

Misper-Bayes: A Bayesian Network Model for Missing Person Investigations

DENIS REILLY¹, MARK TAYLOR, PAUL FERGUS¹, CARL CHALMERS¹,
AND STEVEN THOMPSON

School of Computer Science and Mathematics, Liverpool John Moores University, Liverpool L3 3AF, U.K.

Corresponding author: Denis Reilly (d.reilly@ljmu.ac.uk)

ABSTRACT Bayesian Networks are probabilistic graph models that can be used for classification, prediction, diagnosis and parameter learning. Probabilities can be inferred from the models and missing values can be imputed, based on probability theory. Missing person cases place a strain on the already overstretched resources of Police Forces. Such cases predominantly come from at risk groups such as children in care and people suffering from depression or dementia. Current approaches for dealing with such cases are manual and rely upon empirical studies and domain knowledge. This paper proposes the use of a Bayesian Network model, which can be used to predict the likely location of a missing person (*misper*) for a number of at risk groups. The model is evaluated using a set of misper cases and results compare very favourably with those of the manual processes currently used by UK Police forces. The novel approach described provides both a theoretical foundation and a practical framework for the future development of a decision support system. In addition to the model, a contribution is made through guidelines, which recount experiences in learning a Bayesian Network from data.

INDEX TERMS Algorithms, Bayesian methods, computer and information processing, decision support systems, modeling, probability.

I. INTRODUCTION

According to National Guidelines set out for UK Police Forces [1], A *missing person* is defined as:

“Anyone whose whereabouts cannot be established and where the circumstances are out of character or the context suggests the person may be subject of crime or at risk of harm to themselves or another.”

When someone is categorized as missing, the police will investigate their disappearance and try to find and safeguard them.

In 2013 the Guidelines introduced a second *absent person* category, defined as:

“A person not at a place where they are expected or required to be” and perceived to be “not at any apparent risk”.

When someone is categorised as absent, no police response is required except to monitor and review the situation.

Typically absent cases involve individuals who go missing frequently (often referred to as frequent fliers). They are

The associate editor coordinating the review of this manuscript and approving it for publication was Claudio Zunino¹.

likely to be designated a missing person for the first few times that they are missing, but, if they return unharmed, thereafter they may be designated absent.

Dealing with missing person cases consumes a large proportion of Police time and resources, particularly in urban areas.

This paper distinguishes a missing person (colloquially referred to as a *misper*) from a person who has lost their way, perhaps due to getting lost during a hiking expedition.

Fig. 1 highlights the scale of misper cases facing UK Police Forces and the volume of calls generated from dealing with such cases.

Mispers come from a spectrum of the population. Many are children (typically teenagers) who go missing from care homes, others are adults with mental illness or depression. Cases also include elderly people suffering from dementia related conditions. Murder (homicide) cases, manslaughter cases and death by misadventure often start out as misper cases until a body is located. Current practice relies on heuristics and localized domain knowledge. Social science studies will often interview mispers in the hope of eliciting knowledge in relation to the mispers thought processes, while

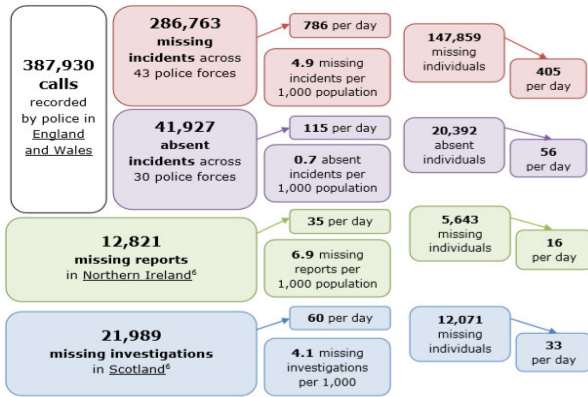


FIGURE 1. Key statistics for missing person in the UK, 2016-17 [1].

they were missing. Typically, Police rely heavily on historical data and behavioural patterns. For example, many teenagers who go missing are found in local parks, where teenagers are known to congregate. Elderly people suffering from dementia may travel to a location associated with a past event in their lives.

Bayesian Networks (BN) are directed graphical models, which have been used extensively in the fields of cognitive science and artificial intelligence throughout the latter half of the 20th and early 21st centuries. The models are based on the theorem of Thomas Bayes [2], which allows probabilities to be updated in light of new evidence. BN have been used for some time within the AI community and more recently amongst the machine learning community [3].

This paper describes the use of a BN model to capture the causal relationships that exist in a misper case and go on to show how the model can be used to impute missing values, such as the likely distance travelled and the likely location, where they may be found.

The paper makes a novel contribution through the development of the BN model, which is learned from data. The work also makes a contribution through the guidance proposed for determining the structure of the model. Current procedures, as used by UK Police forces, rely on iFIND [4], which is a PDF document, based on an empirical study. The BN model allows likely locations to be imputed from input data and serves as the basis of a computer based system.

II. RELATED WORK

Before describing the development of the model some related work is first reviewed, which is of relevance to the work that is the subject of this paper.

A. BAYESIAN NETWORKS FOR SEARCH AND RESCUE

To date no other work has been found, which is concerned with the use of BN to predict outcomes for at-risk misper groups. There is however, some notable research concerned with the use of BN for Search And Rescue (SAR), in relation to people (and “things”) whom have got lost. The distinction of course is that at risk misper groups have intentionally gone missing, whilst SAR cases have unintentionally got

lost. The most notable use of Bayesian inference for search techniques was that of the search for Air France Flight AF 447, which crashed into the Atlantic on 1st June 2009 [5]. After two years of unsuccessful searching, the team used a Bayesian procedure developed for search planning to produce the posterior target location distribution. The distribution was used to guide the search and the wreckage was located within a week.

Reference [6] describes a Bayesian approach to modeling lost person behaviors based on terrain features in Wilderness Search and Rescue. The approach uses a first-order Markov transition matrix for generating a temporal, posterior predictive probability distribution map. The approach also uses a Bayesian χ^2 test for goodness-of-fit and goes on to show that the model closely fits a synthetic dataset. Reference [7] provides a thorough study of missing person behavior in Australia. The study, conducted by Victoria Police as part of the SARBayes project, considers a large dataset of parameters, some of which are more significant than others. Terrain plays an important role and the range of activities, relating to the missing person, are also considered (e.g. climbing, canoeing, hunting etc.).

These works make valid contributions to further knowledge within the field, but they differ to this work in that they deal with cases of entities (people, planes) that go missing by accident. This work is concerned with people who largely go missing intentionally, which reinforces the choice of BN to help understand the structure of causality relations.

