Robust Bayesian Detection: A Case Study

Patrick de Oude
Informatics Institute
University of Amsterdam
the Netherlands
p.deoude@uva.nl

Gregor Pavlin
D-CIS Lab
Thales Nederland B.V.
the Netherlands
gregor.pavlin@d-cis.nl

Joris de Groot
Informatics Institute
University of Amsterdam
the Netherlands
grootjoris@gmail.com

Abstract – *This paper discusses the use of Bayesian networks in a class of contemporary gas detection/classification problems. In particular, we expose the properties of Bayesian networks which allow creation of detection systems with good performance despite significant deviations between the used models and the underlying true probability distributions. Key to adequate grounding of fusion processes is explicit representation of the locality of causal relations in models of monitoring processes. This provides guidance for a systematic and tractable construction of complex detection systems correlating very heterogeneous information. The resulting Bayesian detection systems are experimentally validated.*

1 Introduction

This paper introduces a robust approach to detection of annoying/harmful gases using large quantities of heterogeneous information originating from noisy sources. The observations can be obtained from sensors and humans using conventional communication infrastructure, such as mobile phone networks, Internet, etc. Interpretation of observed patterns requires domain models which provide a mapping between the observations and the hypotheses of interest. In order to be able to draw correct conclusions about hidden events based on observations, the domain models must adequately capture relations between the relevant variables; i.e., the models must be grounded in the real world. However, construction of adequate domain models can be very challenging. Namely, the heterogeneity and large numbers of information sources require domain models which capture complex relations between many related phenomena.

It turns out that probabilistic causal models facilitate design of robust fusion solutions required for the gas detection problems. Namely, in the targeted domains the observations can often be viewed as outcomes of causal stochastic processes. Such processes can be modeled with the help of causal Bayesian networks (BN) [11] which provide a theoretically rigorous and compact mapping between hidden events of interest and observable events.

However, the models are abstractions associated with significant uncertainties as it is often difficult or even impossible to construct models that perfectly capture processes in the real world. The designers inevitably make mistakes and usually there is not enough data to create perfect models through machine learning. In case of BNs, the deviations occur in parameters as well as in the structure. In other words, we are confronted with a grounding problem and the main question is: 'When do the domain models support good detection?'

In order to answer this we systematically investigate how parametric and structural inaccuracies in domain models influence the estimation processes and the expected classification performance. It turns out that the theory on Bayesian networks supports a systematic analysis of these problems and facilitates solutions to robust detection systems. Particularly important is the locality of causal relations in BNs and the corresponding factorization of probability distributions [11]. By considering the locality of causal relations, design rules supporting construction of inherently robust fusion systems have been derived [10]. In addition, the locality of relations in causal processes and the properties of BNs, justify simplifying assumptions, which make modeling and inference tractable without jeopardizing the performance of the resulting detection systems.

The challenges and solutions are illustrated with the help of a system which can detect the presence of annoying/harmful gases by using sensors and human reports.

In section 2 we provide a rationale for using causal Bayesian networks for the targeted problem. Section 3 discusses construction of domain models while section 4 focuses on the modeling deviations and properties of BNs which can be exploited for building of inherently robust detection systems. In section 5 experimental evidence is provided for the properties of the proposed detection systems.

2 Causal Probabilistic Models

Often monitoring processes can be viewed as causal stochastic processes, where hidden events cause observations according to certain probability distributions. Let's as-

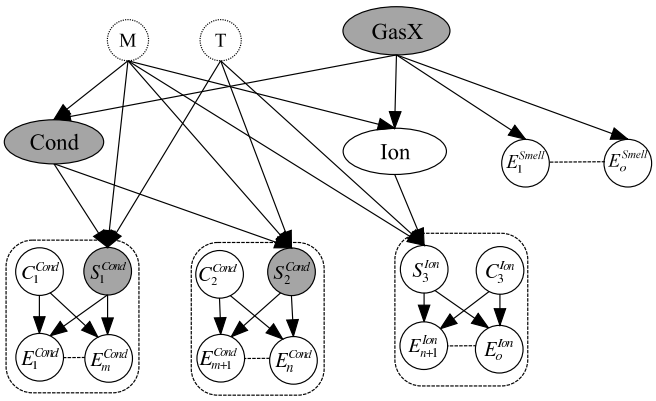


Figure 1: A causal model of a monitoring process capturing relations between the possible states in a particular time slice at a certain location, such as presence/absence of *GasX* and heterogeneous observations represented by leaf nodes E_i^x.

sume a system which detects the presence of a specific gas* at a certain location/area. The presence of the gas causes conditions under which certain type of semiconductors features distinctive conductivity. Introduction of a sensor measuring a particular type of conductivity will introduce a sequence of sensor reports, either confirming or refuting the presence of the gas. Similarly, humans are likely to report about typical symptoms (e.g. smell).

Such a causal process can be described through the graph shown in Figure 1, where each node represents a binary variable; e.g. $GasX = true$ if the gas anomaly is present, otherwise $GasX = false$. Situations under which a semiconductor element and an ionized gas mixture feature typical conductivity is captured by variables *Cond* and *Ion*, respectively. In particular, the states of binary variable *Cond* correspond to the situations† where electrical current under ideal circumstances would either exceed some detection threshold (i.e. $Cond = true$) or remain below that threshold (i.e. $Cond = false$). Variables S_1^{Cond} and S_2^{Cond} denote the measured conductivity on specific semiconductor elements in the first two sensors, while S_3^{Ion} denotes the measured conductivity of the ionized gas mixture in the third sensor. Moreover, binary variables $C_1^{Cond}, C_2^{Cond}, C_3^{Ion}$ represent the status of electronic components of the three sensors evaluating the measured conductivity ‡. If all electronic components of the $j-th$ sensor using a semiconductor element work correctly, then $C_j^{Cond} = ok$. Otherwise $C_j^{Cond} = fail$.

