—技術報告—

犯罪者プロファイリングにおけるベイジアンネットワーク活用の有効性についての一考察
—屋内強姦における犯人の職業に関する確率的推定—

財津 亘^{*1}, 渋谷友祐^{*2}, 長谷川直宏^{*3}

富山県警察本部刑事部科学捜査研究所^{*1}

鳥取県警察本部刑事部科学捜査研究所^{*2}

愛知県警察本部刑事部科学捜査研究所^{*3}

〒930-8570 富山県富山市新総曲輪 1 番 7 号^{*1}

〒680-0911 鳥取県鳥取市千代水二丁目12番地^{*2}

〒460-8502 愛知県名古屋市中区三の丸二丁目 1 番 1 号^{*3}

**Effectiveness of Applying a Bayesian Network to the Offender Profiling:
Stochastic Inference of the Indoor-sex-offender's Employment**

Wataru Zaitu^{*1}, Yu-suke Shibuya^{*2} and Naohiro Hasegawa^{*3}

*Forensic Science Laboratory, Toyama Prefectural Police H.Q.^{*1}
1-7 Shinsogawa, Toyama-shi, Toyama 930-8570, Japan*

*Forensic Science Laboratory, Tottori Prefectural Police H.Q.^{*2}
2-12 Chiyomi, Tottori-shi, Tottori 680-0911, Japan*

*Forensic Science Laboratory, Aichi Prefectural Police H.Q.^{*3}
2-1-1 Sannomaru, Nakaku, Nagoya-shi, Aichi 460-8502, Japan*

(Received 20 November 2006; accepted 22 June 2007)

Offender profiling is one of the tools of decision making for criminal investigation. It is a set of techniques to infer characteristics of an unknown offender, such as sex, age bracket, lifestyle, psychological feature, previous crime, inhabited area, from the information which is left at the crime scene.

In this article, we proposed a tool of decision-making for criminal investigation from the perspective of prediction of an uncertain event by the use of a Bayesian Network (BN). BN is a probability model that describes causal structure of events as chain networks of conditional probability, and is capable to predict the possibility of uncertain events.

To examine the validity of the constructed model, firstly, we divided previous offenders' information of the indoor-sex-offence cases into a training data (9,859 cases) and validation data (50 cases). Secondly, we constructed a model from the training data by means of K2 and MDL (minimum description length) as search-algorithm and information criteria, respectively. Finally, the validity of the model was examined by the validation data as virtual cases.

According to the model, 21 target variables (16 behavioral variables, 2 vehicle variables and 3 victim variables) linked the explanatory variable (employment) directly, and most of these variables related to the employment. The results of the model validity showed that the accuracy of predicting the employment increased 10 % higher when the age bracket could be estimated from the testimony of the victim.

The results indicated that the BN model of the offender profiling would be able to provide valuable information for decision making for crime investigation. To predict characteristics of an unknown offender more accurately, it is crucial to select more appropriate information criteria and develop the search-algorithm, as well as to construct the database from more accurate information.

Key words: Offender Profiling, Sex Offences, Bayesian Network, K2-algorithm, MDL

緒 言

犯罪者プロファイリングとは、FBIのDouglas and Burgess (1986)によると「犯行の分析に基づいて、被疑者の性格特徴や行動特性を特定する捜査手法¹⁾とされている。また、警察庁では、「犯罪現場から得られた資料及び被害者に関する情報等から、犯人の性別、年齢層、生活スタイル、心理学的特徴、犯罪前歴の有無、居住地域等、犯罪捜査に役立つ情報を推定すること²⁾と定義しており、犯罪者プロファイリングは行動科学を捜査に応用した一手法であるとしている。

アメリカでは、犯罪者プロファイリングと呼称される1970年代以前は、精神科医などが犯罪者に対して様々な分析を行っていた³⁾。しかし、1970年代半ばからFBIが犯罪者に対する面接調査を始め、体系的な捜査手法である犯罪者プロファイリングを確立した。FBIは、犯罪者との面接調査や公的記録(犯罪記録、公判記録など)のデータから、性的殺人犯を、犯行現場の状況や犯人の行動によって秩序型と無秩序型に分類できることを見出し、類型論による手法を発展させたのである⁴⁾。これは、主に性的ファンタジーが犯行動機の背景にある事件を対象として、犯行現場の状況や犯人の行動を分析し、犯人の性格特徴や日常生活における行動様式を推定する手法といえる。