B. MACHINE LEARNING AND FORMAL APPROACHES

There are also several other machine learning related approaches for dealing with missing person cases. For example, [8] compares the use of neural networks and rule-based systems for missing person cases in Australia. In later work [9] considers the use of J48 to derive rules, based on the popular C4.5 decision tree generator.

In previous work conducted by the authors [10], a missing person model was developed based on Situation Calculus. The approach represented the state changes that take place over time, whilst the person is missing. The formalisms help to provide a consistent means to represent the uncertainty present in such investigations.

C. EMPIRICAL STUDIES

Two notable related empirical approaches are those of the UK booklet ‘Missing Persons: Understanding, Planning and Responding’ (colloquially referred to as the Grampian Study) [11] and the iFIND System [4], which is currently used by a number of UK Police forces.

The Grampian Study considers a similar set of at risk groups to this work. For each group the study provides a number of tables, which portray useful information, such as likely time periods of missing, distance travelled and likely places for a misper to be found. The Grampian Study also translates data into useful search ranges, which can be superimposed on a map.

TABLE 1. iFIND table of likely locations for 5-8 year olds.

	Where located	Further information
37%	Outside in local area	Streets; parks; shops; fields.
18%	Returned to home address	Mostly playing with friends in the local area either on the street or at a friend's address; one had played in a park; another played in a woodland area alone.
	At a friend's	Playing.
8%	Went to home address	Mostly from school; others mainly from a street or shops.
5%	Found on premises	Playing or hiding.
3%	Home of a family member	Usually grandparents (where mentioned).
	At a public building	Local shops.
1%	Returned to school	Walking the streets in local area before returning.
	Pub	One was playing pool with his brother; another was taken to pub by grandmother.
Individual cases	Woodland	Had become lost.
	In vehicle elsewhere	Playing with a friend within a parked vehicle on a neighbour's drive.
	Bus station	Had been at an older boy's house overnight.

iFIND follows a similar structure, but is based on more recent data to provide a more through coverage. iFIND provides more detail in terms of likely locations. Both Grampian and iFIND place emphasis on Time, Distance and Likely Location and these parameters also feature predominantly in this work. Table 1 shows a typical excerpt from iFIND, which highlights the places where mispers for the category were located. The majority being found outside locally, with a smaller proportion either returning home or being found at a friend's house.

The Grampian Study and iFIND are manual solutions to misper cases. They provide lookup tables that Police officers and call handlers can use to plan search strategies. The model described in this paper provides a computer-based solution, which calculates location probabilities in response to input data related to a misper case.

D. GEOSPATIAL REASONING

In previous work conducted by the authors, the CASPER System (Computer Assisted Search Prioritization and Environmental Response) [12] was developed to study the Geographies of Missing Persons. CASPER combines data analysis and GIS to develop a Google map application to assist investigative and strategic decision making. CASPER was developed to a prototype stage and demonstrated to several Police forces as a viable alternative to their existing case management systems, namely COMPACT and NICHE.

CASPER (illustrated in Fig. 2) was rich in terms of the geospatial information it provided, being able to display heatmaps, places of interest and even live CCTV footage. CASPER allows the search team to overlay a range of different layers onto a map region of interest. For example, the team may choose to overlay information of ATM cash machines if it is known that a misper may be short of money. Alternatively, suicide hotspots can be overlaid (from precompiled suicide data) when dealing with a potential suicide case.

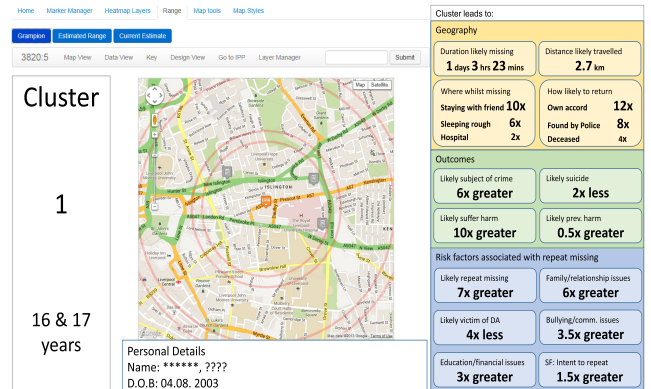


FIGURE 2. CASPER missing persons prototype.

However, the algorithms used in CASPER were largely rule-based algorithms (based on empirical data) which were not optimal.

III. PRELIMINARIES

Formally, for a discrete random variable $X = \{X_1, \dots, X_n\}$, a BN is an annotated directed acyclic graph, which encodes a joint probability distribution (JPD) over X . Formally, a BN is expressed as the pair $N = \langle G, \Theta \rangle$. The first element in N , is a directed acyclic graph, $G = (V, E)$. V denotes the random variables in X , and E denotes the edges, which represent direct dependencies between the variables. The second element Θ denotes the set of parameters, which quantify the network, via conditional probability tables. Each node is annotated with a conditional probability distribution, $P(X_i | Pa(X_i))$, representing the conditional probability of the node X_i given its parents in G . The network N defines a unique JPD over X given by:

$$P(X_1, \dots, X_n) = \prod_{i=1}^N P(X_i | Pa(X_i)) \tag{1}$$

In a BN a conditional probability $P(X | Y)$ is the probability of an event X occurring given that Y occurs. A marginal probability is effectively an unconditional probability. A marginal probability is a distribution formed by calculating the subset of a larger probability distribution. For example, for a JPD $P(X, Y)$ the probability of X can be determined simply by summing all the values for $X = False$ and $X = True$ in the joint table. For a query on a node in a BN, the result is often referred to as the marginal for that node.

For BNs, inference, is the computational methods for deriving answers to queries given a probability model expressed as a BN. Inference in BNs can take on several different forms [13]–[15], broadly speaking, it may be exact or approximate, depending upon the structure of the graph. Exact inference is not always possible when the number of combinations and paths are excessively large. However, it is often possible to refactor a BN graph (i.e. alter the graph structure) before resorting to approximate inference.

Let U be the set of random variables. Let $U^e \subseteq U$ be the set of known (evidence) variables. Let $X^q \in U \setminus U^e$ be the

variables of interest (queries) and let $U^r = U \setminus (U^e \cup X^q)$ be the set of remaining variables.

