and Analytics to Generate Knowledge* (John Wiley and Sons, 2016), www.wiley.com/go/information_quality.] The three examples used in the chapter are complementary in scope. The first example is based on expert opinion assessments of risks in the operation of health care monitoring systems in a hospital environment. The second example is from the monitoring of an open source community and is a data rich application that combines expert opinion, social network analysis and continuous operational variables. The third example is totally data driven and is based on an extensive customer satisfaction survey of airline customers. The first section is an introduction to BNs, Sec. 2 provides a theoretical background on BN. Examples are provided in Sec. 3. Section 4 discusses sensitivity analysis of BNs, Sec. 5 lists a range of software applications implementing BNs. Section 6 concludes the chapter.

Keywords: Graphical models; Bayesian networks; expert opinion; big data; causality models.

1. Bayesian Networks Overview

Bayesian networks (BNs) implement a graphical model structure known as a directed acyclic graph (DAG) that is popular in statistics, machine learning and artificial intelligence. BN enable an effective representation and computation of the joint probability distribution (JPD) over a set of random variables.[40] The structure of a DAG is defined by two sets: the set of nodes and the set of directed arcs. The nodes represent random variables and are drawn as circles labeled by the variables names. The arcs represent links among the variables and are represented by arrows between nodes. In particular, an arc from node X_i to node X_j represents a relation between the corresponding variables. Thus, an arrow indicates that a value taken by variable X_j depends on the value taken by variable X_i. This property is used to reduce the number of parameters that are required to characterize the JPD of the variables. This reduction provides an efficient way to compute the posterior probabilities given the evidence present in the data.[4,22,32,41,44] In addition to the DAG structure, which is often considered as the "qualitative" part of the model, a BN includes "quantitative" parameters. These parameters are described by applying the Markov property, where the conditional probability distribution (CPD) at each node depends only on its parents. For discrete random variables, this conditional probability is represented by a table, listing the local probability that a child node takes on each of the feasible values — for each combination of values of its parents. The joint distribution of a collection of variables is determined uniquely by these local conditional probability tables (CPT). In learning the network structure, one can include *white lists* of forced causality links imposed by expert opinion and *black lists* of links that are not to be included in the network. For examples of BN application to study management efficiency, web site usability, operational risks, biotechnology, customer satisfaction surveys, healthcare systems and testing of web services see, respectively.[2,24–27,31,42] For examples of applications of BN to education, banking, forensic and official statistics see Refs. 10 and 34, 43, 46, 47. The next section provides theoretical details on how BNs are learned and what are their properties.

2. Theoretical Aspects of Bayesian Networks

2.1. *Parameter learning*

To fully specify a BN, and thus represent the JPDs, it is necessary to specify for each node X the probability distribution for X conditional upon X's parents. The distribution of X, conditional upon its parents, may have any form with or without constraints.

These conditional distributions include parameters which are often unknown and must be estimated from data, for example using maximum likelihood. Direct maximization of the likelihood (or of the posterior probability) is usually based on the expectation–maximization (EM) algorithm which alternates computing expected values of the unobserved variables conditional on observed data, with maximizing the complete likelihood assuming that previously computed expected values are correct. Under mild regularity conditions, this process converges to maximum likelihood (or maximum posterior) values of parameters.[19]

A Bayesian approach treats parameters as additional unobserved variables and computes a full posterior distribution over all nodes conditional upon observed data, and then integrates out the parameters. This, however, can be expensive and leads to large dimension models, and in practice classical parameter-setting approaches are more common.[37]

2.2. *Structure learning*

BNs can be specified by expert knowledge (using white lists and black lists) or learned from data, or in combinations of both. The parameters of the local distributions are learned from data, priors elicited from experts, or both. Learning the graph structure of a BN requires a scoring function and a search strategy. Common scoring functions include the posterior probability of the structure given the training data, the Bayesian information criteria (BIC) or Akaike information criteria (AIC). When fitting models, adding parameters increases the likelihood, which may result in over-fitting. Both BIC and AIC resolve this problem by introducing a penalty term for the number of parameters in the model with the penalty term being larger in BIC than in AIC. The time requirement of an exhaustive search, returning back a structure that maximizes the score, is super-exponential in the number of variables. A local search strategy makes incremental changes aimed at improving the score of the structure. A global search algorithm like Markov Chain Monte Carlo (MCMC) can avoid getting trapped in local minima. A partial list of structure learning algorithms includes Hill-Climbing with score functions BIC and AIC Grow-Shrink, Incremental Association, Fast Incremental Association, Interleaved Incremental association, hybrid algorithms and Phase Restricted Maximization. For more on BN structure learning, see Ref. 36.

2.3. *Causality and Bayesian networks*

Causality analysis has been studied from two main different points of view, the "probabilistic" view and the "mechanistic" view. Under the probabilistic view, the causal effect of an intervention is judged by comparing the evolution of the system when the intervention is present and when it is not present. The mechanistic point of view focuses on understanding the mechanisms determining how specific effects come about. The interventionist and mechanistic viewpoints are not mutually exclusive. For examples, when studying biological systems, scientists carry out experiments where they intervene on the system by adding a substance or by knocking out genes. However, the effect of a drug product on the human body cannot be decided only in the laboratory. A mechanistic understanding based on pharmacometrics models is a preliminary condition for determining if a certain medicinal treatment should be studied in order to elucidate biological mechanisms used to intervene and either prevent or cure a disease. The concept of potential outcomes is present in the work of randomized experiments by Fisher and Neyman in the 1920s and was extended by Rubin in the 1970s to non-randomized studies and different modes of inference.[35] In their work, causal effects are viewed as comparisons of potential outcomes, each corresponding to a level of the treatment and each observable, had the treatment taken on the corresponding level with at most one outcome actually observed, the one corresponding to the treatment level realized. In addition, the assignment mechanism needs to be explicitly defined as a probability model for how units receive the different treatment levels. With this perspective, a causal inference problem is viewed as a problem of missing data, where the assignment mechanism is explicitly modeled as a process for revealing the observed data. The assumptions on the assignment mechanism are crucial for identifying and deriving methods to estimate causal effects.[16]

Imai *et al.*[21] studied how to design randomized experiments to identify causal mechanisms. They study designs that are useful in situations where researchers can directly manipulate the intermediate variable that lies on the causal path from the treatment to the outcome. Such a variable is often referred to as a 'mediator'. Under the parallel design, each subject is randomly assigned to one of the two experiments. In one experiment, only the treatment variable is randomized whereas in the other, both the treatment and the mediator are randomized. Under the crossover design, each experimental unit is sequentially assigned to two experiments where the first assignment is conducted randomly and the subsequent assignment is determined without randomization on the basis of the treatment and mediator values in the previous experiment. They propose designs that permit the use of indirect and subtle manipulation. Under the parallel encouragement design, experimental subjects who are assigned to the second experiment are randomly encouraged to take (rather than assigned to) certain values of the mediator after the treatment has been randomized. Similarly, the crossover encouragement design employs randomized encouragement rather than the direct manipulation in the second experiment. These two designs generalize the classical parallel and crossover designs in clinical trials, allowing for imperfect manipulation, thus providing informative inferences about causal mechanisms by focusing on a subset of the population.