しかし、このような手法は次に述べる点において科学性が疑問視されるようになった。まず、FBIの手法では、「犯行現場は犯人の人格を反映してい

る⁵⁾ことを前提に、犯行現場全体を心の中でイメージし、犯人の人格や動機を推定する。このように直感的洞察を重視しているため、科学原則である整合性・実証性・再現性に乏しく、分析者の技量に依存しがちであることが挙げられる⁶⁾。また、FBIの手法の基礎となる面接調査については、面接技術における技量に個人差があることや、被面接者の供述内容にどの程度の信憑性があるのかという問題がある。その他、面接調査で得られた類型間の境界が曖昧であること、類型に犯罪者プロファイリングの対象となる当該事件を当てはめる方法が恣意的であること、年齢層や使用した乗物など犯人に関する事柄を推定した場合の妥当性が、犯人検挙に至らなければ分からないことが挙げられる。また、FBIの手法は精神医学や心理学を基盤とするため、具体的な年齢・職業・学歴などの犯人属性の推定よりも、心理的特徴や動機の推定を重視する。そのため推定された犯人像については、犯人の心理学的な説明が主であり、捜査支援情報としての有用性が低い。このように、FBIの手法の様々な問題点が浮び上がるなかで、1980年代半ばからイギリスではファセット理論と呼ばれる統計手法を基礎にした犯罪者プロファイリングが実施されるようになった。

ファセット理論とは、Louise Guttmanが提唱した、ファセットに基づいた構造による事前仮説を重要視する仮説検証型の多変量解析である⁷⁾。この理論と調査データに基づいた犯罪者プロファイリングが、英国リバプール大学のDavid Canterを中心に

開発されてきた⁸⁾。これらの手法では、犯罪者の諸行動間の関係を検討し、最小空間分析 (SSA-I) を用いることで犯罪行動を犯行テーマという大枠でとらえ、当該事件と類似の犯行テーマを持つ犯罪者群の特徴から当該事件の犯人属性を推定する⁹⁾。類似した手法に、数量化理論Ⅲ類¹⁰⁾などを用いた犯人像の推定が数多く日本で試みられている。これらの方法では、まず過去に検挙された犯罪者の犯罪行動や被害者属性または環境要因といった変数を空間的に布置する。続いて、互いに近接して布置された変数群に共通する犯行スタイルを見出し、それを基に犯罪者をいくつかの犯行テーマもしくは類型に分類するのである。数量化理論Ⅲ類などで得られた座標値を説明変数としたクラスター分析によって、各変数をいくつかのクラスターに分類し、それらを犯行テーマとすることもある。研究レベルでは、犯行テーマや類型間の犯人属性の違いを検討することが目的であることが多く、この点を χ^2 検定などの仮説検定によって検討することが多い。一方、実務レベルでは、犯行テーマや類型間の差異よりも、当該事件と類似した犯行特徴をもつ犯罪者群の犯人属性の傾向を検討することが重要になる。そのため、当該事件と類似した犯行特徴をもつ犯罪者群の犯人属性の比率を検討することが多いようである。

このようなテーマ (類型) 分析による方法は、犯罪者の犯罪行動や被害者属性または環境要因といった変数の構造や、それらと犯人属性との関係を理解する上では非常に有用であろう。しかし、これらの手法では、数種類に分類された犯行テーマや類型に犯人属性を対応させると犯人のタイプも数種類に限定されてしまうことや、現実には犯行テーマや類型に対応して明確に犯人属性が分類できることはきわめてまれであるといった問題点が指摘されている²⁾。犯罪者プロファイリングとは、いわば“不確定事象の予測”であり、これに関しては、より予測に適した手法を用いる必要があると考えられる。

このような観点から、1990年代後半には、テーマ (類型) 分析ではない様々な統計解析による犯人像推定のための手法が検討されている。例えば、決定木分析を用いたもの¹¹⁾やロジスティック回帰分析を用いたもの^{12,13)}が挙げられる。2002年にはイタリア