The probability distribution of the evidence variables and the query variables can be calculated, via marginalization, as:

$$P(X^q, U^e) = \sum_{U^r} P(X_1, \dots, X_N) \quad (2)$$

The normalization may be calculated as:

$$P(U^e) = \sum_{U^q} P(X^q, U^e) \quad (3)$$

Then conditional probabilities may be calculated as:

$$P(X^q | U^e) = \frac{P(X^q, U^e)}{P(U^e)} \quad (4)$$

Inference can be used to ask a range of different questions, in relation to the probability distribution, depending upon the nature and context of the problem at hand. Typically one or more of the following are likely to be of interest:

- Diagnosis: $P(X = \textit{cause} | U = \textit{symptom})$
- Prediction: $P(X = \textit{symptom} | U = \textit{cause})$
- Classification: $\textit{arg max}_{\textit{class}} P(X = \textit{class} | U = \textit{data})$
- Decision-making (given a cost function)

The particular strengths and weaknesses of BN are covered well in [16]. To summarize:

- They provide a natural way to handle missing data
- Suitable for small and incomplete datasets
- Combine different sources of knowledge
- Explicit treatment of uncertainty and support for decision analysis
- Fast response to queries

Essentially, a BN defines a unique JPD over X and computationally the JPD takes the form of a large table, constructed from the tables defined at individual nodes, in accordance with the graph links. So computationally, inference is the process of scanning the joint table to find a value (or values), which correspond to evidence E , possibly summing values along the way.

Often, the table will take the form of a sparse matrix (i.e. many zero entries) and this property can be exploited to make inference tractable, even when the number of parameters is very large. Subsequent sections in the paper will consider certain legal rearrangements of the JPD table, can be used to marginalize out certain parameters. Such rearrangements allow queries to be satisfied in linear-time methods by identifying a subgraph of the original graph relevant to the query [17].

IV. DESIGN AND IMPLEMENTATION

A. MISPER-BAYES MODEL DEVELOPMENT

The development of a BN model requires the learning of two components: the graph topology (structure) and the parameters of each conditional probability distribution. Both structure and parameters can be learnt from data. However, learning structure is much harder than learning parameters.

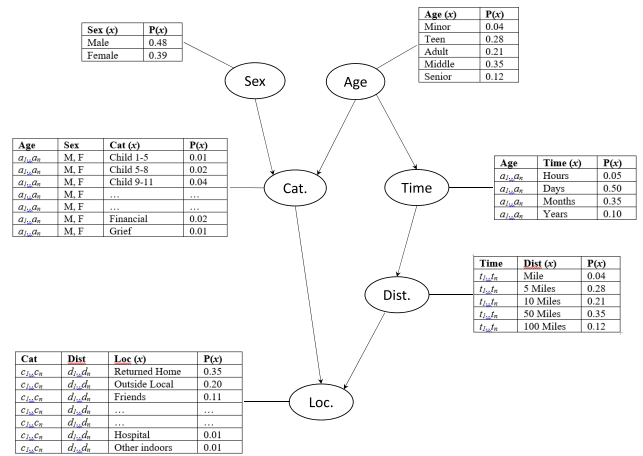


FIGURE 3. Misper-Bayes model.

There are a number of established techniques available to learn both the parameters and the structure [18]. Algorithms for learning a BN structure from data have two components: a scoring metric and a search procedure. The scoring metric computes a score reflecting the goodness-of-fit of the structure to the data. The search procedure tries to identify network structures with high scores and is regarded as NP hard [19].

Typically, the Naïve Bayes classifier provides a good place to start in relation to learning BN structure for a relatively small set of variables. The Naïve Bayes classifier (and its variants) [20] provides a baseline model for many machine learning classification problems. Naïve Bayes gives surprisingly good results, provided the condition of independence amongst variables hold. Unfortunately, in this case independence does not hold. For example, the different mental health categories, under consideration, are age related.

After dismissing Naïve Bayes, the development process went on to develop a Generalized Bayesian Network (GBN) and chose the bnlearn Python library [21] to learn the structure, based on data from iFIND. After several iterations and variable eliminations the development process arrived at the graph similar to that of Fig. 3. bnlearn starts with an empty network structure of all variables, then proceeds by adding, removing and reversing edges between nodes to maximize the goodness of fit of the model. The final structure, learnt by bnlearn, contained an excessive number of edges, likely due to overfitting (i.e. the noise within the data had been represented in the model itself). These unnecessary edges out were thinned out, based on the interpretation of the causal relationships between variables to deliver the final structure of Fig. 3. Finally, bnlearn was used to learn the parameters using iFIND data compiled from the summary table (Table 2).

In summary, automated structure learning is useful to develop the initial structure, but it can lead to overfitting and manual intervention is required to thin out some unnecessary edges. Based on the journey in developing the BN model, it was thought that it would be worthwhile sharing the experiences of learning a BN model in terms of structure and parameters, as presented in Algorithm 1.

TABLE 2. iFIND data summary.

iFIND Summary				
Category	Male	Female	Joint M+F	Totals
Children 1-5			35	35
Children 5-8			158	158
Children 9-11			246	246
Children 12-17	394	561		955
Children 12-17 (In Care)	431	376		807
Autism	118	30		148
Asperger Syndrome			121	121
ADHD	274	74		348
Genetic Condition			36	36
Intellectual Impairment	268	241		509
Dementia	133	80		213
Depression	283	335		618
PTSD			43	43
Suicidal (Alive)	258	130		388
Suicidal (Deceased)	84	19		103
Schizophrenia	201	76		277
Bipolar	97	96		193
Personality Disorder	35	94		129
Eating Disorder			32	32
Financial Problems	94	33		127
Grief			53	53
Totals	2670	2145	724	5539

Fig. 3 represents the model in terms of a digraph and associated conditional probability tables. The nodes in the graph represent the random variables, which are linked through the conditional probability tables.

Most of the tables are fairly self-explanatory, with a couple of exceptions: the Cat(x) table reflects the different categories of at risk mispers and the Loc(x) table reflects the different locations that mispers are likely to be found. Note that there is no edge connecting Cat(x) and Time(x) (although there was an edge in an earlier version of the model). It transpired that Age(x) provides a better predictor of the time spent missing than Cat(x). For example, the age of a young child or an elderly subject has a direct bearing on the time that they are missing. There were other variables that could have been included into the model such as race, ethnicity and deprivation index, but these were seen to have a lesser effect than the variables shown in Fig. 3.

Recalling equation (1), the JPD for the Misper-Bayes graphical model (Fig. 3) can be written as:

$$P(L,D,C,T,S,A) = P(L|D,C)P(D|T)P(C|S,A)P(T|A)P(S)P(A) \tag{5}$$

where:

$P(A)$ represents the probability of the different age groups.

$P(S)$ represents the probability of the sex types male and female.

Algorithm 1 (BN Model Development Guidelines)

1. If variable set is small and variables are independent apply Naïve Bayes (or its variants)
2. Otherwise consider General BN (GBN)
 - 2.1 Learn structure
 - 2.2 Check structure for Markov blanket
 - 2.3 If Markov blanket exists
 - 2.3.1 Discard variables outside of the blanket
 - 2.4 Otherwise
 - 2.4.1 Check structure for excessive edges
 - 2.4.2 Foreach edge:
 - 2.4.2.1 Check causal relationship
 - 2.4.2.2 Remove if not needed
3. Learn parameters
4. Evaluate model accuracy and performance
5. If performance is acceptable
 - 5.1 Validate model
6. Otherwise
 - 6.1 If GBN is not a tree
 - 6.1.1 Construct Tree approximation
 - 6.2.2 Compare Tree against GBN
 - 6.2 Validate Tree or GBN

$P(T | A)$ represents the conditional probability of time missing, based on age.