Causal Bayesian networks are BNs where the effect of any intervention can be defined by a 'do' operator that separates intervention from conditioning. The basic idea is that intervention breaks the influence of a confounder so that one can make a true causal assessment. The established counterfactual definitions of direct and indirect effects depend on an ability to manipulate mediators. A BN graphical representations, based on local independence graphs and dynamic path analysis, can be used to provide an overview of dynamic relations.[1] As an alternative approach, the econometric approach develops explicit models of outcomes, where the causes of effects are investigated and the mechanisms governing the choice of treatment are analyzed. In such investigations, counterfactuals are studied (Counterfactuals are possible outcomes in different hypothetical states of the world). The study of causality in studies of economic policies involves: (a) defining counterfactuals, (b) identifying causal models from idealized data of population distributions and (c) identifying causal models from actual data, where sampling variability is an issue.[20] Pearl developed BNs as the method of choice for reasoning in artificial intelligence and expert systems, replacing earlier *ad hoc* rule-based systems. His extensive work covers topics such as: causal calculus, counterfactuals, Do calculus, transportability, missingness graphs, causal mediation, graph mutilation and external validity.[38] In a heated head to head debate between probabilistic and mechanistic view, Pearl has taken strong standings against the probabilistic view, see for example the paper by Baker[3] and discussion by Pearl.[39] The work of Aalen *et al.*[1] and Imai *et al.*[21] show how these approaches can be used in complementary ways. For more examples of BN applications, see Fenton and Neil.[12–14] The next section provides three BN application examples.

3. Bayesian Network Case Studies

This section presents the applications of BNs to three diverse case studies. The first case study is based on an expert assessment of risks in monitoring of patients in a hospital, the second example is based on an analysis of an open source community and involves social network analysis (SNA) and data related to the development of software code. The third used case is derived from a large survey of customers of an airline company.

3.1. *Patient monitoring in a hospital*

Modern medical devices incorporate alarms that trigger visual and audio alerts. Alarm-related adverse incidents can lead to patient harm and represent a key patient safety issue. Clinical alarms handling procedures in the monitoring of hospitalized patients is a complex and critical task. This involves both addressing the response to alarms and the proper setting of control limits. Critical alarm hazards include (1) Failure of staff to be informed of a valid alarm on time and take appropriate action and (2) Alarm Fatigue, when staff is overwhelmed, distracted and desensitized. Prevention to avoid harm to patients requires scrutinizing how alarms are initiated, communicated and responded to.

To assess the conditions of alarm monitoring in a specific hospital, once can conduct a mapping of current alarms and their impact on patient safety using a methodology developed by ECRI.[11] This is performed using a spreadsheet documenting any or all alarm signals that a team identifies as potentially important (e.g., high priority) for management. Figure 1 is an example of such a worksheet. It is designed to assess the decision processes related to alarms by comparing potentially important signals in such a way that the most important become more obvious.

Care area	Alarm load	Obstacles to effective alarm communication or response	Alarm signal	Device/system	Medical opinion of alarm importance	Risk to patient from alarm malfunction or delayed caregiver response	Contribution to alarm load/fatigue
ICU	moderate	low	low alarm volume	physiologic monitor	high	high	low
ICU	moderate	low	bradycardia	physiologic monitor	high	high	low
ICU	moderate	low	tachicardia	physiologic monitor	moderate	moderate	moderate
ICU	high	low	leads off	physiologic monitor	high	moderate	moderate
ICU	moderate	low	low oxygen saturation	physiologic monitor	high	high	moderate
ICU	moderate	low	low BP	physiologic monitor	high	high	moderate
ICU	moderate	low	high BP	physiologic monitor	moderate	moderate	moderate
RECOVERY	high	low	low alarm volume	physiologic monitor	high	high	low
RECOVERY	high	low	bradycardia	physiologic monitor	high	moderate	high
RECOVERY	high	low	tachicardia	physiologic monitor	moderate	moderate	high
RECOVERY	high	low	leads off	physiologic monitor	high	moderate	high
RECOVERY	moderate	low	low oxygen saturation	physiologic monitor	high	high	moderate
RECOVERY	moderate	low	low BP	physiologic monitor	high	high	moderate
RECOVERY	moderate	low	high BP	physiologic monitor	moderate	moderate	moderate
Med-surg dep	high	high	low alarm volume	physiologic monitor	high	high	low
Med-surg dep	high	high	bradycardia	physiologic monitor	high	moderate	high
Med-surg dep	high	high	tachicardia	physiologic monitor	moderate	moderate	high
Med-surg dep	high	high	leads off	physiologic monitor	high	moderate	high
Med-surg dep	moderate	high	low oxygen saturation	physiologic monitor	high	high	high
Med-surg dep	moderate	high	low BP	physiologic monitor	high	high	high
Med-surg dep	moderate	high	high BP	physiologic monitor	moderate	moderate	high

Fig. 1. Spreadsheet for alarm monitoring risks.

Filling in the spreadsheet requires entering ratings derived from experts, other tools (e.g., Nursing Staff Survey) as well as from other sources (e.g., input from medical staff). The first column in Fig. 1, specifies the care area. In the figure, one can see information related to the intensive care unit (ICU), the recovery room and the surgery department. The second column reflects the alarm lad intensity. The third column shows us an assessment of obstacles to effective alarm communication or response. The next column lists the different types of alarm signals from a specific device. In Fig. 1, we see only the part related to a physiology monitor. The last three columns correspond to a medical opinion of alarm importance, the risk to patient from an alarm malfunction or delay, and the contribution of the alarm to load and fatigue of the medical staff.

We show next how data from such a clinical alarm hazard spreadsheet can be analyzed with a BN that links care area, device, alarm signal, risks to patients and load on staff.

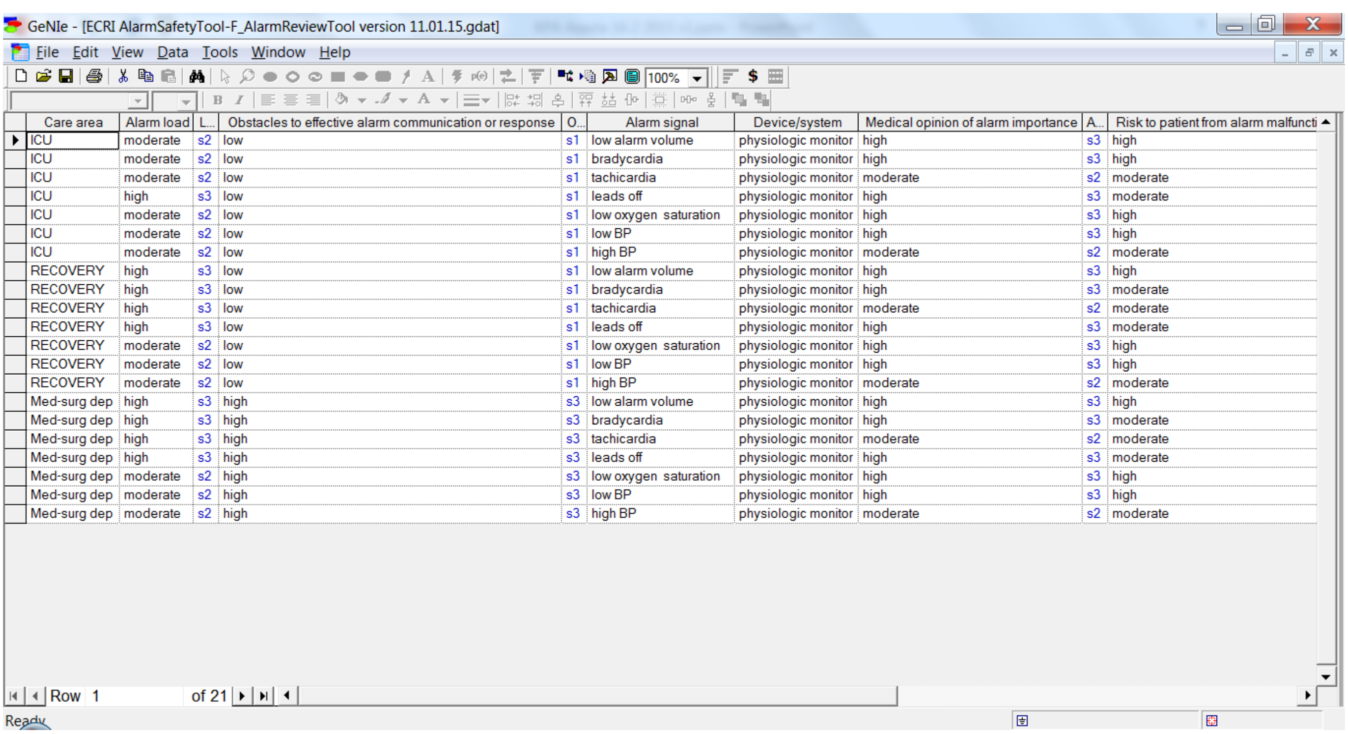
The software we will use in the analysis is GeNie version 2.0 from the university of Pittsburgh (http://genie.sis.pitt.edu). Figure 2 shows the GeNie data entry screen corresponding to the data shown in Fig. 1. The low, medium and high levels in Fig. 1 are represented here as s1, s2 and s3. Figure 3 shows the BNs derived from the data in Fig. 2.