A sequence of reports produced by a sensor is denoted by binary variables $E_1^x, \ldots, E_n^x$, where x represents the information type. $E_i^x = true$ if a sensor report supports the hypothesis $GasX = true$. In this example, subgraphs containing nodes $S_1^{Cond}, C_1^{Cond}, E_1^{Cond}, \ldots, E_m^{Cond}$ and $S_2^{Cond}, C_2^{Cond}, E_{m+1}^{Cond}, \ldots, E_n^{Cond}$ describe processes in two sensors measuring the conductivity on semiconductor elements. The subgraph consisting of nodes S_3^{Ion}, C_3^{Ion} and $E_{n+1}^{Ion}, \ldots, E_o^{Ion}$, on the other hand, corresponds to the third sensor measuring conductivity of the ionized gas mixture. In addition, humans who have olfactory reactions to *GasX* can submit reports on what they smell via an automated call service or a web-interface. The states of binary variable *Smell* represent situations in which people familiar with a typical smell of *GasX* either do or do not recognize the smell. Moreover, each individual report is represented by a node E_i^{Smell}.

All nodes mentioned above correspond to variables whose states are outcomes of a causal stochastic process. We call these variables *Process variables*. In Figure 1 all non-root nodes and node *GasX* correspond to process variables. *Context variables*, on the other hand, represent phenomena that influence a causal process of interest while they can be considered independent of the same process. Consequently, in a causal model such variables are always represented through root nodes. Context variables represent boundary conditions of a causal process. In the example from Figure 1 the root nodes corresponding to variables *T* and *M* are context variables.

Note that in this case information fusion can be viewed as the estimation of distributions over the states of *process variables* which are not directly observable.

2.1 Bayesian Networks

Stochastic relations in causal processes can be described through conditional probabilities. Consequently, the presented observation generating processes can efficiently be modeled with the help of discrete Causal Bayesian networks (BN) [11]. A BN consists of a qualitative and quantitative part. The *qualitative part* is specified through a directed acyclic graph (DAG), where each variable in the domain is represented as a node§. The nodes are connected through directed edges, which represent direct dependencies between the represented variables. An example of a DAG can be found in Figure 1. The strength of stochastic relations between discrete variables is specified by conditional probability distributions over variables given their parents in the DAG. These distributions are captured by conditional probability tables (CPTs), which represent the *quantitative part* of a BN.

In principle, each BN encodes a joint probability distribution (JPD) $P(\mathbf{V})$ over a set of variables $\mathbf{V}$ through a product of conditional probabilities $\prod_i P(X_i|\pi(X_i))$, where $X_i \in \mathbf{V}$ correspond to a node in the DAG and $\pi(X_i)$ is the set of parents of X_i in the DAG. In other words, the factorization of the JPD $P(\mathbf{V})$ corresponds to the topology of the DAG.

The JPD factorization is exploited for efficient inference with Bayesian networks [6], i.e. computation of marginal

*For the sake of simplicity we say that the gas is present if its concentration exceeds some critical value. Otherwise we say that the gas is absent.

†a gas mixture in which partial pressure of *GasX* exceeds a threshold causing significant currents in a perfect semiconductor element

‡In this case each variable represents the global status of the entire electronic circuitry. We could also explicitly represent the states of each individual electronic component with interconnected variables of lower dimensionality, which would result in more incoming links to the variables representing reports.

§Terms node and variable are used interchangeably in this paper.

distributions over arbitrary variables in the model, given the observed states (i.e. evidence). In the presented detection system, the inference in BNs is used for the computation of the posterior probability $P(GasX|\mathcal{E})$, where $\mathcal{E}$ denotes a set of observations (i.e. instantiations of certain variables or an observation pattern). If $P(GasX = true|\mathcal{E}) > P(GasX = false|\mathcal{E})$ the detection system assumes that *GasX* is present. Note that the same principle can be used for the classification of gases, where each state of the hypothesis variable would correspond to a particular gas.

Exact inference is in principle possible only if the factorization of a JPD corresponds to a factor tree [7]. This is the case if (i) a BN has a DAG with a tree topology or (ii) a multiply connected BN is compiled into a secondary structure, such as for example a Junction tree. Alternatively, by instantiating appropriate variables in a multiply connected BN we obtain a model which is equivalent to a factor tree without any compilation of secondary structures [3]. For example, this is achieved if variables *M* and *T* in the model from Figure 2 are instantiated; we obtain a system which corresponds to a set of factors that share a single process variable *GasX*.

