Decision Support Systems 80 (2015) 42–55

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Causal inference for violence risk management and decision support in forensic psychiatry

Anthony Costa Constantinou [a,b,*], Mark Freestone [a,c], William Marsh [b], Jeremy Coid [a]

[a] Violence Prevention Research Unit, Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary, University of London, London EC1A 7BE, UK
[b] Risk and Information Management Research Group, School of Electronic Engineering and Computer Science, Queen Mary, University of London, London E1 4NS, UK
[c] Forensic Personality Disorder Service, East London NHS Foundation Trust, London, UK

ARTICLE INFO

Article history:
Received 31 October 2014
Received in revised form 11 July 2015
Accepted 28 September 2015
Available online 8 October 2015

Keywords:
Bayesian networks
Belief networks
Causal interventions
Criminology
Forensic psychiatry
Mental health

ABSTRACT

The purpose of medium secure services (MSS) is to provide accommodation, support, and treatment to individuals with enduring mental health problems who usually come into contact with the criminal justice system. These individuals are, therefore, believed to pose a risk of violence to themselves as well as to other individuals. Assessing and managing the risk of violence is considered to be a critical component for discharged decision making in MSS. Methods for violence risk assessment in this area of research are typically based on regression models or checklists with no statistical composition and which naturally demonstrate mediocre predictive performance and, more importantly, without providing genuine decision support. While Bayesian networks have become popular tools for decision support in the medical field over the last couple of decades, they have not been extensively studied in forensic psychiatry. In this paper, we describe a decision support system using Bayesian networks, which is mainly parameterised based on questionnaire, interviewing and clinical assessment data, for violence risk assessment and risk management in patients discharged from MSS. The results demonstrate moderate to significant improvements in forecasting capability. More importantly, we demonstrate how decision support is improved over the well-established approaches in this area of research, primarily by incorporating causal interventions and taking advantage of the model's ability in answering complex probabilistic queries for unobserved variables.

1. Introduction

Adequate management of offenders released from *medium secure services* (MSS) is crucial in preventing violent crime and ensuring efficient allocation of resources. To this end, many MSS in the UK and elsewhere make use of risk assessment tools in the ongoing management of patients. However, in this application domain, the current state of the art is represented by regression-based models and checklists with no statistical composition. Forensic medical practitioners have remained unimpressed by the current state of the art predictors, in terms of decision support in managing patients with serious mental illness problems, and have identified the need to examine new ways of modelling information [6].

Bayesian networks (BNs), which are probabilistic graphical models based on causal relationships, are especially well-suited for decision making scenarios that require as to consider multiple pieces of uncertain evidence. Over the last couple of decades, there has been a renewed interest in Bayesian inference, especially for real-world applications. This is because Bayesian inference, which used to be computationally intractable, now allows us to develop large-scale BN models using specialised software that takes advantage of efficient BN inference propagation algorithms [26,38].

Since then, successful applications of BNs for decision support have been witnessed in various application domains. These include

a) *Law and forensics*: [16] proposed the use of BNs as a tool for avoiding probabilistic fallacies in legal practice, which continue to occur despite that many of the fallacies have been well documented. [27] used BNs as part of a novel approach to triage for digital forensics for collecting and reusing past digital forensic investigation information in order to highlight likely evidential areas on a suspect operating system.
b) *Medical and biomedical informatics*: [54] presented a methodology for developing BN models that predict and reason with latent variables, in order to provide information that is useful for clinical decision makers, using a combination of expert knowledge and data, and in [55], the authors described a decision support BN system for assisting clinicians in making better decisions in Warfarin therapy management. Numerous other applications in biomedicine are

* Corresponding author at: Risk and Information Management Research Group, School of Electronic Engineering and Computer Science, Queen Mary, University of London, London E1 4NS, UK.
E-mail address: anthony@constantinou.info (A.C. Constantinou).

http://dx.doi.org/10.1016/j.dss.2015.09.006
0167-9236/

covered in [21,25,35,36].

c) *Safety*: Naderpour et al. [37] presented a decision support system to help with the management of abnormal situations in safety-critical environments and demonstrated, based on a case taken from US Chemical Safety reports, how the system provided support for operators in maintaining the risk of dynamic situations at acceptable levels. [40] proposed a BN for cascading crisis events, such as typhoons, rainstorms, and floods, that provides the capability to analyse the chain reaction path of such an event and potential losses, with experimental results indicating that this BN-based method improved forecasting accuracy compared to existing classical methods.

d) *Software development, project management, and information technology*: [32] developed a BN based on information technology implementations and demonstrated how the BN model can be incorporated into a decision support system to support *what-if* analysis. [28] demonstrated how a BN model, that was learned from data but which considered expert causality constraints, was able to perform better, in terms of predicting project management risks, than many previously proposed well-known algorithms and models. Yet et al. [56] proposed a dynamic BN modelling framework for calculating the costs and benefits of a project over a specified time period, allowing for changing circumstances and trade-offs.

e) *Sports prediction, betting, and psychology*: [7,8] demonstrated how an expert constructed BN model, that combined both data and expert knowledge, was able to outperform purely data-driven statistical models and generate profit against the gambling market. In sports psychology, Constantinou et al. [9] employed a BN to infer referee bias diagnostically by examining whether relevant causal factors during a football match could explain referee decisions.