Figure 4 shows a diagnostic analysis where we condition the BN on high level of fatigue. We can now see what conditions are the main contributor to this hazardous condition. Specifically, one can see that the percentage of surgery department increased from 33% to 42% and high load from 36% to 43%.

Figure 5 is an example of predictive analysis where the BN is conditioned on ICU and bradycardia. We see that the contribution of an alarm indicates a slower than normal heart rate (bradycardia) in ICU, if not treated properly, increases the potential high level harm to patients from 39% to 52%, a very dramatic increase. In parallel, we see that the impact of this condition on light load to staff increased from 30% to 49% and the option of low level of obstacles to communicate the alarm increased from 55% to 78%.

In other words, the bradycardia alarm in ICU does not increase fatigue, is very effectively communicated with a low level of obstruction and is considered higher than average alarm in terms if impact on patient safety. This analysis relies on expert opinion assessments and does not consider unobservable latent variables.

In this case study, the data was derived from expert opinion assessment of seven types of alarms from a patient monitoring device placed in three care delivery areas. The alarms are evaluated on various criteria such as contribution to staff fatigue, risk to patient from malfunction, obstacles to communicate alarms etc. This assessment is feeding a spreadsheet which is then analyzed using a BN. The analysis provides both diagnostic and predictive capabilities that supports various improvement initiatives aimed at improving monitoring effectiveness.


GeNIe - [ECRI AlarmSafetyTool-F_AlarmReviewTool version 11.01.15.gdat]

File Edit View Data Tools Window Help

Care area	Alarm load	L...	Obstacles to effective alarm communication or response	O...	Alarm signal	Device/system	Medical opinion of alarm importance	A...	Risk to patient from alarm malfuncti
ICU	moderate	s2	low	s1	low alarm volume	physiologic monitor	high	s3	high
ICU	moderate	s2	low	s1	bradycardia	physiologic monitor	high	s3	high
ICU	moderate	s2	low	s1	tachicardia	physiologic monitor	moderate	s2	moderate
ICU	high	s3	low	s1	leads off	physiologic monitor	high	s3	moderate
ICU	moderate	s2	low	s1	low oxygen saturation	physiologic monitor	high	s3	high
ICU	moderate	s2	low	s1	low BP	physiologic monitor	high	s3	high
ICU	moderate	s2	low	s1	high BP	physiologic monitor	moderate	s2	moderate
RECOVERY	high	s3	low	s1	low alarm volume	physiologic monitor	high	s3	high
RECOVERY	high	s3	low	s1	bradycardia	physiologic monitor	high	s3	moderate
RECOVERY	high	s3	low	s1	tachicardia	physiologic monitor	moderate	s2	moderate
RECOVERY	high	s3	low	s1	leads off	physiologic monitor	high	s3	moderate
RECOVERY	moderate	s2	low	s1	low oxygen saturation	physiologic monitor	high	s3	high
RECOVERY	moderate	s2	low	s1	low BP	physiologic monitor	high	s3	high
RECOVERY	moderate	s2	low	s1	high BP	physiologic monitor	moderate	s2	moderate
Med-surg dep	high	s3	high	s3	low alarm volume	physiologic monitor	high	s3	high
Med-surg dep	high	s3	high	s3	bradycardia	physiologic monitor	high	s3	moderate
Med-surg dep	high	s3	high	s3	tachicardia	physiologic monitor	moderate	s2	moderate
Med-surg dep	high	s3	high	s3	leads off	physiologic monitor	high	s3	moderate
Med-surg dep	moderate	s2	high	s3	low oxygen saturation	physiologic monitor	high	s3	high
Med-surg dep	moderate	s2	high	s3	low BP	physiologic monitor	high	s3	high
Med-surg dep	moderate	s2	high	s3	high BP	physiologic monitor	moderate	s2	moderate

Row 1 of 21

Ready

Fig. 2. Data entry of GeNie software Version 2.0.

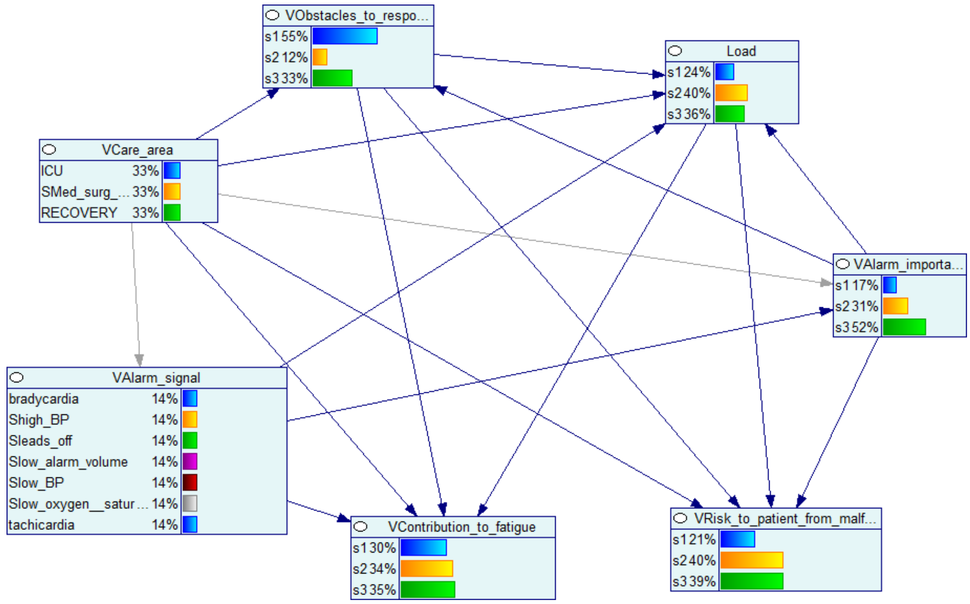


Fig. 3. BN of data collected in spreadsheet presented in Fig. 1.