$P(C | S, A)$ represents the conditional probability of the different categories, based on sex type and age group.

$P(D | T)$ represents the conditional probability of distance travelled, based on time missing.

$P(L | D, C)$ represents the conditional probability of the likely location, based on the different categories and the distance travelled.

In Fig. 3 the Sex(x) category values (Male/Female) do not total to 1 because for certain categories there is no distinction between gender types in the iFIND data (Table 1). This is the case for young children for whom gender is not significant.

The current model represents time missing as a discrete variable (Time(x)) and as such, the model does not provide a continuous representation of time. Section V.D alludes to misper cases for which time may prove significant and suggest alterations to the model to accommodate time as a continuum.

B. MISPER-BAYES IMPLEMENTATION AND INFERENCE

The Misper-Bayes model was implemented in Python using the pomegranate machine learning package [22], which provides an easy to use abstraction of BN modelling. Pomegranate uses a *belief propagation* algorithm to satisfy conditional probability queries, which gives exact marginals when the graph is a tree (i.e. has no loops), but only approximates the true marginals in cyclic (loopy) graphs. Under these circumstances the algorithm is referred to as *loopy belief propagation*.

Pomegranate, provides a useful `predict_proba()` function, which uses loopy belief propagation in order to query probabilities for different parts of the graph and calculate marginals. For example (in the Python interpreter) a set of facts may be entered as a collection of dictionary entries to represent a realistic misper case.

```
>>>facts = { 'Sex': 'M', 'Age': 'T', 'Cat': 'C5',
'Time': 'D' }
>>>beliefs = map( str, model.predict_proba
(facts))
>>>print("\n".join("{}\t\t{}".format(state.name,
belief) for state, belief in zip(model.states,
beliefs)))
```

The encoding for observations is such that:

```
Sex: Male ('M')
Age: Teenager ('T')
Category: Child in Care (aged 12-17) ('C5')
Time: Missing Several Days ('D')
```

The beliefs map will provide probability predictions for likely Distance travelled and likely Location to be found. After some Python pretty printing, the output is as shown below.

```
Sex           M
Age           T
Cat.          C5
Time          D
Distance      {
    "class": "Distribution",
    "dtype": "str",
    "name": "DiscreteDistribution",
    "parameters": [
        {
            "S": 0.383,
            "M": 0.344,
            "L": 0.273,
            "X": 0.0
        }
    ],
    "frozen": false
}
Location      {
    "class": "Distribution",
    "dtype": "str",
    "name": "DiscreteDistribution",
    "parameters": [
        {
            "Pc": 0.051,
            "Ts": 0.02,
            "Aa": 0.01,
            "Fr": 0.081,
            "Oe": 0.04,
            "Ol": 0.162,
            "Fa": 0.091,
            "Re": 0.545
        }
    ],
    "frozen": false
}
```

```
"frozen": false
}
```

The figures of interest here are the imputed values for likely Distance travelled and likely Location to be found.

According to the encoding the figures tell us:

```
Distance: 'S' - Short Distance (up to 2 miles)
           = 0.383
           'M' - Medium Distance (up to 10 miles) = 0.344
           'L' - Long Distance (up to 50 miles) = 0.273
           'X' - Extra Long Distance (> 50 miles) = 0.0

Location: 'Re' - Returned Home (i.e. Care Home) =
           0.55
           'Ol' - Outside Local (e.g. Street, Park) = 0.16
           'Fa' - Family Members House = 0.09
           'Fr' - Friends House = 0.08
           'Pc' - Place of Care (e.g. Police/social care) =
           0.05
           'Oe' - Outside Elsewhere = 0.04
           'Ts' - Train Station = 0.02
           'Aa' - Another Address = 0.01
```

The BN model was designed and implemented using data relating to UK misper cases. As such, the current implementation will have a bias towards UK cases, although it is felt that the model itself (Fig. 3) defines the nodes and the graph structure that are sufficiently generic to apply internationally.

V. RESULTS AND DISCUSSION

A. MODEL EVALUATION

The Misper-Bayes model was evaluated using a series of queries with a set of misper cases. For each query, the results of the model were cross checked against the results of the iFIND system. The result comparisons are summarized in Table 3.

Table 3 shows the three most likely locations for each misper case (other locations with very low probabilities are omitted or brevity). The Misper-Bayes column has probabilities rounded to 2 decimal places. As can be seen the majority of results give the same value as iFIND and several results that are not the same are accurate within $\pm 1\%$. This indicates that the model converges to the results provided in iFIND. Details of how the iFIND results were rounded is unknown, also in the majority of iFIND Location tables there is a final row termed 'Individual Cases' for which no numerical values are given. These 'Individual Cases' could explain the very slight variation, between or own Misper-Bayes results and the results of iFIND.

Although data from iFIND was used in the development of the model, it is important to stress that the Location values for Misper-Bayes were *imputed* from the model, based on the conditional relationships of Fig. 3. Typically a Police Officer using iFIND would need to perform a manual lookup to locate the appropriate table in order to determine the likely Location.

B. MODEL REFINEMENT-TREE APPROXIMATION

As shown in Table 3, the Misper-Bayes model (Fig. 3) provides a very similar set of results to those of iFIND. The model was examined further to see if it could be improved or made more computationally efficient.

TABLE 3. Results of queries for misper cases (iFIND and Misper-Bayes comparison).

(A) CHILDREN AND YOUNG ADULTS

Sex	Age	Cat.	Dist	iFIND Loc.	Misper-Bayes Loc.
M	12	Child 12-17 (In Care) C5	S: 0.38 M: 0.34 L: 0.27 X: 0.0	1. Returned Home (0.54) 2. Outside Local (0.16) 3. Family Member (0.09)	1. Returned Home (0.55) 2. Outside Local (0.17) 3. Family Member (0.09)
F	12	Child 12-17 (In Care) C5	S: 0.41 M: 0.35 L: 0.23 X: 0.0	1. Returned Home (0.33) 2. Friends (0.15) 3. Outside Local (0.13)	1. Returned Home (0.34) 2. Friends (0.15) 3. Outside Local (0.13)
M	16	ADHD	S: 0.39 M: 0.34 L: 0.25 X: 0.01	1. Returned Home (0.35) 2. Outside Local (0.23) 3. Friends (0.13)	1. Returned Home (0.35) 2. Outside Local (0.24) 3. Friends (0.13)
F	15	ADHD	S: 0.39 M: 0.35 L: 0.25 X: 0.0	1. Returned Home (0.39) 2. Outside Local (0.24) 3. Friends (0.12)	1. Returned Home (0.39) 2. Outside Local (0.24) 3. Friends (0.13)
M	21	Depression	S: 0.27 M: 0.32 L: 0.3 X: 0.1	1. Returned to Last Place Seen (0.34) 2. Outside Local (0.08) 3. In Vehicle (0.07)	1. Returned to Last Place Seen (0.35) 2. Outside Local (0.09) 3. In Vehicle (0.07)
F	25	Depression	S: 0.29 M: 0.33 L: 0.3 X: 0.07	1. Returned to Place Missing From (0.25) 2. Friends (0.13) 3. Outside Elsewhere (0.10)	1. Returned to Place Missing From (0.25) 2. Friends (0.14) 3. Outside Elsewhere (0.11)