The model depicted in Figure 1 corresponds to a single time interval (i.e. time slice) in which a sequence of observations was obtained while the states of non-leaf variables including *GasX* remained constant; within such a time slice the electronic circuitry of a sensor evaluates the measurements and generates a stream of sensor reports. We assume that the domain's hidden phenomena are *quasi-static* in the sense that they do not change during a single time slice. For example, in Figure 1 nodes *GasX* and *Cond* represent the presence/absence of gas and conductivity of the semiconductors, respectively. We assume that for the duration of the time slice either $GasX = true$ or $GasX = false$. On the other hand, within a time slice we obtain several sensor reports at different points in time, which in turn can be seen as a result of a first order Markov process. However, since we assume that the non-leaf nodes are *quasi-static* during a time slice, we can equivalently describe such a process through a set of branches rooted in a single node (see [2]). In other words, for each observation from a sequence obtained within a single time slice we obtain a correct causal model by simply appending a new leaf node (see for example the leaf nodes E_i^x in Figure 1).

3 Model Construction

By explicitly considering causality, it turns out that in domain models capturing monitoring processes, the critical dependencies can easily be determined. In such domains causal processes are well understood, since they are to a great extent created through designers of monitoring systems and operators; such processes are created by (i) putting together various man-made components, such as sensors, communication systems and (ii) exploitation of well known procedures, such as asking people that call an emergency center specific questions about the symptoms and percepts. We can safely assume that sensor developers understand main aspects of causal mechanisms in the sensing devices while the professionals in an estimation process follow well established procedures.

Model construction consists of (i) the development of causal graphs, which capture direct dependencies between the variables, i.e. qualitative knowledge, and (ii) the determination of the parameters which capture the strength of the causal (stochastic) influences, i.e. quantitative knowledge.

3.1 Construction of Causal Graphs

Construction of causal domain models for gas detection is exploiting the fact that with each sensor or a human reporter a new local causal process is introduced to the domain. By introducing a new gas sensor, the gas initiates processes on the measuring element and in the sensor's circuitry which eventually produces sensor reports. For example, the process introduced by the third sensor is captured by the graph fragment consisting of variables S_3^{Ion}, C_3^{Ion} and $E_{n+1}^{Ion}, \ldots, E_o^{Ion}$ shown in Figure 1. In principle, the local processes within a sensor become part of the overall mechanism generating observations. Similar reasoning applies to human observers.

Moreover, each type of sensors usually provides observations about a single phenomenon represented by a *process variable*, such as, for example, *Cond* and *Ion* in Figure 1. Thus several sensors of the same type are influenced by a single phenomenon resulting from a causal process and a few boundary conditions represented by *context variables*, such as *M* and *T* in Figure 1. In addition, components and reports of one sensor do not influence components and reports of another sensor. Thus, we can represent a process generating observations of a certain type with a network fragment which is sparsely connected to other fragments (see, for example, Figure 1). Only variables for which direct dependencies exist are connected directly. *In other words, the resulting causal graph describing the overall observation generating process is sparsely connected.*

3.2 Determination of Parameters

The modeling parameters, i.e. the conditional probabilities defining the CPTs, are obtained with the help of the Maximum-likelihood estimation method [5], for example, or through expert introspection. The former is a sound solution, however, it is viable if all variables in the model can be observed. In case of partially observable models EM algorithm can be used [1]. Sometimes, experts can estimate the conditional probabilities.

In principle, the CPTs relating variables corresponding to the sensor components, observed phenomena and the observations can be estimated by repeated experiments. In some cases, this approach can be also practical for the estimation of CPTs relating human reports and the presence of gases. For example, we could extract such relations from the database compiled by the DCMR milieudienst in Rijnmond, an environmental protection agency in the Port of Rotterdam. The database captures complaints/reports of citizens collected during incidents with known causes. The

$P(E_i^{Smell}\|GasX)$	$GasX = true$	$GasX = false$
$E_i^{Smell} = yes$	0.75	0.34
$E_i^{Smell} = no$	0.25	0.66

Table 1: Conditional probability tables capturing relations $P(E_i^{Smell}|GasX)$.

database allowed estimation of conditional probability distributions for simple reports from citizens indicating the presence or absence of an industrial air pollution, i.e. an anomaly caused by an abnormal concentration of any substance that can be classified as gas, chemical vapor or oil derivatives, which are typically released by the industry[¶]. In this case $GasX = true$ corresponds to the presence of such an anomaly. In principle we counted how often complainers responded with yes/no to a question "Does the smell remind you of a chemical or oil or gas?" This question could be asked via a web-page or by an automated response system when complainers call a special number provided by the DCMR. In the latter case the complainers respond by pressing various options, such as 1 for yes, 2 for no and 3 for don't know. By using the DCMR database we extracted the perception model $P(E_i^{Smell}|GasX)$ for citizens shown in Table 1. The estimation of the parameters in Table 1 was based on 586 incidents for which the cause of complaints was known. The 95% confidence intervals for the parameters were ±0.03 for $P(E_i^{Smell}|GasX = true)$ and ±0.037 for $P(E_i^{Smell}|GasX = false)$, respectively[‖].

4 Modeling Deviations

In real world settings it is impossible to obtain perfect models. The structure as well as the parameters inevitably deviate from the reality. In this section we discuss the origin of such deviations and mitigation of their impact.

4.1 Structural Deviations

Structural deviations are likely to be present in a causal model either due to the lack of knowledge of the true relations and relevant variables or they are introduced on purpose in order to simplify modeling and reasoning processes.

Figure 2 shows a simplified model derived from the model depicted in Figure 1. The influence of the sensor components C_j^x on sensor reports was ignored and variables *Cond* and *Ion* were omitted.