において Neural Network for Psychological Criminal Profiling (NNPCP) Project が始動し、ニューラルネットワークによる犯罪者プロファイリングのソフトウェア開発が進められている¹⁴⁾。“不確定事象の予測”という観点からは、本研究のテーマであるベイジアンネットワークもまた、犯罪者プロファイリングに適した手法であると考えられる。以下にベイジアンネットワークの概要及びベイジアンネットワークが犯罪者プロファイリングに適した手法の1つであると考えられる理由を述べる。

まず、ベイジアンネットワークとは、主観的確率を用いるベイズ統計学の中でも近年特に注目されている手法の1つである¹⁵⁻¹⁷⁾。現在、心理学の分野で主に用いられている t 検定や χ^2 検定などの仮説検定は、客観的確率による頻度論に基づいたものであるが、近年では有意水準設定の曖昧さなど様々な問題点^{18,19)}が指摘されている。それに対して、主観的確率を用いるベイズ統計学は、客観的確率の問題点に対処できうる統計学として、様々な分野で急速に普及しており、例えば、科学的根拠に基づく医療 (Evidence-Based Medicine: EBM) を推進する医療統計の分野²⁰⁾や意思決定^{21,22)}などの応用分野で積極的に活用されている。

ベイズ統計学の基礎概念は、Thomas Bayes の提唱したベイズの定理である^{23,24)}。ベイズの定理は条件付確率を求める式で、 $P(H|D) = P(D|H)P(H)/P(D)$ と表される。この式で、 H は仮説を、 D はデータを示している。 $P(H|D)$ はデータ D を得た後の仮説 H が真である確率を、 $P(H)$ はデータ D を得る前の仮説 H が真である確率を示し、それぞれ事後確率、事前確率という。また、 $P(D|H)$ は、「これが原因でデータ D が得られるだろう」という仮説 H の尤もらしさという意味から尤度といわれ、例えば「高熱の原因はインフルエンザである」という仮説 H の下で「インフルエンザの検査薬が陽性」というデータ D が得られる確率を示している。換言すると、ベイズの定理はあるデータが与えられた場合に、ある仮説が成り立つ確率を求める式であるといえる。

ベイジアンネットワークは、事象の因果構造を条件付確率の連鎖ネットワークで表現する確率モデル

である^{25,26)}。モデルは個々の変数であるノード、複数の変数間の定性的な依存関係を有向リンクで示したグラフ構造、個々の変数の定量的な関係を示した条件付確率で構成される。また、構築したモデル上で確率推論を行うことによって、不確実な事象の起こりやすさやその可能性を予測できる²⁶⁾。なお、確率推論とは観測された情報を使って、観測されていない情報の確率分布を求めることである。ベイジアンネットワークの特徴を以下に述べる。

まず、ベイジアンネットワークでは、共分散構造分析のように相関関係から因果構造を説明するのではなく、モデルそのものが因果構造を表している。そのため、ベイズ推定によって原因から結果、または結果から原因を確率的に推測できる。また、ニューラルネットワークのように構造がブラックボックスではなく、モデルをグラフィカルに表現できるため、定性的に理解しやすいといえる。モデルの構築に関しては、統計的学習機能を有しており、K2などの探索アルゴリズムやAIC (Akaike information criterion)やMDL (minimum description length)などの情報量基準を用いて、データベースからモデルを自動構築することが可能である。また、現在ある情報から合理的な意思決定を下すことができる“当座の結論”や、新しい情報の入力ごとにベイズ推定で判断を修正・更新することができる“逐次合理性”があり、新しい情報の追加学習が可能となる。さらに、データが無い場合でも専門知識によってモデルを構築できることが従来の統計手法とは大きく異なる。シミュレーションも可能なことから、モデル構築後の確率推論ではあらゆるパターンで情報を入力して、様々な角度から各変数内のカテゴリの確率やその変化を捉えることもできる。