(B) ADULTS AND SENIORS

Sex	Age	Cat.	Dist	iFIND Loc.	Misper-Bayes Loc.
M	31	Suicidal (Alive)	S: 0.33 M: 0.34 L: 0.26 X: 0.06	1. Returned Home (0.22) 2. Outside Local (0.20) 3. Family Member (0.07)	1. Returned Home (0.23) 2. Outside Local (0.20) 3. Family Member (0.08)
F	37	Suicidal (Alive)	S: 0.33 M: 0.37 L: 0.25 X: 0.04	1. Outside Local (0.29) 2. Returned Home (0.16) 3. Outside Elsewhere (0.09)	1. Outside Local (0.29) 2. Returned Home (0.17) 3. Outside Elsewhere (0.09)

As mentioned previously in section 1.1, there are certain legal rearrangements of the joint probability distribution (JPD) table through which certain parameters may be marginalized

TABLE 3. (Continued.) Results of queries for misper cases (iFIND and Misper-Bayes comparison).

M	32	Bipolar	S: 0.41 M: 0.34 L: 0.24 X: 0.0	1. Returned to Last Place Seen (0.31) 2. Presented to Place of Care (0.09) 3. Outside Elsewhere (0.07)	1. Returned to Last Place Seen (0.31) 2. Presented to Place of Care (0.09) 3. Outside Elsewhere (0.07)
F	34	Bipolar	S: 0.39 M: 0.36 L: 0.24 X: 0.0	1. Returned to Last Place Seen (0.27) 2. Friends (0.1) 3. Family Member (0.08)	1. Returned to Last Place Seen (0.28) 2. Friends (0.1) 3. Family Member (0.08)
M	47	Financial	S: 0.32 M: 0.35 L: 0.30 X: 0.02	1. Returned Home (0.32) 2. Outside Local (0.15) 3. Family Member (0.06)	1. Returned Home (0.32) 2. Outside Local (0.15) 3. Family Member (0.07)
F	43	Financial	S: 0.34 M: 0.36 L: 0.28 X: 0.01	1. Returned to Last Place Seen (0.28) 2. Outside Local (0.15) 3. Family Member (0.12)	1. Returned to Last Place Seen (0.28) 2. Outside Local (0.15) 3. Family Member (0.12)
M	73	Dementia	S: 0.33 M: 0.34 L: 0.32 X: 0.01	1. Outside Local (0.44) 2. Outside Elsewhere (0.14) 3. In Vehicle (0.09)	1. Outside Local (0.45) 2. Outside Elsewhere (0.15) 3. In Vehicle (0.09)
F	71	Dementia	S: 0.37 M: 0.34 L: 0.27 X: 0.01	1. Outside Local (0.48) 2. Outside Elsewhere (0.1) 3. Returned to Last Place Seen (0.09)	1. Outside Local (0.48) 2. Outside Elsewhere (0.11) 3. Returned to Last Place Seen (0.09)

out of a graph and such rearrangements allow queries to be satisfied in linear-time.

There is a large body of work concerned with the manipulation and transformation of probabilistic graphical models to improve structured learning, inference and information storage and representation. There are various techniques, rooted in graph theory, such as moralization [23], d-separation [17] and tree-decompositions [24]. Of particular interest to this work is the idea of a tree approximation to the original model and determining how good this approximation actually is, based on the use of cross-entropy. Further still, what is the optimum approximation and how well does it match the results of the original model?

As mentioned previously, in BNs tree structures are desirable because belief propagation is exact. A tree with n vertices ($(n-1)$ edges) only requires $((d-1) + d(d-1)(n-1))$

Algorithm 2 (Chow Liu)

1. From the given distribution $P(x)$ compute the joint distribution $P(x_i, x_j)$ for all $i \neq j$
2. Using the pairwise distributions from step 1, compute the mutual information for each pair of nodes and assign it as the weight to the corresponding edge.
3. Compute the maximum-weight spanning tree (MST):
 - 3.1 Start from the empty tree over n variables.
 - 3.2 Insert the two largest-weight edges.
 - 3.3 Find the next largest-weight edge and add it to the tree if no cycle is formed; otherwise, discard the edge and repeat this step.
 - 3.4 Repeat step 3.3 until $n - 1$ edges have been selected (a tree is constructed).
4. Select an arbitrary root node, and direct the edges outwards from the root.
5. Tree approximation $Q(x)$ can be computed as a projection of $P(x)$ on the resulting directed tree (using the product-form of $Q(x)$).

parameters, where d is the domain size. The actual complexity of inference in a BN is proportional to its tree-width [25] which measures how closely the network resembles a tree.

The further examination of the model considered the Chow Liu theorem [26], which relies on Kullback-Leibler divergence (KL-divergence or cross entropy), which is expressed in (6).

$$D(P||Q) = P \sum_x P(x) \log \frac{P(x)}{Q(x)} \quad (6)$$

The KL-divergence, quantifies how much one probability distribution differs (or diverges) from another probability distribution ($||$ is the divergence operator, indicating that P deviates from Q). K-L divergence expresses the amount of information lost when Q is used to approximate P , proof of which is provided in [16]. The Chow Liu theorem is stated below.

Lemma:

For a JPD $P(X = x)$ and a tree structure T , the best approximation $Q(X = x)$ (i.e., $Q(X = x)$ that minimizes $D(P||Q)$) satisfies:

$$Q(x_i | x_{j(i)}) = P(x_i | x_{j(i)}) \quad \text{for all } i = 1, \dots, n \quad (7)$$

Such $Q(x)$ is called the projection of $P(x)$ on T .