Decision support benefits from the use of Bayesian models have also been reported in other more specialised applications. For instance, [52] proposed a hierarchical naive Bayes model that improves existing identity matching techniques in terms of searching effectiveness, and [53] developed a BN, as part of a framework for model integration and holistic modelling of socio-technical systems, and demonstrated decision support benefits based on an airport inbound passenger facilitation case study. [17] illustrate how BNs can be applied to model knowledge in many more diverse fields.

However, there have been limited previous attempts in developing decision support systems using BNs in forensic psychiatry, as well as from questionnaire and interviewing data in general. [45] acknowledged this by stating that BNs have been rarely used to analyse customer survey data. More specifically, these previous relevant attempts focused on analysing survey data for customer complaints and satisfaction [1,44,45] and for marketing purposes [29,46] also used survey data to extract general information from the British general household survey, which provides a continuous information on a range of social fields such as population, housing, education, employment, health, and income. All of these previous studies have reported a number of advantages in using BNs for analysing this kind of data. Most notably, these include that BNs a) offer a rich and descriptive overview of the broader customer behaviour by providing insights into determinants and subsequent behaviour, b) provide a causal explanation using observable variables within a single nonlinear multivariate model, c) provide the ability to conduct probabilistic inference for both prediction and diagnosis, and d) provide a graphical representation and outputs that be easily understood by professionals.

Despite the significant benefits demonstrated, BNs are still underexploited in forensic psychiatry. Therefore, it was felt that causal BNs could improve on the current state of the art. In [10] we presented the first BN model for preventing violent re-offence in released prisoners with serious history of violence. This paper is an extension of that study, but which focuses on mentally ill patients and provides decision support for discharged decision making from MSS. The paper is organised as follows: Section 2 describes the data and methodology, Section 3 describes model validation and discusses the results, and Section 4 provides our concluding remarks and directions for future work.

2. Data and methodology

We make use of a dataset referred to as VoRAMSS (*The Validation of New Risk Assessment Instruments for Use with Patients Discharged from Medium Secure Services*; [15]). The dataset consists of questionnaire, interviewing and assessment data from 386 patients, out of whom 343 are males and 43 are females. Interviews were performed at 6 and 12 months post-discharge. At 6 months post-discharge, the occurrences for general violence[†] and violent convictions are 13.73% and 2.33%, respectively, while at 12 months (i.e. between 6 and 12 months after release), the respective occurrence rates are 11.40% and 3.12%. The cumulative rates (i.e. 0–12 months) for general violence and violent convictions are 22.28% and 5.18%, respectively.

In addition to the VoRAMSS dataset mentioned above, we have also made use of a small part of a second dataset which is referred to as the Prisoner Cohort Study (PCS) [5]. This is because the PCS dataset provided information for a small number of model parameters that were considered important for decision analysis, but which the VoRAMSS dataset failed to capture (details in Section 2.2). However, the PCS dataset is somewhat different to the VoRAMSS dataset in the sense that it involves released prisoners, rather than patients discharged from MSS. However, many of those released prisoners also suffered from mental health problems (i.e. severe depression, anxiety, psychotic disorder). In an attempt to maintain relevant to the VoRAMSS dataset, we have restricted the cases considered by this second dataset to mentally ill individuals. The PCS dataset consists of questionnaire, interviewing and assessment data from 953 prisoners (before and after release), 778 males and 175 females, with a reconviction rate of 25.18% over a ~5 year period post-release. There were 594 cases of mentally ill individuals, and which were used to learn the causal relationships for the following model factors: a) anger management, b) drug misuse treatment, c) alcohol misuse treatment, d) cocaine dependence, e) cannabis dependence, f) stimulants dependence, and g) alcohol dependence.

All of these data that had been extracted from questionnaires, interviews, and assessments of patients with a specialist was then combined with relevant patient data retrieved by the Police National Computer (PNC), which mainly consisted of criminal records. As a result, we were presented with a set of unstructured patient data that had been collected independently of the requirements of a BN model, with a large number of variables consisting of repetitive, redundant, and in many cases, contradictory information. This meant that the initial format of the data was inappropriate for causal analysis. In order to make the data adequate for causal inference, we had to restructure the dataset.

Fig. 1 presents a diagram which demonstrates the practices we had to use in order to move from the unstructured data into a BN model capable of simulating interventions for risk management decisions, indicating that expert knowledge played a crucial role at every step of the process. Expert knowledge was provided by two clinically active experts in forensic psychiatry (JC) and forensic psychology (MF), each with at least 8 years' experience in forensic mental health research, having published widely on criminal justice outcomes [2,3,19], psychopathy and personality disorder [4,20], and mental illness [2]. We discuss these development stages in turn in the subsections that follow.

[†] The definition of *general violence* in this paper is identical to that of violence as defined in [15] for the VoRAMSS dataset, which includes sexual assaults, assaultive acts that involved the use of a weapon, or threats that are made with a weapon in hand as well as all the acts of battery, regardless of whether or not have resulted in injury.