In this model we can identify two types of structural errors resulting in the introduction of confounding variables [12]. We facilitate the discussion by introducing two types of modeling fragments which can be identified in the true model in Figure 1: Confounding Modeling Fragments and Interactive Fork Modeling Fragments, respectively.

[¶]From the decision maker's point of view, it is important to differentiate between complaints indicating an industrial air pollution and other pollution types, since the former is likely to involve harmful substances.

[‖]The computation of these intervals was an approximation based on the fact that for large numbers binomial distributions approach normal distributions [8].

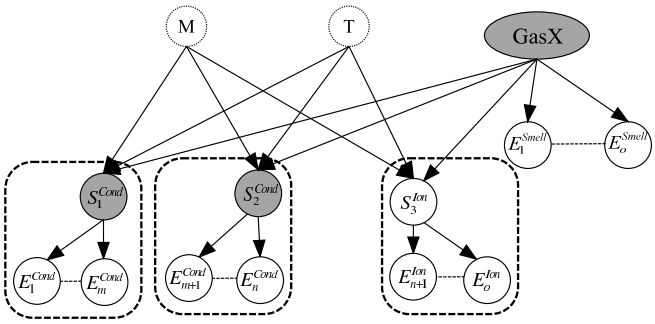
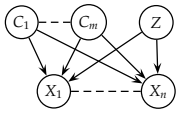


Figure 2: A simplified causal model of a monitoring process.

Definition 1 (Confounding Modeling fragment)**.** *A* **Confounding Modeling Fragment** *(CM fragment) is a graph defined over a set of multinomial random variables in* $\mathbf{V}$ *consisting of a hypothesis variable* $Z \in \mathbf{V}$*, a set of confounding variables* $\mathbf{C} \subset \mathbf{V}$*, where* $\mathbf{C} = \{C_1, \ldots, C_m\}$*, and a set of observable variables* $\mathbf{X} \subset \mathbf{V}$*, where* $\mathbf{X} = \{X_1, \ldots X_n\}$*, with the following dependence structure:*



The parameters of the model are denoted by $\boldsymbol{\theta}^c$ *and represent the CPT values corresponding to discrete (conditional) probability distribution values of the variables* $\mathbf{V}$ *in the model.*

Examples of CM fragments can be found in Figure 1 (see the dashed rounded rectangles). Note that the variables representing the status of electronic components are not represented in the simplified model from Figure 2. Consequently, C_j^{Cond}'s become confounders [12] in the simplified model which *incorrectly* implies independence $(E_p^{Cond} \perp\!\!\!\perp E_q^{Cond}|S_j^{Cond})_P$[**] for an arbitrary pair of measurements E_p^{Cond} and E_q^{Cond} from the $j-th$ sensor. Such a simplification is often inevitable due to the lack of knowledge and modeling resources. For example, it might be difficult to identify all influences of the electronic circuitry, determine the probability of failure of a sensor represented by C_j^{Cond} and to find the exact strength of their influence on the sensor reports.

The modeling errors influence the computation of $P(E_1^{Cond}, \ldots, E_m^{Cond}|S_j^{Cond})$, the likelihood of S_j^{Cond}, given the observations influenced by this variable. These likelihoods are passed in the inference algorithm as λ-messages [11]. In order to investigate the impact of such errors we compute the absolute difference between (i) the likelihoods $P(\mathcal{E}|S_j^{Cond})$ obtained with the correct CM fragment (see dashed rectangles in Figure 1) and (ii) the likelihood $P^*(\mathcal{E}|S_j^{Cond})$ obtained with simplified fragments (see dashed rectangles in

[**]Notation $(\mathbf{X} \perp\!\!\!\perp \mathbf{Y}|\mathbf{Z})_P$ means that a set of variables $\mathbf{X}$ is conditionally independent of variables $\mathbf{Y}$, given that the variables from set $\mathbf{Z}$ are observed (see [12]).

Figure 2):

$$f^{c}_{\Delta(\mathcal{E}|s_j)}(\boldsymbol{\theta}^c) = |P_c(\mathcal{E}|s_j) - P^*_c(\mathcal{E}|s_j)|, \quad (1)$$

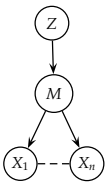
where s_j is a state of the variable S_j^{Cond} while $\mathcal{E}$ represents evidence instantiation for the observation variables E_i^{Cond} which are directly influenced by S_j^{Cond}. The difference $f^{c}_{\Delta(\mathcal{E}|s_j)}(\boldsymbol{\theta}^c)$ is computed for a specific parameterization $\boldsymbol{\theta}^c$ of the model.

It turns out that the impact of the omission of the confounding variables depends on the prior distribution over the states of the confounding variables. The influence of the prior is shown in Figure 3, where for a particular state s_j the *expected* value of $f^{c}_{\Delta(\mathcal{E}|s_j)}$ is plotted over different values of prior probability $P(C_j^{Cond} = true) = \alpha$ for different numbers of observation nodes (see the embedded legend in Figure 3). This computation took into account distributions over the instantiations of the evidence variables (i.e. observations), which were based on the ground truth parameterizations captured by the CPTs $P(E_i^{Cond}|S_j^{Cond}, C_j^{Cond})$. The ground truth CPT parameters were sampled from a uniform distribution. The CPTs $P(E_i^{Cond}|S_j^{Cond})$ of the erroneous model, on the other hand, were obtained by marginalizing C_j^{Cond} out of the product $P(C_j^{Cond})P(E_i^{Cond}|S_j^{Cond}, C_j^{Cond})$, where $P(C_j^{Cond})$ denotes the true prior probability over C_j^{Cond}. Note that, the prior probability over $P(S_j^{Cond})$ is not relevant, since likelihoods are computed. The plot shows that with increasing numbers of observation nodes also the absolute difference of likelihoods increases. However, this difference remains small if priors are close to 1 or 0. Similar is true for the sensors measuring the phenomena represented by *Ion*.