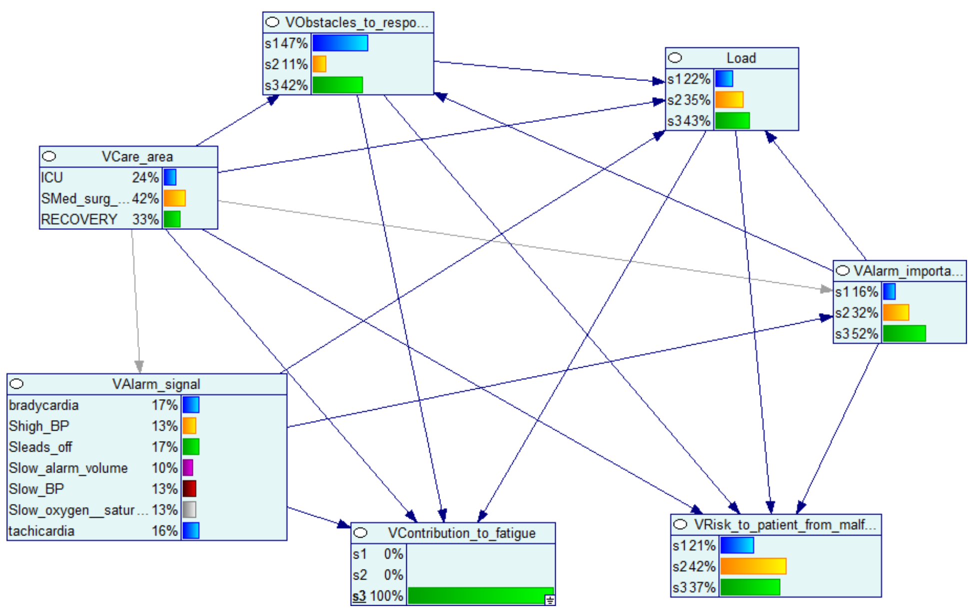


Fig. 4. BN conditioned on high contributors to fatigue (diagnostic analysis).

3.2. *Risk management of open source software*

The second example is based on data collected from a community developing open source software (OSS). Risk management is a necessary and challenging task for organizations that adopt OSS in their products and in their software development process.[33] Risk management in OSS adoption can benefit from data that is available publicly about the adopted OSS components, as well as data that describes the behavior of OSS communities. This use case is derived from the RISCOSS project (www.riscoss.eu), a platform and related assessment methodology for managing risks in OSS adoption.[15]

As a specific example, we aim at understanding the roles of various members in the OSS community and the relationships between them by the analysis of the mailing lists or forums of the community. The analysis should provide us

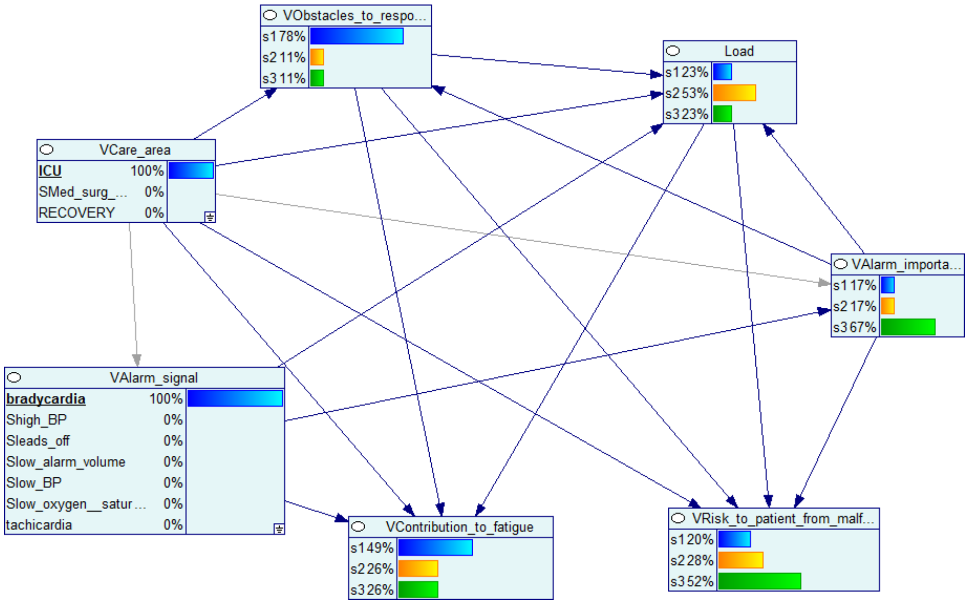


Fig. 5. BN conditioned on ICU and bradycardia (predictive analysis).

information on dimensions such as timeliness of an OSS community that can be measured by its capacity of following a roadmap or to release fixes and evolutions of the software in time. The methods illustrated here are based on data coming from the XWiki OSS community (http://www.xwiki.org), an Open Source platform for developing collaborative applications and managing knowledge using the wiki metaphor. XWiki was originally written in 2003 and released at the beginning of 2004; since then, a growing community of users and contributors started to gather around it. The data consists of: user and developer mailing lists archives, IRC chat archives, code commits and code review comments, and information about bugs and releases. The community is around 650,000 lines of code, around 95 contributors responsible for around 29,000 commits and with more than 200.000 messages, and 10.000 issues reported since 2004.

Specifically, we apply a SNA of the interactions between members of the OSS community. Highlighting actors of importance to the network is a common task of SNA. Centrality measures are ways of representing this importance in a quantifiable way. A node (or actor)'s importance is considered to be the extent of the involvement in a network, and this can be measured in several ways. Centrality measures are usually applied to undirected networks, with indices for directed graphs termed prestige measures. The degree centrality is the simplest way to quantify a node's importance which is used to consider the number of nodes it is connected to, with high numbers interpreted to be of higher importance. Therefore, the degree of a node provides local information of its importance. For a review of SNA see Ref. 45.

As an example of an SNA, we analyze the data from XWiki community over five years using data preprocessing of the IRC chat archives so extracting the dynamics of the XWiki community over time. Some of the challenges included a chat format change towards the end of 2010 and ambiguous names of a unique user (e.g., Vincent, VincentM, Vinny, Vinz). Eventually names were fixed manually. Figure 6 represents the visual rendering of the dynamics, over time, of the community in terms of intensity and kind of relationships between the different groups of actors (mainly contributors and manager of the community), that can be captured by community metrics such as degree of centrality. The analysis has been performed using NodeXL (http://nodexl.codeplex.com) a tool and a set of operations for SNA. The NodeXL-Network Overview, Discovery and Exploration add-in for Excel adds network analysis and visualization features to the spreadsheet. The core of NodeXL is a special Excel workbook template that structures data for network analysis and visualization. Six main worksheets currently form the template. There are worksheets for "Edges", "Vertices", and "Images" in addition to worksheets for "Clusters," mappings of nodes to clusters ("Cluster Vertices"), and a global overview of the network's metrics. NodeXL workflow typically moves from data import through steps such as import data, clean the data, calculate graph metrics, create clusters, create sub-graph images, prepare edge lists, expand worksheet with graphing attributes, and show graph such as the graphs in Fig. 6.