ベイジアンネットワークには従来の統計手法との類似点も多い。例えば、共分散構造分析と同様に潜在変数を扱えることが挙げられる。また、2つ以上の変数間の関係を線形やロジスティック曲線などの関数によってモデル化するのではなく、決定木分析の樹形図のように変数のネットワークを構築する。そのため、自由度が比較的高く、従来の多変量解析(主成分分析や因子分析など)のような変数間の線形関係に基づくモデル化だけでなく、ニューラルネ

ットワークと同様に非線形関係にも適用可能である²⁶⁾。他に、ノンパラメトリック検定と同様に離散型変数やその母集団の正規性を必要としないことなども挙げられる。

このような特徴をもつベイジアンネットワークは、次に述べるように犯罪者プロファイリングにも有効であると考えられる。

第1に、テーマ(類型)分析では、逐次発生する当該事件の情報を追加・更新して再分析するには再解釈が必要となる。しかし、前述のとおり、ベイジアンネットワークでは新しい情報の追加学習が可能なことから、当該事件が連続発生した場合のデータ更新が容易である。また、当該事件に限らず将来発生するあらゆる事件の情報も逐次更新が可能のため、新しいタイプの事件にも対応することができる。

第2に、実際の事件では目撃者の証言や被害者の供述などから、例えば性別などの犯人属性についての情報が得られることがある。このように、事件情報を収集した時点で犯人属性について既知の情報がある場合、その情報を考慮することで、推定すべき犯人属性の推定結果の精度が高まることが期待される。これについては、クロス集計表を用いた検討も可能であるが、この方法は変数の増加に伴ってクロス集計表の作成に多くの時間を要するため、迅速性や効率性が求められる実務レベルでは不向きといえよう。しかし、ベイジアンネットワークでは、モデルに組み込んだ全ての変数を犯人属性推定に扱えることや、情報を得た後の手続きが変数の確率分布の設定と事後確率の算出のみであることから、迅速かつ効率的に実務に対応できる。

第3に、テーマ(類型)分析による方法では、犯人属性の推定に有効な変数を事前に把握していることが少なく、データベースからの変数選択が恣意的にならざるをえない。しかし、ベイジアンネットワークでは、ノード間のリンクは因果関係を示すため、いったん全ての変数を用いてモデルを構築すれば、各犯人属性と直接リンクする変数を、犯人属性の推定に有効な変数として選択することが可能である。また、手続きで後述する効果分析という手法から、各変数が犯人属性の推定に寄与する程度を知る

こともできる。さらに、モデル全体を概観することで、モデル構築に用いた全ての変数間の因果関係を知ることも可能なため、テーマ(類型)分析同様に、各変数の構造やそれらの関係性の理解にも有用である。

他に、犯罪者プロファイリングにベイジアンネットワークを用いる利点として、まずバイアスのない意思決定が可能な点が挙げられる。ベイジアンネットワークでは、過去の犯罪者のデータベースから算出される事前確率と当該事件の情報から得られる尤度によって犯人属性を確率的に推定するため、過去の経験に囚われることや、当該事件の情報を過剰に重視するという誤判断がなく、より適切な意思決定ができる。また、犯人属性の推定結果である確率は、そのまま予測の妥当性指標として用いることができるため、捜査員は、推定された各犯人属性の主観的確率を比較して、捜査上優先すべき犯人属性の順位付けに用いることができ、捜査支援に有効である。さらに、ベイジアンネットワークが意思決定の分野²²⁾において応用されてきたことが示すように、将来的にはベイジアンネットワークによって従来までの捜査員の経験や勘をモデル化することが可能であり、相対していた捜査員の経験とデータの統計的分析を融合することができると期待される。

このような理由から、近年デューク大学の LISC (Laboratory for Intelligent Systems and Controls) では犯罪者プロファイリングを題材に、ベイジアンネットワークの探索アルゴリズムの開発や研究が積極的に行われており²⁷⁾、次世代の手法として注目されている。

これらのことから、本研究では日本の犯罪者情報及び事件情報を用いてベイジアンネットワークを構築し、日本で犯罪者プロファイリングを実施することを想定して、モデルの検証を行った。本研究では屋内強姦事件を分析対象とした。その理由は、犯人と被害者が対面する事件では、被害者から犯人に関する目撃情報を得られる場合があることから、犯人の年齢層に関する情報がある場合とない場合で、犯人属性推定の精度を比較検討するためである。また、本研究では、犯人属性の推定に関する研究の一環として、まず犯人の職業に関する推定精度につい