In general, it is plausible to assume that sensor components are reliable and the prior probability of a failure of the sensor's circuitry (i.e. $C_j^{Cond} = false$) is low. Therefore, in the presented use case, C_j^{Cond} does not have a significant impact on the computation of λ-messages and, consequently, on the overall classification performance.

Another important fragment type in the presented problem is the Interactive Fork Modeling Fragment (The term *interactive fork* originates from [12]).

Definition 2 (Interactive Fork Modeling Fragment)**.** *An* **Interactive Fork Modeling Fragment** *(IFM fragment) is a graph defined over a set of multinomial random variables in* $\mathbf{V}$*, consisting of a hypothesis variable* $Z \in \mathbf{V}$*, a mediating variable* $M \in \mathbf{V}$ *and multiple observable variables* $\mathbf{X} \subset \mathbf{V}$*, where* $\mathbf{X} = \{X_1, \ldots, X_n\}$ *with the following dependence structure:*



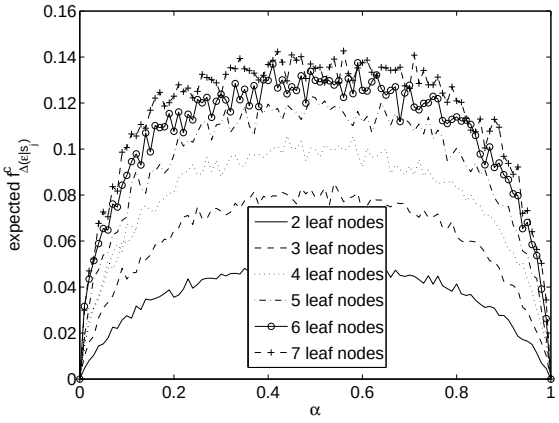


Figure 3: The expected likelihood deviation for the binary CM fragment as a function of the prior probability distribution over a confounding variable.

The parameters of the model are denoted by $\boldsymbol{\theta}^m$ *and represent the CPT values corresponding to discrete (conditional) probability distribution values of the variables* $\mathbf{V}$ *in the model.*

An example of an IFM can be found in Figure 1 (see the subgraph spanning *GasX*, *Cond*, S_1^{Cond} and S_2^{Cond} shown as gray nodes). Note that *Cond* is not represented in the simplified model from Figure 2 and, therefore, it becomes a confounding variable. In the true model the variables S_1^{Cond} and S_2^{Cond} are mediated by the variable *Cond*. However, without explicitly representing *Cond*, the simplified model *incorrectly* asserts the following conditional independence $(S_1^{Cond} \perp\!\!\!\perp S_2^{Cond}|GasX)_P$ in the underlying probability distribution. The omission of the mediating variable *Cond* influences the computation of the likelihood of *GasX*, given the observations from all sensors measuring the phenomenon corresponding to *Cond*. We investigate this effect by computing the absolute difference between the likelihoods $P(\mathcal{E}|GasX)$, obtained with the true model from Figure 1 and the likelihood $P^*(\mathcal{E}|GasX)$, based on the erroneous model shown in Figure 2:

$$f^{m}_{\Delta(\mathcal{E}|gasx)}(\boldsymbol{\theta}^m) = |P_m(\mathcal{E}|gasx) - P^*_m(\mathcal{E}|gasx)|, \quad (2)$$

where *gasx* is a state of variable *GasX* and $\mathcal{E}$ represents the evidence obtained by the sensors measuring *Cond*. The difference $f^{m}_{\Delta(\mathcal{E}|gasx)}(\boldsymbol{\theta}^m)$ is again computed for a specific parameterization $\boldsymbol{\theta}^m$ of the model.

It turns out that the impact of the omission of the mediating variables depends on the strength of the CPTs capturing $P(S_1^{Cond}|GasX)$ and $P(S_2^{Cond}|GasX)$, respectively. To show this we used the correct and the erroneous fragments consisting of gray nodes in Figure 1 and Figure 2, respectively. The influence of the CPT is shown in Figure 4, where for a particular state of *GasX* the *expected* likelihood difference $f^{m}_{\Delta(\mathcal{E}|gasx)}(\boldsymbol{\theta}^m)$ was computed over different values of the conditional probability $P(Cond = true|GasX = true) =$

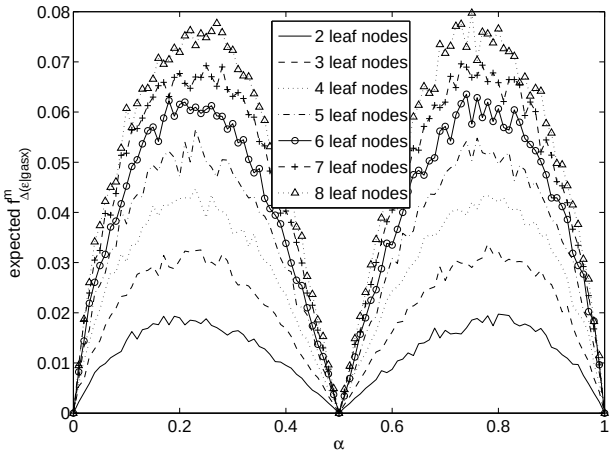


Figure 4: The average likelihood deviation for the binary IFM as a function of the CPT strength.