Each of the social networks presented in Fig. 6 is characterized by a range of measures such as the degree of centrality of the various community groups. The dynamics of a social network is reflected by a changing value of such measures, over time. We call these, and other measures derived from the OSS community, risk drivers. These risk

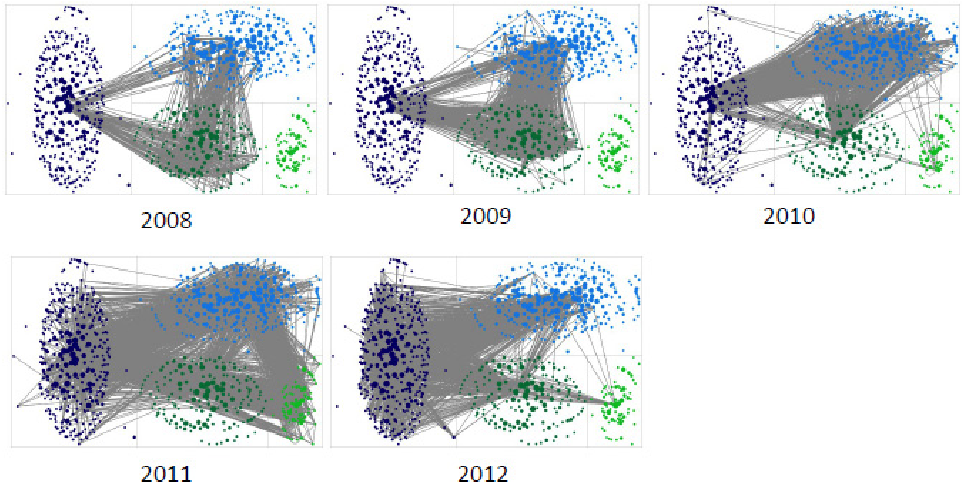


Fig. 6. SNA of the XWiki OSS community chats.

drivers form the raw data used in the risk assessment, in this case of adopting the XWiki OSS. This data can be aggregated continuously using specialized data collectors.

The data sources used in this analysis consisted of:

(1) Mailing lists archives:
  - XWiki users mailing list: http://lists.xwiki.org/pipermail/users
  - XWiki devs mailing list: http://lists.xwiki.org/pipermail/devs
(2) IRC chat archives: http://dev.xwiki.org/xwiki/bin/view/IRC/WebHome
(3) Commits (via git): https://github.com/xwiki
(4) Code review comments available on GitHub
(5) Everything about bugs and releases: http://jira.xwiki.org

From this data and SNA measures, one derives risk drivers that are determining risk indicators. Examples of risk indicators include Timeliness (is the community responsive to open issues), Activeness (what is the profile and number of active members) and Forking (is the community likely to split). These risk indicators are representing an interpretation of risk drivers values by OSS experts. To link risk drivers to risk indicators, the RISCOSS project developed a methodology based on workshops where experts are asked to rate alternative scenarios. An example of such scenarios is presented in Fig. 7.

The risk drivers are set at various levels and the expert is asked to evaluate the scenario in terms of the risk indicator. As an example, in scenario 30 of Fig. 7, the scenario consists of two forums posts per day, 11 messages per thread, a low amount of daily mails and a small community with a high proportion of developers, a med size number of testers and companies using the OSS. These conditions were determined as signifying low Activeness by the expert who did the rating. After running about 50 such scenarios, one obtains data that can be analyzed with BN, like in the alarm monitoring use case. Such an analysis is shown in Fig. 8.

The BN linking of risk drivers to risk indicators is providing a framework for ongoing risk management of the OSS community. It provides a tool for exploring the impact of specific behavior of the OSS community and evaluate risk mitigation strategies. For more on such models, see Ref. 15.

A	B	C	D	E	F	AI	AJ	AK	AL	AM	AN
Risk Driver	**State 1**	**State 2**	**State 3**	**State 4**	**State 5**	**29**	**30**	**31**	**32**	**33**	**34**
Forum posts per day	0	1	4	9	12	10	2	6	6	7	11
Forum messages per thread	0	1	4	9	19	14	11	9	19	16	4
Mail per day	low	medium	high			medium	low	low	high	medium	high
Overall community size	small	medium	high			low	low	low	medium	medium	medium
Number of developers involved	small	medium	high			high	high	high	high	high	medium
Number of testers (individuals providing feedback)	small	medium	high			high	medium	high	high	low	medium
Number of companies using the software	small	medium	high			low	medium	medium	medium	medium	low
Companies supporting the project (adding to code)	small	medium	high			high	high	low	high	low	medium
Activeness	1	2	3	4	5	4	1	3	4	4	5

Fig. 7. Alternative scenarios for linking risk drivers to the Activeness risk indicator.

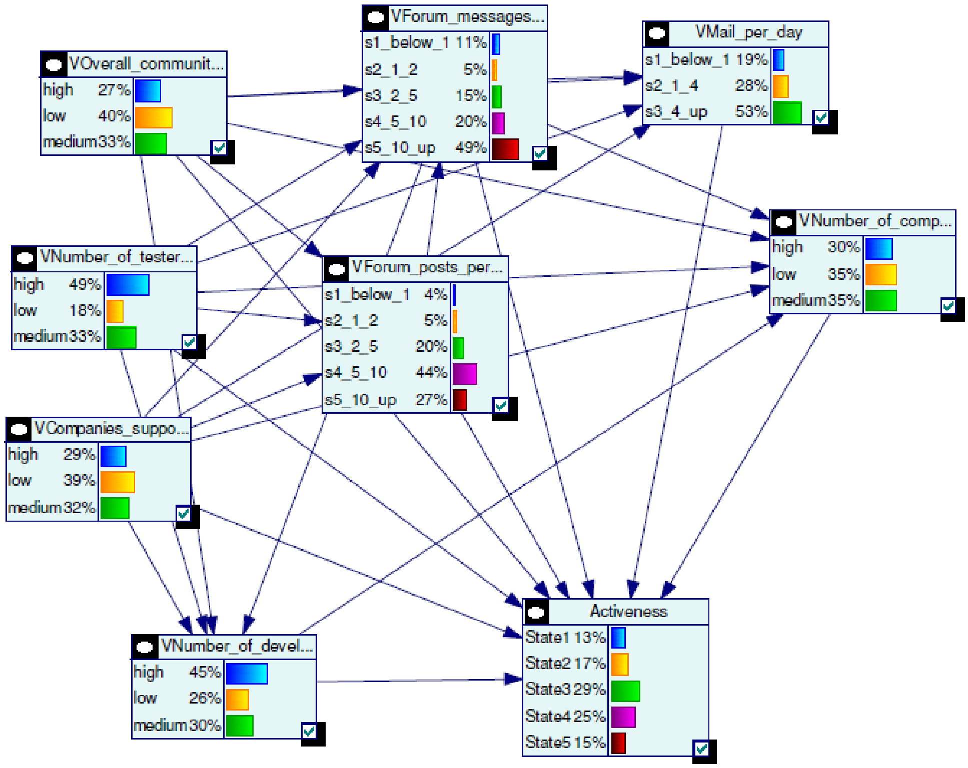


Fig. 8. BN linking risk drivers to risk indicators.

3.3. *Satisfaction survey from an airline company*

The third case study is about a customer satisfaction survey presented in Ref. 8. The example consists of a typical customer satisfaction questionnaire directed at passengers of an airline company to evaluate their experience. The questionnaire contains questions (items) on passengers' satisfaction from their overall experience and from six service elements (departure, booking, check-in, cabin environment, cabin crew, meal). The evaluation of each item is based on a four-point scale (from 1= extremely dissatisfied to 4 = extremely satisfied). Additional information on passengers was also collected such as gender, age, nationality and the purpose of the trip. The data consists of responses in $n = 9720$ valid questionnaires. The goal of the analysis is to evaluate these six dimensions and the level of satisfaction from the overall experience, taking into account the interdependencies between the degree of satisfaction from different aspects of the service. Clearly, these cannot be assumed to be independent of each other, and therefore a BN analysis presents a particularly well-suited tool for this kind of analysis. For more examples of BN applications to customer survey analysis see Refs. 5 and 27. Figure 9 is a BN of this data constructed using the Hill-Climbing algorithm with score functions AIC. The proportion of customers expressing very high overall satisfaction (a rating of "4") is 33%.

In Fig. 10, one sees the conditioning of the BN on extremely dissatisfied customer (left) and extremely satisfied customers (right).