Theorem (Chow Liu):

For a JPD $P(x)$, the KL-divergence $D(P||Q)$ is minimized by projecting $P(x)$ on a Maximum-Weight Spanning Tree (MST) over nodes in X , where the weight on the edge (X_i, X_j) is defined by the *mutual information measure*:

$$I(X_i, X_j) = \sum_{x_i x_j} P(x_i, x_j) \log \frac{P(x_i, x_j)}{P(x_i) P(x_j)} \quad (8)$$

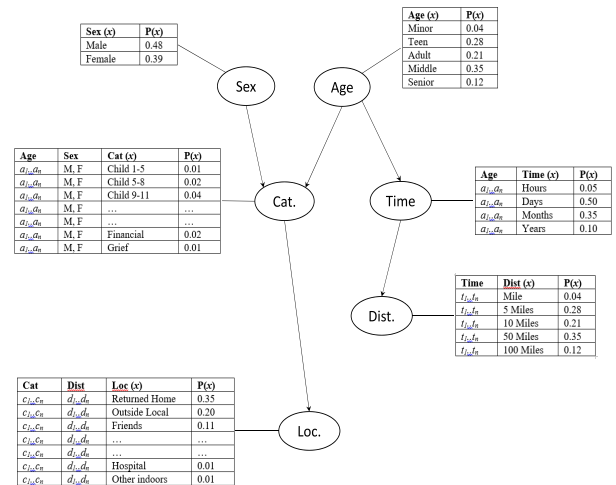


FIGURE 4. Misper-Bayes polytree representation.

If the Chow Liu algorithm is applied to the previous Misper-Bayes model, the model is transformed into the polytree structure shown below in Fig. 4.

Effectively, the algorithm produces an approximation, which is always a tree. It works by computing the weight $I(X_i, X_j)$ of each edge between nodes X_i, X_j and finding the MST. When the algorithm is applied to the graph of Fig. 3, the MST eliminates the edge connecting nodes 'Dist.' and 'Loc.'

To evaluate the revised polytree approximation the same set of queries from Table 3 were executed. Numerically, the marginals were almost identical (barring some rounding errors).

C. MODEL ASSESSMENT (QUANTITATIVE)

Given that the results of the queries were largely the same, the execution time for queries on the polytree model were then examined against those for the original acyclic graph. It turned out that the execution times for the polytree model were only marginally less than those of the original acyclic graph model. On average a query for the original graph model took 0.036 seconds, whereas the same query for the polytree model took 0.03 seconds, executed on an Intel Core i7, 16GB RAM (no GPU). As mentioned previously, pomegranate uses a loopy belief propagation algorithm for the implementation of the `predict_proba()` function, which is an inexact algorithm that converges to the exact solution on BNs which have a tree structure.

The additional time for the original graph model (0.006 seconds) is likely to be down to the time taken for convergence. By realistic BN standards both of the models are relatively small in terms of the number of nodes and edges. However, the models contain relatively large, sparse conditional probability tables (e.g. $Loc(x)$ is 700×4 elements). The literature provides specialised algorithms for dealing with sparseness in BN, which restrict the search space based on heuristics that either bound the search space by limiting the degree of nodes within the network or by limiting the set of possible edges [27]. Many BN models typically consist of

a large number of variables (e.g. medical applications may use several hundred variables) with small probability tables. In contrast, the model is based on a small number of variables with large probability tables.

Computationally, the latter type of model is beneficial for satisfying queries. For models with many variables (i.e. nodes), more effort is expended traversing the nodes as opposed to the values in a probability table associated with a given node. To improve query performance the model was transformed into a polytree using the Chow Liu algorithm, which is based on mutual information. The Chow Liu algorithm is known to reduce sparseness [27] but results showed only a minimal performance improvement, which suggests that sparseness in models with few variables yet large probability tables does not have a major impact on performance.

After examining the revised polytree model, it was concluded that the original acyclic graph is preferable as it is semantically richer. For example, the location of a missing person is likely to depend on how far a misper has travelled. If a misper has travelled a considerable distance, they may be less likely to frequent certain locations that they would do if they had stayed local.

D. MODEL ASSESSMENT (QUANTITATIVE)

The model performs well for cases that fit the discrete categories and exhibit normal behavior within that category. Results of such cases compare very closely to those of iFIND. However, it is unknown how well the model performs for cases that are a combination of categories. For example, if an individual is categorized, equally, as suffering from ADHD, but with a history of schizophrenia then, manually, one could impute the outcome for each case and look for common ground between the outcomes to inform the search strategy. However, for many cases, one category will be more dominant than the other.

There was no data available to consider such cases because UK Police Forces record data based on the discrete categories considered in the paper. Hybrid cases are, at present, largely dealt with based on the domain knowledge of the officers involved in the investigation. However, to consider a sample of such cases the tables in the Python implementation were modified to allow values to be imputed for queries with two categories of equal (50:50) weighting. This did not require any alteration to the model itself (Fig. 3), only the implementation. Based on this revised implementation, queries for two categories could be issued as below.

Sex: Male ('M')
 Age: Teenager ('T')
 Category: Child in Care (aged 12-17) ('C5'),
 ADHD ('Ad')
 Time: Missing Several Days ('D')

The results for a small sample of queries are shown in Table 4.

The Combined Category results are comparable with those imputed from real data in the Individual Category (Misper-Bayes) column. For the Combined Category results

TABLE 4. Results of queries for misper cases (individual categories in column 5 and Combined (50:50) categories in column 6).

(A) CHILDREN					
Sex	Age	Cat.	Dist	Individual Category (Misper-Bayes) Loc.	Combined Category Loc.
M	12	Child 12-17 (In Care) C5	S: 0.38 M: 0.34 L: 0.27 X: 0.0	1. Returned Home (0.55) 2. Outside Local (0.17) 3. Family Member (0.09)	1. Returned Home (0.44) 2. Outside Local ((0.21) 3. Friends (0.1) 4. Family Member (0.09)
		ADHD	S: 0.39 M: 0.34 L: 0.25 X: 0.01	1. Returned Home (0.35) 2. Outside Local (0.23) 3. Friends (0.13)	
F	12	Child 12-17 (In Care) C5	S: 0.41 M: 0.35 L: 0.23 X: 0.0	1. Returned Home (0.33) 2. Friends (0.15) 3. Outside Local (0.13)	1. Returned Home (0.32) 2. Outside Local (0.16) 3. Friends (0.15)
		Autism	S: 0.38 M: 0.36 L: 0.21 X: 0.0	1. Returned Home (0.3) 2. Outside Local (0.18) 3. Family Member (0.15)	4. Family Member (0.14)
M	16	ADHD	S: 0.39 M: 0.34 L: 0.25 X: 0.01	1. Returned Home (0.35) 2. Outside Local (0.23) 3. Friends (0.13)	1. Returned Home (0.3) 2. Outside Local (0.21) 3. Returned to Place Missing From (0.2)
		Schizophrenia	S: 0.32 M: 0.3 L: 0.29 X: 0.07	1. Returned to Place Missing From (0.25) 2. Friends (0.13) 3. Outside Elsewhere (0.10)	4. Friends (0.12)
F	15	ADHD	S: 0.39 M: 0.35 L: 0.25 X: 0.0	1. Returned Home (0.39) 2. Outside Local (0.24) 3. Friends (0.12)	1. Returned Home (0.33) 2. Outside Local (0.27) 3. Friends (0.11)
		Eating Disorder	S: 0.29 M: 0.33 L: 0.3 X: 0.05	1. Outside Local (0.33) 2. Returned Home (0.26) 3. Friends (0.10)	4. Outside Elsewhere (0.07)

the four most likely locations for each case are shown. It is important to stress that the Combined Category results were not arrived at by simply taking an average of the Individual category results, they were imputed from the tables of the Misper-Bayes BN model. These results cannot be compared

TABLE 4. (Continued.) Results of queries for misper cases (individual categories in column 5 and Combined (50:50) categories in column 6).