$P(Cond = true|GasX = false) = \alpha$ for different numbers of observation nodes. The ground truth CPTs relating the highlighted nodes in Figure 1 were used for the determination of the distributions over the evidence instantiations. These CPTs were sampled from a uniform distribution. The conditional probabilities $P(S_j^{Cond}|GasX)$ used in the erroneous model, on the other hand, were computed from the true distributions by marginalizing out variable *Cond*. Note that the parameters for $P(GasX)$ are irrelevant since we are interested in the likelihood of *Gas*X.

The plot in Figure 4 shows that the *expected* difference in likelihood $f^m_{\Delta(\mathcal{E}|gasx)}$ is the smallest when α is close to 0, 0.5 and 1. Thus, if the CPT relating *GasX* and *Cond* is strong then the omission of *Cond* does not have a significant impact on the computation of λ-messages and, consequently, on the overall classification performance.

4.2 Deviations of Modeling Parameters

In real world problems it is very likely that the amount of data is limited and detailed knowledge about all aspects of the domain is not available. Consequently, the conditional probability distributions encoded in the model (i.e. the parameters) significantly deviate from the true conditional probability distributions. For example, it is difficult to find parameters that precisely describe the distributions over the combinations of the sensor states and states of the related phenomena, such as for example $P(S_1^C|Cond)$. Similarly, it is often difficult to obtain modeling parameters that precisely describe the true probability of obtaining certain types of reports from humans, given the presence of various chemical substances; systematic databases with sufficient reports, as the one used for the estimation of the CPT shown in table 1, are rather an exception.

However, by considering the theory of BNs it was shown that reliable classification can be achieved even if the parameters significantly deviate from the true probabilities, as long as the used BNs satisfy simple conditions [4, 10]. *Key to robust inference are relations between the true conditional probability distributions and the distributions captured by the used CPTs [10].* Let's assume a true conditional probability distribution $P(E|C)$ between variables C and E whose states c_i and e_j correspond to causes and effects, respectively. It can be shown that a system is robust if the used CPTs $P^*(E|C)$ and the true distributions $P(E|C)$ satisfy simple conditions:

$$\forall e_j \in E : \operatorname*{argmax}_{c_i} P^*(e_j|c_i) = \operatorname*{argmax}_{c_i} P(e_j|c_i). \quad (3)$$

and

$$0.5 < \sum_{e_j \in B_{c_i}} P(e_j|c_i), \quad (4)$$

where B_{c_i} is the set of all states of E for which the likelihood of state c_i is maximum: $B_{c_i} = \{e_k|\forall c_j \neq c_i : P^*(e_k|c_i) > P^*(e_k|c_j)\}$. In case of binary variables the relations (3) and (4) are satisfied if, in both, the CPT describing the true conditional probabilities and the CPT from the model, the elements of the same diagonal exceed 0.5. For example, let's assume a true distribution $P(E|C)$ over binary variables E and C: $P(e|c) = 0.7$, $P(\overline{e}|c) = 0.3$, $P(e|\overline{c}) = 0.4$ and $P(\overline{e}|\overline{c}) = 0.6$. We say that a CPT $P^*(E|C)$ in a model correctly captures relations between the true probabilities if its parameters satisfy $P^*(e|c) > 0.5 \wedge P^*(\overline{e}|\overline{c}) > 0.5$.

In the targeted domains it is often plausible to assume that relations (3) and (4) can reliably be identified by experts or extracted from relatively small data sets with the help of machine learning techniques.

Moreover, in [10] it was shown that if the BN corresponds to a *factor tree* whose root is the hypothesis variable, then the expected classification accuracy asymptotically approaches 1 with the growing number of branches rooted in the hypothesis variable if relations (3) and (4) are satisfied for all CPTs in the BN. This is a consequence of the inherent properties of the BNs [10] and it is supported by the experimental results from section 5. A factorization corresponds to a factor tree if the BN has a DAG with a tree topology or appropriate variables in multiply connected BNs are instantiated. Example of the latter is instantiation of variables M and T in the multiply connected model from Figure 2.

If many information sources are available, we can obtain BNs corresponding to factor trees with large branching factors, which makes fusion reliable even if we use CPT parameters that deviate from the true distributions significantly. This can be illustrated with a simple example, where the detection is based on complaints only and the true distribution $P(E_i^{Smell}|GasX)$ is given by table 1. Thus, we would instantiate only the leave nodes corresponding to smell observations in the model from Figure 2. This would be equivalent to reasoning with a naive BN. If 0.5 were used as the decision threshold, the detection would be equivalent to simple majority voting [10]. Consequently, we can show that the lower bound on the detection performance would asymptotically approach 1 for arbitrary CPTs $P^*(E^{Smell}|GasX)$ as long as they satisfy the following relations: $P^*(E^{Smell} =$

$true|GasX = true) = P^*(E^{Smell} = false|GasX = false) >$ 0.5, given the table 1. For 11 reports the expected detection accuracy would exceed 0.86 while the accuracy would exceed 0.96 if we obtained more than 30 reports.