The examples in Fig. 10 show how a BN provides an efficient profiling of very satisfied and very unsatisfied customers. This diagnostic information is a crucial input to initiatives aimed at improving customer satisfaction. Contrasting very satisfied with very unsatisfied customers is typically an effective way to generate insights for achieving this goal. Like in the previous two examples, the BN considered here is focused on observable data. This can be expanded by including modeling and latent variable effects. For an example of latent variables in the analysis of a customer satisfaction survey, see Ref. 17. For more considerations on how to determine causality, see Sec. 2.

4. Sensitivity Analysis of Bayesians Networks

The three examples presented above shop how a BN can be used as a decision support tool for determining which predictor variables are important on the basis of their effect on target variables. In such an analysis, choosing an adequate BN structure is a critical task. In practice, there are several algorithms available for determining the BN structure, each on with its specific characteristics. For example, the

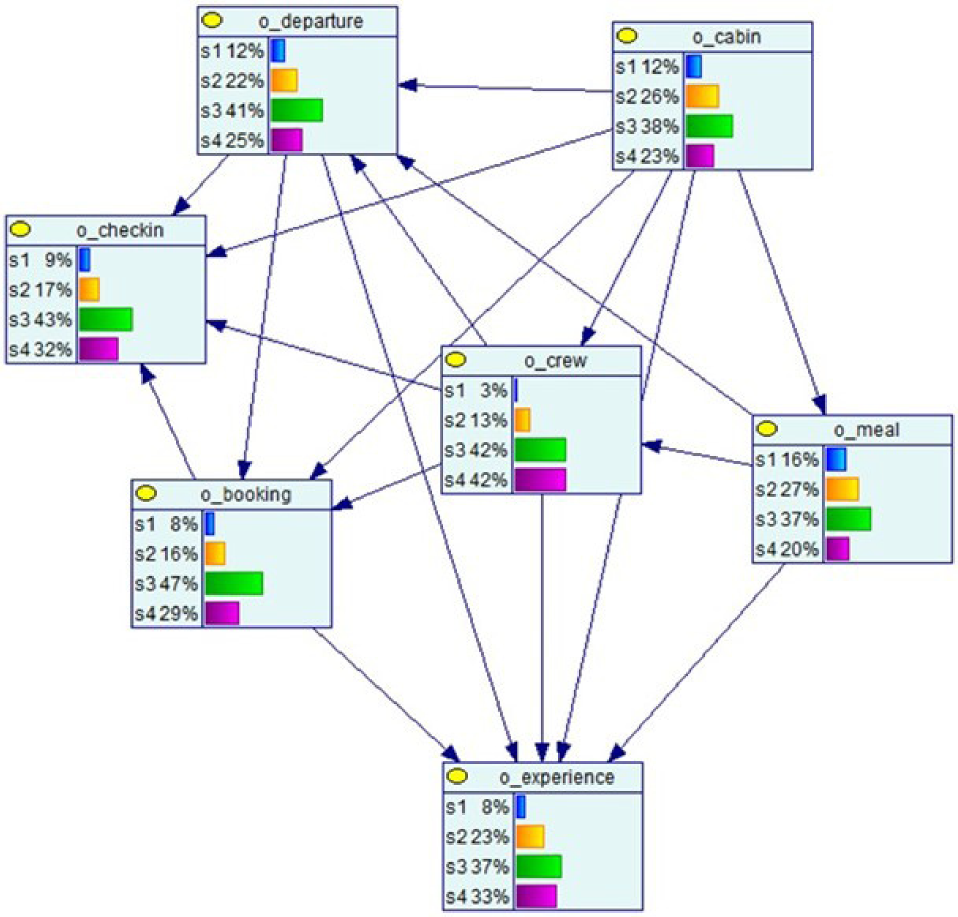


Fig. 9. BN of airline passenger customer satisfaction survey.

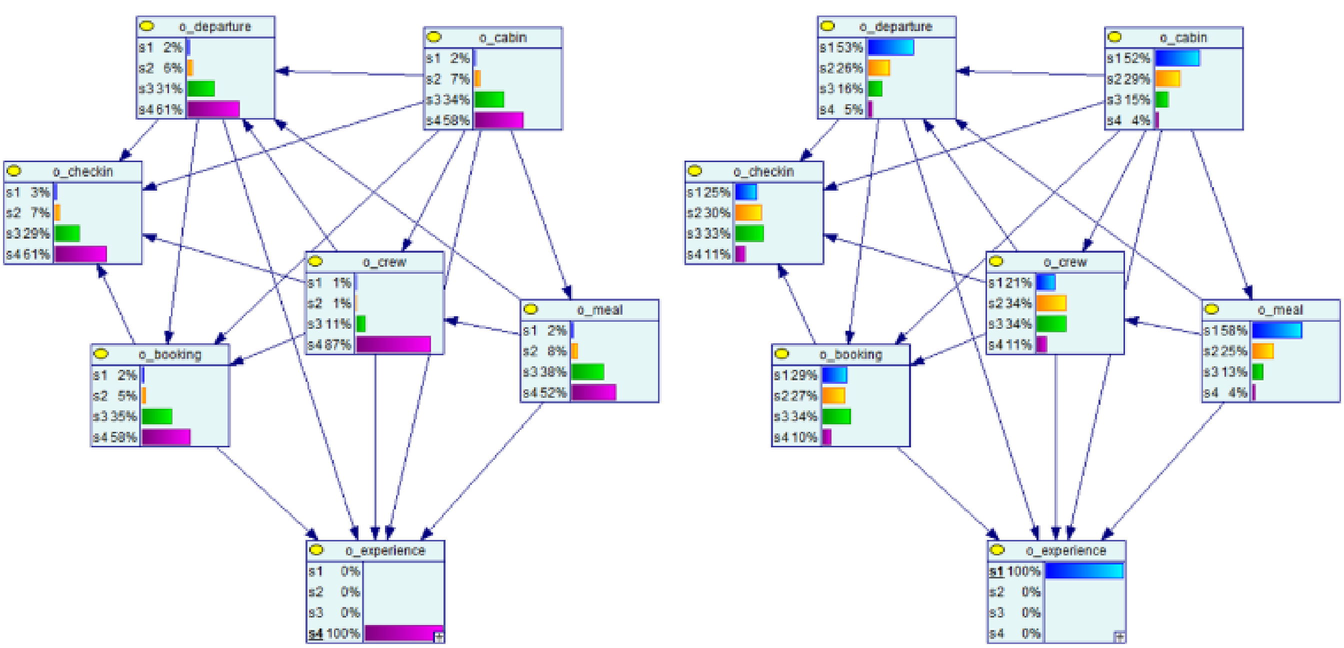


Fig. 10. BN conditioned on extremely dissatisfied customer (left) and extremely satisfied customers (right).