(B) ADULTS					
Sex	Age	Cat.	Dist	Individual Category (Misper-Bayes) Loc.	Combined Category Loc.
M	31	Suicidal (Alive)	S: 0.33 M: 0.34 L: 0.26 X: 0.06	1. Returned Home (0.22) 2. Outside Local (0.20) 3. Family Member (0.07)	1. Returned to Last Place Seen (0.31) 2. Outside Local (0.19) 3. Family Member (0.17)
		Depression	S: 0.25 M: 0.33 L: 0.31 X: 0.08	1. Returned to Last Place Seen (0.34) 2. Outside Local (0.08) 3. In Vehicle (0.07)	4. In Vehicle (0.06)
F	37	Suicidal (Alive)	S: 0.33 M: 0.37 L: 0.25 X: 0.04	1. Outside Local (0.29) 2. Returned Home (0.16) 3. Outside Elsewhere (0.09)	1. Returned to Last Place Seen (0.29) 2. Returned Home (0.17) 3. Outside Elsewhere (0.1)
		Bipolar	S: 0.37 M: 0.38 L: 0.2 X: 0.0	1. Returned to Last Place Seen (0.27) 2. Friends (0.1) 3. Family Member (0.08)	4. Friends (0.09)

against any tangible results used in Police cases, but they serve only to show the versatility of the model in that it is amendable to change to satisfy specific requirements.

One drawback with this approach is that it is difficult to estimate the balance between the two (or more) categories. Ratios of 50:50 were chosen, although it could be 60:40 or 30:70. Police officers alone would not be able to judge the balance, it would require input from a Psychologist, which adds strength to the argument that misper cases require input from many sources not just Police. Typically, misper cases require a multi-agency approach with input from Police, Psychologists, Health-care professionals, family members and friends, to name but a few.

Inevitably, some misper cases prove difficult, particularly those in which the subject endeavors to stay missing. Under such circumstances the model will only provide a baseline from which search strategy decisions can be made. The model was chosen for its simplicity and accuracy in predicting location and the variables defined in the model are still valid, even for the more difficult cases. The success in dealing with the more difficult cases often depends on the response time of the Police officers involved and the thoroughness of their search of a particular location. These more difficult cases could be assisted by a model that accommodates changes that may occur over time. Such temporal effects exceed the scope of the current work, but this is something that will be pursued in future work.

In misper cases Police forces refer to the “golden hour”, particularly for cases involving children. This is the time period in which the application of standardized approaches will, in theory, yield a desirable result. It is also the time period in which tangible evidence and information is abundant. As time progresses the case may become more complex and standardized approaches alone may not yield a desirable result and Police may have to think “outside of the box”. Periodically, misper cases are reviewed to see if any further evidence has come to light to assist search strategies. The passage of time may lead to significant changes such as a misper failing to take vital medication, or falling short of money, food or accommodation. The scope of the current work was the development of a Bayesian Network model to impute likely locations for different categories of mispers. In future work, the model will be extended using Dynamic Bayesian Networks to incorporate the timeline of events associated with a misper investigation.

VI. CONCLUSION

This research has described the design and implementation of our Misper-Bayes model to assist Police forces in determining the whereabouts of a missing person. The work makes a novel contribution as it is the first computer-based solution to assist in actively dealing with misper cases.

Misper-Bayes provides a powerful tool, which can be used to good effect to whittle down the likely locations where the missing person may be found. The results of likely location queries on the Misper-Bayes model delivered very similar results to those of the iFIND system. The model was examined to see if a tree approximation provides a better alternative. It was concluded that a tree approximation is not needed, assuming the current implementation based on the Python Pomegranate package.

The strength of the model lies in its simplicity yet versatility. The model can accommodate some variation to the discrete categories through some changes to the Python implementation. When combined with a geospatial front-end (e.g. CASPER), the Misper-Bayes model can be used to very good effect to assist Police Officers with the prioritization of their search strategy. The approach demonstrated has scope to support evidence-based policing beyond that of missing person cases.

In addition to the development of the model, guidelines were provided that may prove useful for others faced with learning a BN model from data.

REFERENCES

- [1] National Crime Agency. (2017). *Missing Person Data Report 2016/2017*. [Online]. Available: <https://www.nationalcrimeagency.gov.uk/who-we-are/publications/304-2016-17-ukmpu-data-report-v1/file>
- [2] J. V. Stone, *Bayes Rule: A Tutorial Introduction to Bayesian Analysis*. Sheffield, U.K.: Univ. Sheffield, 2013.
- [3] B. D’Ambrosio, “Inference in Bayesian networks,” *AI Mag.*, vol. 20, no. 2, pp. 21–36, 1999.
- [4] N. Eales. (2015). *iFIND*. National Crime Agency (NCA), London, U.K. [Online]. Available: <https://missingpersons.police.uk/en-gb/resources/downloads/iFIND>