5 Experimental Results

The properties of the proposed Bayesian detection systems were demonstrated in a series of experiments. We created several detection systems with different configurations of chemical sensors and reports from humans. For each configuration we created a ground truth model, capturing the true distributions over the variables in a monitoring process. Such a ground truth model was used for the computation of distributions over all possible observation patterns which were then fed to different versions of detection systems and the expected classification performance for *GasX* was computed. We consider three types of detection systems:

Detection system 1 had a perfect domain model[††], that is, the structure and parameters are both correct.

Detection system 2 used the domain model with a perfect structure, however, the CPT parameters differed from the true parameters. For all nodes, except *GasX*, C_i^{Cond} and C_i^{Ion}, the CPT parameters deviated from the true probability up to 0.2, while the greater than/smaller than relations between parameters of the true CPTs were preserved.

Detection system 3 had a simplified structure and the CPT parameters differed significantly from the true parameters. In the simplified structure the variables of sensor components, *Ion* and *Cond* were omitted. All new CPTs were computed from the true model and additional deviations were introduced similarly to the case 2.

For each state z' of *GasX* we computed the probability that a certain observation pattern $\mathcal{E}$ will materialize, i.e. $P(\mathcal{E}|GasX = z')$. Moreover, for each observation pattern and each detection system with parameters $\boldsymbol{\theta}$ and structure G the posterior probability $P(GasX|\mathcal{E}, \boldsymbol{\theta}, G)$ was computed. This posterior was used for Bayesian classification [5]; the classifier returns state $\hat{z}$ for which $\hat{z} = \text{argmax}_z P(GasX = z|\mathcal{E}, \boldsymbol{\theta}, G)$. Based on the classified state $\hat{z}$ and the computed probability for a given observation pattern $P(\mathcal{E}|GasX = z')$ the expected classification accuracy is computed. The overall expected classification accuracy $E(acc)$ of a system is obtained by summing over all $P(\mathcal{E}|GasX = z')P(GasX = z')$, for which the classifier using a particular model with parameters $\boldsymbol{\theta}$ and structure G returned the true state z': $z' = \text{argmax}_z P(GasX = z|\mathcal{E}, \boldsymbol{\theta}, G)$; i.e. the expected classification accuracy $E(acc)$ is the sum of all probabilities of the situations for which the ground truth state z' equals to the classified state $\hat{z}$.

constellation of information sources	reports	detection system 1	2	3
1x Cond	3	0.6721	0.6721	0.6721
1x Cond, 1x Ion and 1x Complaint	7	0.7612	0.7452	0.7408
3x Cond, 1x Ion and 1x Complaint	13	0.7978	0.7894	0.7797
3x Cond, 1x Ion and 3x Complaint	15	0.8489	0.8402	0.8402

Table 2: Expected classification accuracy of gas anomaly detectors as a function of the number of reports.

For the three detection systems, i.e. systems 1, 2 and 3, the expected classification accuracy is given in Table 2 for different constellations of information sources. From this table we see that, compared to the perfect detection system 1, the classification performance of system 2 is reduced through deviating parameters. Structural deviations of system 3 have even greater impact on the performance. However, as more information sources are added to the system the classification performance increases for all three detection systems and would eventually converge to 1 as the number of sensors and reporters would approach infinity. Thus, the impact of modeling deviations can be mitigated by increasing the types and the number of information sources. Note also, that in system 3 the impact of the omission of *Cond* is not negligible with respect to the computed likelihood of *Cond*. Still, the impact on the overall system was mitigated by adding other types of sources, such as reports on smell (see last row in Table 2).

6 Discussion

In this paper we show that causal Bayesian networks allow efficient construction of fusion systems which support robust gas detection in real world applications. The focus is on the causal perspective which facilitates construction of well grounded models in several ways and provides guidance for the construction of inherently robust systems.

By explicitly using the knowledge about causal relations in the domain we can often construct models which are simplifications of the true processes, while they contain all variables and relations which are critical for good detection/classification. By adopting a view in which the observations are outcomes of causal stochastic processes we construct models which include variables that have a physical meaning and whose states represent measurable[‡‡] outcomes of stochastic causal processes.

In addition, by knowing direct influences between the relevant phenomena in the domain we obtain a sound rationale

[††] Since the numbers in columns of a CPT sum up to one, the models were fully defined by the following parameters: $P(Cond = true|GasX = true) = P(Cond = false|GasX = false) = 0.9$, $P(S_i^{Cond} = true|Cond = true) = P(S_i^{Cond} = false|Cond = false) = 0.95$, $P(Ion = true|GasX = true) = P(Ion = false|GasX = false) = 0.95$ and $P(S_i^{Ion} = true|Ion = true) = P(S_i^{Ion} = false|Ion = false) = 0.9$. Moreover, $P(GasX) = 0.5$, $P(C_i^{Cond}) = 0.99$, $P(C_i^{Ion}) = 0.99$ and for $P(E_i^{Smell}|GasX)$ we use the CPT in Table 1. For the sensor observation models we used $P(E_j^{Cond} = true|C_k^{Cond} = true, S_i^{Cond} = true) = 0.9$, $P(E_j^{Cond} = true|C_k^{Cond} = false, S_i^{Cond} = true) = 0.1$, $P(E_j^{Cond} = true|C_k^{Cond} = false, S_i^{Cond} = true) = 0.2$, $P(E_j^{Cond} = true|C_k^{Cond} = false, S_i^{Cond} = false) = 0.05$, $P(E_j^{Ion} = true|C_k^{Ion} = true, S_i^{Ion} = true) = 0.95$, $P(E_j^{Ion} = true|C_k^{Ion} = false, S_i^{Ion} = true) = 0.05$, $P(E_j^{Ion} = true|C_k^{Ion} = false, S_i^{Ion} = true) = 0.15$ and $P(E_j^{Ion} = true|C_k^{Ion} = false, S_i^{Ion} = false) = 0.1$.