Table 1. BN with proportion of occurrence of each are in the bootstrap replicates.

	hc-bic	hc-aic	tabu-bic	tabu-aic	gs	iamb	fiamb	intamb	mmhc-bic	mmhc-aic	rsmax	tot
Booking Checkin	1.0	1.0	1.0	1.0	1.0	0.5	0.5	0.5	1.0	1.0	1.0	9.5
Cabin crew	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	6.0
Cabin departure	1.0	1.0	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	9.0
Cabin experience	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	0.0	9.0
Cabin meal	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	11.0
Crew booking	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	6.0
Crew departure	1.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	1.0	0.0	7.0
Crew experience	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	0.0	9.0
Departure booking	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	6.0
Departure checkin	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	1.0	1.0	0.0	7.0
Departure experience	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	0.0	9.0
Meal crew	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	7.0
Cabin booking	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	3.0
Crew checkin	0.0	1.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	4.0
Meal departure	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	3.0
Meal experience	0.0	1.0	0.0	1.0	0.0	1.0	1.0	1.0	0.0	1.0	0.0	6.0
Booking departure	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	0.0	1.0	5.0
Crew meal	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	4.0
Departure meal	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	2.0
Checkin booking	0.0	0.0	0.0	0.0	0.0	0.5	0.5	0.5	0.0	0.0	0.0	1.5
Checkin departure	0.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	4.0
Booking experience	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	2.0
Checkin experience	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	2.0
Checkin crew	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
Departure cabin	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0

R package *bnlearn* includes eleven algorithms: two-scored based learning algorithms (Hill-Climbing with score functions BIC and AIC and TABU with score functions BIC and AIC), five constraint-based learning algorithms (Grow-Shrink, Incremental Association, Fast Incremental Association, Interleaved Incremental association, Max-min Parents and Children), and two hybrid algorithms (MMHC with score functions BIC and AIC, Phase Restricted Maximization). In this section, we present an approach for performing a sensitivity analysis of BN, across various structure learning algorithms, in order to assess the robustness of the specific BN, one plans to use. Following the application of different learning algorithms to set up a BN structure, some arcs in the network are recurrently present and some are not. As a basis for designing a robust BN, one can compute how often an arc is present, across various algorithms, with respect to the total number of networks examined.

Table 1 shows the impact of the 11 learning algorithms implemented in the *bnlearn* R application. The last column represents the total number of arcs across the 11 algorithms. The robust structure is defined by arcs that appear in a majority of learned networks. For these variables, the link connection does not depend on the learning algorithm and the derived prediction and is therefore considered robust. For more on this topic and related sensitivity analysis issues, see Ref. 8.

After selection of a robust network, one can perform what-if sensitivity scenario analysis. These scenarios are computer experiments on a BN performed by conditioning on specific variable combinations and predicting the target variables using empirically estimated network. We can then analyze the effect of variable combinations on target distributions using the type of conditioning demonstrated in Sec. 4. The next section provides an annotated listing of various software products implementing BNs.

5. Software for Bayesian Network Applications

(i) Graphical Network Interface (GeNIe) is the graphical interface to Structural Modeling, Inference, and Learning Engine (SMILE), a fully portable Bayesian inference engine developed by the Decision Systems Laboratory of the University of Pittsburgh and thoroughly field tested since 1998. Up to version 2.0 GeNIe could be freely downloaded from http://genie.sis.pitt.edu with no restrictions on applications or otherwise. Version 2.1 is now available from http://www.bayesfusion.com/ with commercial and academic versions, user guides and related documentation.

(ii) Hugin (http://www.hugin.com) is a commercial software which provides a variety of products for both research and nonacademic use. The HUGIN Decision Engine (HDE) implements state-of-the-art algorithms for BNs and influence diagrams such as object-oriented modeling, learning from data with both continuous and discrete variables, value of information analysis, sensitivity analysis and data conflict analysis.

(iii) IBM SPSS Modeller (http://www-01.ibm.com/software/analytics/spss) is a general application for analytics that has incorporated the Hugin tool for running BNs (http://www.ibm.com/developerworks/library/wa-bayes1). IBM SPSS is not free software.

(iv) The R bnlearn package is powerful and free. Compared with other available BN software programs, it is able to perform both constrained-based and score-based methods. It implements five constraint-based learning algorithms (Grow-Shrink, Incremental Association, Fast Incremental Association, Interleaved Incremental association, Max–min Parents and Children), two scored-based learning algorithms (Hill-Climbing, TABU) and two hybrid algorithms (MMHC, Phase Restricted Maximization).

(v) Bayesia (http://www.bayesia.com) developed proprietary technology for BN analysis. In collaboration with research labs and big research projects, the company develops innovative technology solutions. Its products include (1) BayesiaLab, a BN publishing and automatic learning program which represents expert knowledge and allows one to find it among a mass of data, (2) Bayesia Market Simulator, a market simulation software package which can be used to compare the influence of a set of competing offers in relation to a defined population, (3) Bayesia Engines, a library of software components through which can integrate modeling and the use of BNs and (4) Bayesia Graph Layout Engine, a library of software components used to integrate the automatic position of graphs in specific application.

(vi) Inatas (www.inatas.com) provides the Inatas System Modeller software package for both research and commercial use. The software permits the generation of networks from data and/or expert knowledge. It also permits the generation of ensemble models and the introduction of decision theoretic elements for decision support or, through the use of a real time data feed API, system automation. A cloud-based service with GUI is in development.

(vii) SamIam (http://reasoning.cs.ucla.edu/samiam) is a comprehensive tool for modeling and reasoning with BNs, developed in Java by the Automated Reasoning Group at UCLA. Samiam includes two main components: a graphical user interface and a reasoning engine. The graphical interface lets users develop BN models and save them in a variety of formats. The reasoning engine supports many tasks including: classical inference; parameter estimation; time-space tradeoffs; sensitivity analysis; and explanation-generation based on MAP and MPE.

(viii) BNT (https://code.google.com/p/bnt) supports many types of CPDs (nodes), decision and utility nodes, static and dynamic BNs and many different inference algorithms and methods for parameter learning. The source code is extensively documented, object-oriented, and free, making it an excellent tool for teaching, research and rapid prototyping.

(ix) Agenarisk (http://www.agenarisk.com/) is able to handle continuous nodes without the need for static discretization. It enables decision-makers to measure and compare different risks in a way that is repeatable and auditable and is ideal for risk scenario planning.

6. Discussion

This chapter presents several examples of BNs in order to illustrate their wide range of relevance, from expert opinion-based data to big data applications. In all these cases, BNs has helped enhance the quality of information derived from an analysis of the available data sets. In Secs. 1 and 2, we describe various technical aspects of BNs, including estimation of distributions and algorithms for learning the BN structure. In learning the network structure, one can include *white lists* of forced causality links imposed by expert opinion and *black lists* of links that are not to be included in the network, again using inputs from content experts. This essential feature permits an effective dialogue with content experts who can impact the model used for data analysis. We also briefly discuss statistical inference of causality links, a very active area of research. In general, BNs provide a very effective descriptive causality analysis, with a natural graphical display. A comprehensive approach to BNs, with application sports, medicine and risks is provided by the Bayes knowledge project (http://bayes-knowledge.com/). Additional example of applications of BNs in the context of the quality of the generated information are included in Refs. 23 and 29.

References

[1] O. Aalen, K. Røysland and JM. Gran, Causality, mediation and time: A dynamic viewpoint, *J. R. Stat. Soc. A*, **175**(4), 831 (2012).
[2] X. Bai, R. S. Kenett and W. Yu, Risk Assessment and adaptive group testing of semantic web services, *Int. J. Softw. Eng. Knowl. Eng.* **22**(5), 595 (2012).
[3] S. Baker, Causal inference, probability theory, and graphical insights, *Stat. Med.* **2**(25), 4319 (2013).
[4] I. Ben Gal, Bayesian networks, in *Encyclopedia of Statistics in Quality and Reliability*, eds. F. Ruggeri, R. S. Kenett and F. Faltin (Wiley, UK, 2007).

[5] A. C. Constantinou, N. Fenton, W. Marsh and L. Radlinski, From complex questionnaire and interviewing data to intelligent Bayesian network models for medical decision support, *Artif. Intell. Med.* **67**, 75 (2016).

[6] G. F. Cooper, The computational complexity of probabilistic inference using Bayesian belief networks, *Artif. Intell.* **42**, 393 (1990).

[7] C. Cornalba, R. S. Kenett and P. Giudici, Sensitivity Analysis of Bayesian Networks with Stochastic Emulators, *ENBIS-DEINDE Proc.* University of Torino, Torino, Italy (2007).