- [5] L. D. Stone, C. M. Keller, T. M. Kratzke, and J. P. Strumpfer, "Search for the wreckage of air France flight AF 447," *Stat. Sci.*, vol. 29, no. 1, pp. 69–80, Feb. 2014.
- [6] L. Lin and M. A. Goodrich, "A Bayesian approach to modeling lost person behaviors based on terrain features in wilderness search and rescue," *Comput. Math. Org. Theory*, vol. 16, no. 3, pp. 300–323, Sep. 2010.
- [7] E. Sava, C. Twardy, R. Koester, and M. Sonwalkar, "Evaluating lost person behavior models," *Trans. GIS*, vol. 20, no. 1, pp. 38–53, Feb. 2016.
- [8] K. Blackmore, T. Boosomaier, S. Foy, and D. Thomson, "Data mining of missing persons data," in *Proc. 1st Int. Conf. Fuzzy Syst. Knowl. Discovery, Comput. Intell. E-Age*, Singapore, 2002, pp. 305–314.
- [9] K. Blackmore and T. Boosomaier, "Comparison of See5 and J48: PART algorithms for missing person profiling," in *Proc. 1st Int. Conf. Inf. Technol. Appl. (ICITA)*, Sydney, NSW, Australia, 2002, pp. 1–7.
- [10] M. Taylor and D. Reilly, "Knowledge representation for missing persons investigations," *J. Syst. Inf. Technol.*, vol. 19, no. 2, pp. 138–150, Mar. 2017.
- [11] G. Gibb and P. Woolnough, "Missing persons: Understanding, planning, responding—a guide for police officers," Grampian Police, Aberdeen, U.K., Tech. Rep., 2017. [Online]. Available: http://www.searchresearch.org.uk/downloads/ukmpbs/GGibb_missing_person_report.pdf
- [12] D. Reilly, C. Wren, S. Giles, L. Cunningham, and P. Hargreaves, "CASPER: Computer assisted search prioritisation and environmental response application," in *Proc. 6th Int. Conf. Develop. eSystems Eng.*, Abu Dhabi, UAE, Dec. 2013, pp. 225–230.
- [13] A. Darwiche, *Modelling and Reasoning With Bayesian Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [14] A. Darwiche, "A differential approach to inference in Bayesian networks," *J. ACM*, vol. 50, no. 3, pp. 280–305, May 2003.
- [15] P. Larrañaga, H. Karshenas, C. Bielza, and R. Santana, "A review on evolutionary algorithms in Bayesian network learning and inference tasks," *Inf. Sci.*, vol. 233, pp. 109–125, Jun. 2013.
- [16] L. Uusitalo, "Advantages and challenges of Bayesian networks in environmental modelling," *Ecolog. Model.*, vol. 203, nos. 3–4, pp. 312–318, May 2007.
- [17] D. Geiger, T. Verma, and J. Pearl, "D-separation: From theorems to algorithms," in *Proc. 5th Workshop Uncertainty Artif. Intell.*, ON, Canada, 1989, pp. 118–125.
- [18] M. Scanagatta, A. Salmerón, and F. Stella, "A survey on Bayesian network structure learning from data," *Prog. Artif. Intell.*, vol. 8, no. 4, pp. 425–439, 2019.
- [19] D. M. Chickering, D. Heckerman, and C. Meek, "Large-sample learning of Bayesian networks is NP-hard," *J. Mach. Learn. Res.*, vol. 5, pp. 1287–1330, Oct. 2004.
- [20] J. Cheng and R. Greiner, "Comparing Bayesian network classifiers," in *Proc. 15th Int. Conf. Uncertainty Artif. Intell.*, Stockholm, Sweden, 1999, pp. 101–108.
- [21] *Bnlearn—Graphical Structure of Bayesian Networks*. Accessed: Mar. 26, 2021. [Online]. Available: <https://pypi.org/project/bnlearn>
- [22] J. Schreiber, "Pomegranate: Fast and flexible probabilistic modeling in Python," *J. Mach. Learn. Res.*, vol. 18, no. 164, pp. 1–6, 2018.
- [23] G. Elidan and S. Gould, "Learning bounded treewidth Bayesian networks," *J. Mach. Learn. Res.*, vol. 9, no. 2, pp. 2699–2731, 2008.
- [24] U. Kjaerulff, "Reduction of computational complexity in Bayesian networks through removal of weak dependencies," in *Proc. 10th Conf. Uncertainty Artif. Intell.*, CA, USA, 1994, pp. 374–382.
- [25] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951.
- [26] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 3, pp. 462–467, May 1968.
- [27] N. Friedman, I. Nachman, and D. Peer, "Learning Bayesian network structure from massive datasets: The 'sparse candidate' algorithm," in *Proc. 15th Conf. Uncertainty Artif. Intell. (UAI)*, 1999, pp. 206–215.



DENIS REILLY received the B.Eng. degree (Hons.) in electrical and electronic engineering and the M.Sc. degree (Hons.) in computer science and software engineering from the University of Liverpool, in 1989 and 1991, respectively, and the Ph.D. degree in distributed system and middleware from Liverpool John Moores University, in 2003.

He is currently a Principal Lecturer with the Department of Computer Science, Liverpool John Moores University. His research interests include

machine learning, data analytics, distributed systems and middleware, medical informatics, and cloud computing and security. In addition to his research interests he is a Course Leader for B.Sc. degree (Hons.) in computer studies and the B.Sc. degree (Hons.) in computer networks.



MARK TAYLOR received the B.Sc. degree (Hons.) in mathematics from Warwick University, U.K., in 1984, and the Ph.D. degree in computer science from Salford University, U.K., in 1999.

He is currently a Chartered IT Professional, a Chartered Engineer, and a Chartered Scientist. He is also a Senior Lecturer with the Department of Computer Science, Liverpool John Moores University, an Examiner of the British Computer Society, and an Academic Reference Group Member of

His Majesty's Inspectorate of Constabularies and Fire and Rescue Services, U.K. His research interests include data analytics, statistical modeling, risk analysis, and fire prevention.



PAUL FERGUS is currently a Professor in machine learning. He is also the Head of the Data Science Research Centre. He has competitively won external grants to support his research from HEFCE, Royal Academy of Engineering, Innovate U.K., Knowledge Transfer Partnership, North West Regional Innovation Fund and Bupa. He is also conducting research with Mersey Care NHS Foundation Trust looking at the use of smart meters to detect activities of daily living in people living alone with Dementia by monitoring the use of home appliances to model habitual behaviours for early intervention practices and safe independent living at home. His main research interests include machine learning for detecting and predicting preterm births. He is also interested in the detection of foetal hypoxia, electroencephalogram seizure classification and bioinformatics, such as polygenetic obesity, Type II diabetes, and multiple sclerosis. He has published over 200 peer-reviewed articles in these areas.



CARL CHALMERS is currently a Senior Lecturer with the Department of Computer Science, Liverpool John Moores University. He is also leading a three-year project on smart energy data and dementia in collaboration with Mersey Care NHS Trust. As a part of the project, a six month patient trial is underway within the NHS with future trials planned. The current trial involves monitoring and modelling the behaviour of dementia patients to facilitate safe independent living. He is also working in the area of high performance computing and cloud computing to support and improve existing machine learning approaches, while facilitating application integration. His main research interests include the advanced metering infrastructure, smart technologies, ambient assistive living, machine learning, high-performance computing, cloud computing, data visualisation, remote patient monitoring, and ICT-based healthcare.



STEVEN THOMPSON received the B.Sc. degree (Hons.) in information systems and the M.Sc. degree (Hons.) in wireless and mobile computing from Liverpool John Moores University, in 2008 and 2013, respectively, where he is currently pursuing the Ph.D. degree.

He is currently a Senior Technical Officer with the Department of Computer Science, Liverpool John Moores University, where he is performing part-time research. His research interests include

data analytics, machine learning, deep learning, and neural networks.

...