[‡‡] There exist measurement devices or we know coarse tendencies from our experience

for the choice of the modeled relations between the variables; e.g. we know that components and reports from different passive chemical sensors do not influence each other. Consequently, the resulting causal models are not fully connected. In fact, as we show in this paper, the models can have sparse structures, which are associated with relatively low numbers of parameters. This has important implications regarding the tractability of the parameter estimation processes; good parameter estimates can be obtained from limited data sets.

On the other hand, with the help of rigorous theory of Bayesian networks we can understand the impact of different types of modeling simplifications and inevitable deviations with respect to the true processes. This allows further reduction of the model complexity and construction effort without a significant impact on the grounding of the fusion processes and, consequently, the detection performance.

We identified two common types of structural violations, which are inevitable in real world settings and showed that, under realistic assumptions, they do not have a significant impact on the overall detection performance. Moreover, in [10] it was shown that factorization captured by the domain models influences the robustness of Bayesian classification/detection processes. It turns out that BNs relevant for the discussed detection problems factorize the joint probability distribution, such that robust fusion can be achieved with respect to the deviations between the modeling parameters and the true distributions over the variables of interest. This is important since in the targeted domains it is often very difficult to obtain parameters which precisely describe the true distributions. For example, it is difficult to obtain enough data from all possible situations in which a sensor is operating. Even more challenging is to obtain useful sensor models for human reporters. However, as it was shown in [10], we can achieve good detection performance if we are able to identify simple relations between the true probabilities. Estimation of such relations is often tractable. This was confirmed in a recent pilot study carried out in cooperation with the University of Amsterdam and the DCMR, which showed that reports from humans in an industrial area indeed support effective detection.

Theoretical discussion in this paper is supported by experimental results based on an example of a gas detection system using complex patterns originating from heterogeneous information sources, such as different types of chemical sensors and humans. The experimental results show that the detection performance is affected by the parametric and structural model deviations, but the performance loss can be compensated by adding more information sources to the system.

An initial prototype of the presented detection system has been created and is currently being evaluated with chemical sensors. The sensors are not calibrated for specific gases but can be used for simple anomaly detectors which are incorporated into a distributed Bayesian gas detection system. Moreover, we implemented a system supporting automated querying of people. In order to avoid automated parsing and understanding of spoken language, we introduced querying methods which reduce the human responses to **yes**, **no**, **don't know**.

The presented detection system is part of a larger gas tracking framework which also makes use of dynamic Bayesian networks modeling the gas propagation processes [9]. A preliminary validation of the detection and tracking system was carried out with real world sensor data and human reports collected by the DCMR Milieudienst Rijnmond. While the initial results are encouraging, more experiments are currently underway to obtain statistically significant evidence for the overall system performance.

7 Acknowledgments

We would like to thank Cor van der Kooij from the DCMR Milieudienst Rijnmond for helping with the complaint database. The work presented in this paper is partially funded by (i) the European Union under the Information and Communication Technologies (ICT) theme of the 7th Framework Program for R&D (DIADEM project, ref. no: 224318) and (ii) ICIS project ref. no: BSIK03024.

References

[1] C. Bishop. *Pattern Recognition and Machine Learning*. Springer-Verlag, 2006.

[2] P. de Oude, B. Ottens, and G. Pavlin. Information fusion in distributed probabilistic networks. In *Artificial Intelligence and Applications*, pages 195–201, Innsbruck, Austria, 2005.

[3] P. de Oude and G. Pavlin. Efficient distributed bayesian reasoning via targeted instantiation of variables. In *IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, pages 323–330, 2009.

[4] P. Domingos and M. J. Pazzani. Beyond independence: Conditions for the optimality of the simple bayesian classifier. In *International Conference on Machine Learning*, pages 105–112, 1996.

[5] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience, 2nd edition, Nov. 2000.

[6] F. V. Jensen. *Bayesian Networks and Decision Graphs*. Springer-Verlag, New York, 2001.

[7] Kschischang, Frey, and Loeliger. Factor graphs and the sum-product algorithm. *IEEETIT: IEEE Transactions on Information Theory*, 47, 2001.

[8] T. M. Mitchell. *Machine Learning*. McGraw-Hill, New York, 1997.

[9] G. Pavlin, F. C. A. Groen, and P. de Oude. *A Distributed Approach to Gas Detection and Source Localization Using Heterogeneous Information*, volume 281 of *Studies in Computational Intelligence*. Springer, Berlin, Germany, 2010.

[10] G. Pavlin and J. Nunnink. Inference meta models: Towards robust information fusion with bayesian networks. In *The Ninth International Conference on Information Fusion*, Florence, Italy, 2006.

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

[12] J. Pearl. *Causality: Models Reasoning and Inference*. Cambridge University Press, 2000.