[8] F. Cugnata, R. S. Kenett and S. Salini, Bayesian networks in survey data: Robustness and sensitivity issues, *J. Qual. Technol.* **48**, 253 (2016).

[9] L. DallaValle, Official statistics data integration using vines and non parametric. Bayesian belief nets, *Qual. Technol. Quant. Manage.* **11**(1), 111 (2014).

[10] M. Di Zio, G. Sacco, M. Scanu and P. Vicard, Multivariate techniques for imputation based on Bayesian networks, *Neural Netw. World* **4**, 303 (2005).

[11] ECRI, *The Alarm Safety Handbook Strategies, Tools, and Guidance* (ECRI Institute, Plymouth Meeting, Pennsylvania, USA, 2014).

[12] N. E. Fenton and M. Neil, The use of Bayes and causal modelling in decision making, uncertainty and risk, *UPGRADE, Eur. J. Inf. Prof. CEPIS (Council of European Professional Informatics Societies)*, **12**(5), 10 (2011).

[13] N. E. Fenton and M. Neil, *Risk Assessment and Decision Analysis with Bayesian Networks* (CRC Press, 2012), http://www.bayesianrisk.com.

[14] N. E. Fenton and M. Neil, Decision support software for probabilistic risk assessment using Bayesian networks, *IEEE Softw.* **31**(2), 21 (2014).

[15] X. Franch, R. S. Kenett, A. Susi, N. Galanis, R. Glott and F. Mancinelli, Community data for OSS adoption risk management, in *The Art and Science of Analyzing Software Data*, eds. C. Bird, T. Menzies and T. Zimmermann (Morgan Kaufmann, 2016).

[16] B. Frosini, Causality and causal models: A conceptual perspective, *Int. Stat. Rev.* **74**, 305 (2006).

[17] M. Gasparini, F. Pellerey and M. Proietti, Bayesian hierarchical models to analyze customer satisfaction data for quality improvement: A case study, *Appl. Stoch. Model Bus. Ind.* **28**, 571 (2012).

[18] A. Harel, R. S. Kenett and F. Ruggeri, Modeling web usability diagnostics on the basis of usage statistics, in *Statistical Methods in eCommerce Research*, eds. W. Jank and G. Shmueli (John Wiley & Sons, New Jersey, USA, 2009).

[19] D. Heckerman, A tutorial on learning with Bayesian networks, Microsoft research technical report MSR-TR-95-06, Revised November 1996, from http://research.microsoft.com.

[20] J. Heckman, Econometric causality. *Int. Stat. Rev.* **76**, 1 (2008).

[21] K. Imai, D. Tingley and T. Yamamoto, Experimental designs for identifying causal mechanisms, *J. R. Stat. Soc. A* **176**(1), 5 (2013).

[22] F. V. Jensen, *Bayesian Networks and Decision Graphs* (Springer, 2001).

[23] R. S. Kenett, On generating high infoQ with Bayesian networks, *Qual. Technol. Quant. Manag*. **13**(3), (2016), http://dx.doi.org/10.1080/16843703.2016.11891.

[24] R. Kenett, A. De Frenne, X Tort-Martorell and C. McCollin, The statistical efficiency conjecture, in *Applying Statistical Methods in Business and Industry — The State of the Art*, eds. T. Greenfield, S. Coleman and R. Montgomery (John Wiley & Sons, Chichester, UK, 2008).

[25] R. S. Kenett, Risk analysis in drug Manufacturing and Healthcare, in *Statistical Methods in Healthcare*, eds. F. Faltin, R. S. Kenett and F. Ruggeri (John Wiley & Sons, 2012).

[26] R. S. Kenett and Y. Raanan, *Operational Risk Management: A Practical Approach to Intelligent Data Analysis* (John Wiley & Sons, Chichester, UK, 2010).

[27] R. S. Kenett and S. Salini, *Modern Analysis of Customer Satisfaction Surveys: With Applications Using R* (John Wiley & Sons, Chichester, UK, 2011).

[28] R. S. Kenett and G. Shmueli, On information quality, *J. R. Stat. Soc. A* **177**(1), 3 (2014).

[29] R. S. Kenett and G. Shmueli, *Information Quality: The Potential of Data and Analytics to Generate Knowledge* (John Wiley & Sons, 2016). www.wiley.com/go/information_quality.

[30] R. S. Kenett and S. Zacks, *Modern Industrial Statistics: With Applications Using R, MINITAB and JMP*, 2nd edn (John Wiley & Sons, Chichester, UK, 2014).

[31] R. S. Kenett, A. Harel and F. Ruggeri, Controlling the usability of web services, *Int. J. Softw. Eng. Knowl. Eng.* **19**(5), 627 (2009).

[32] T. Koski and J. Noble, *Bayesian Networks — An Introduction* (John Wiley & Sons, Chichester, UK, 2009).

[33] J. Li, R. Conradi, O. Slyngstad, M. Torchiano, M. Morisio and C. Bunse, A state-of-the-practice survey of risk management in development with off-the-shelf software components, *IEEE Trans. Softw. Eng*. **34**(2), 271 (2008).

[34] D. Marella and P. Vicard, Object-oriented Bayesian networks for modelling the respondent measurement error, *Commun. Stat. — Theory Methods* **42**(19), 3463 (2013).

[35] F. Mealli, B. Pacini and D. B. Rubin, Statistical inference for causal effects, in *Modern Analysis of Customer Satisfaction Surveys: with Applications using R*, eds. R. S. Kenett and S. Salini (John Wiley and Sons, Chichester, UK, 2012).

[36] F. Musella, A PC algorithm variation for ordinal variables, *Comput. Stat.* **28**(6), 2749 (2013).

[37] E. R. Neapolitan, *Learning Bayesian Networks* (Prentice Hall, 2003).

[38] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, 1988).

[39] J. Pearl, Comment on causal inference, probability theory, and graphical insights (by Stuart G. Baker). UCLA Cognitive Systems Laboratory, *Stat. Med.* **32**(25), 4331 (2013).

[40] J. Pearl, Bayesian networks: A model of self-activated memory for evidential reasoning (UCLA Technical Report CSD-850017). *Proc. 7th Conf. Cognitive Science Society* (University of California, Irvine, CA), pp. 329–334.

[41] J. Pearl, *Causality: Models, Reasoning, and Inference*, 2nd edn. (Cambridge University Press, UK, 2009).

[42] J. Peterson and R. S. Kenett, Modelling opportunities for statisticians supporting quality by design efforts for pharmaceutical development and manufacturing, *Biopharmaceut. Rep.* **18**(2), 6 (2011).

[43] L. D. Pietro, R. G. Mugion, F. Musella, M. F. Renzi and P. Vicard, Reconciling internal and external performance in a holistic approach: A Bayesian network model in higher education, *Expert Syst. Appl.* **42**(5), 2691 (2015).

[44] O. Pourret, P. Naïm and B. Marcot, *Bayesian Networks: A Practical Guide to Applications* (John Wiley & Sons, Chichester, UK, 2008).

[45]M. Salter-Townshend, A. White, I. Gollini, T. B. Murphy, Review of statistical network analysis: Models, algorithms, and software, *Stat. Anal. Data Min.* **5**(4), 243 (2012).
[46]C. Tarantola, P. Vicard and I. Ntzoufras, Monitoring and improving Greek banking services Using Bayesian networks: An analysis of mystery shopping data, *Expert Syst. Appl.* **39**(11), 10103 (2012).
[47]P. Vicard, A. P. Dawid, J. Mortera and S. L. Lauritzen, Estimation of mutation rates from paternity casework, *Forensic Sci. Int. Genet.* **2**, 9 (2008).