て検討した。ベイジアンネットワーク構築には探索アルゴリズムに **K2**、情報量基準に **MDL** を用いた。**K2** はノード間の時間的順序についての条件を与えることで子ノードごとに親ノードの候補を限定し、グラフ構造を適度に探索するヒューリスティクスなアルゴリズムである。また、**MDL** はベイジアン情報量基準に一致もしくは漸近収束する情報量基準であり、予測などの応用に最適な基準とされているため²⁵⁾、犯罪者プロファイリングに適していると考えた。続いて、構築したモデルの精度を検証するために、モデル構築前にデータベースをモデル構築用学習データとモデル検証用評価データに分類し、評価データを仮想未解決事件として学習データから構築したモデルを評価する方法を採用した。

方 法

分析対象データ

警察庁保有の屋内強姦事件で検挙された犯罪者情報からデータベースを構築し、分析の対象とした(10,016件, 52変数)。また、データベースをモデル構築用の学習データ(9,966件)とモデル検証用の評価データ(50件)に分類した。年齢層別に分類した学習データ及び評価データの有職率、つまり有職者のベースレートを Table 1 に示す。学習データ中 107件で職業情報が欠損していたため、Table 1 に示す学習データは当該107件を除いた全9,859件であった。なお、この学習データのベースレートはベイ

Table 1 Base rates of employment according to the age bracket in the training data and the validation data.

年齢層	学習データ (n=9859) における有職率	評価データ (n=50) における有職率
10代	41.3% (294/712)	33.3% (1/3)
20代	77.4% (2752/3556)	70.8% (17/24)
30代	81.8% (2490/3043)	90.9% (10/11)
40代以上	78.9% (2010/2548)	83.3% (10/12)
各データの 全体の割合	76.5% (7546/9859)	76.0% (38/50)

ジアンネットワークにおける事前確率にあたる。
ソフトウェア

モデル構築及び検証には、株式会社数理システム製ベイジアンネットワーク構築支援システム Bayo-Net ver. 3.0を用いた。

手続き

学習データの犯罪行動に関する43変数（偽計，目を覆うなど），犯行時使用した乗物に関する3変数（自動車，バイク，自転車使用），被害者属性に関する4変数（対象者が幼児・年少者，女子学生，カップル，高齢者），犯人属性に関する2変数（年代，職業）から，次のようにベイジアンネットワークモデルを構築した。まず，犯人属性変数が説明変数，行動・乗物・被害者変数が目的変数になるように，また，同種の変数間では，時間的先行の関係から，例えば証拠隠滅変数が時間的に先行する接近方法変数の親ノードにならないように制約するという手順でモデル構造の概要設計を行った。概要設計後，探索アルゴリズムに K2，モデル選択基準に MDL を用いてモデルを構築し，MDL を基準に因果構造の修正を行った。

続いて50件の評価データを仮想未解決事件とし，それらの行動・乗物・被害者変数から犯人の職業（有職もしくは無職・学生）の推定をすることでモデル検証を行った。検証方法はハードエビデンスで行い，仮想未解決事件で確認されている行動変数の確率を1に固定し，サンプリング法^{25,26)}により確率推論を行った。また，乗物変数と被害者変数に関しては，確認された変数の確率を1，それ以外の確率を0に固定して確率推論を行った。例えば，自動車の使用が確認されている場合は，自動車変数の確率を1に，自動車以外の乗物変数は0に固定した。ベイジアンネットワークでは，確認されたデータを1に固定して確率推論を行うのが通常である。そのため，犯罪行動などでは確認されていなかった変数を強制的に0に固定するのは不自然であるため行わなかった。ただし，乗物などは車に乗っていたことが確認できれば，他の乗物を使用した可能性が低くなるため，本研究では前述の方法を採用した。職業（有職もしくは無職・学生）の推定は，確率推論前の学習データのベースレートである事前確率と確率

推論後の事後確率との比較によって行った。具体的には，職業変数において有職の事後確率が事前確率に比べて高くなった場合を有職，低くなった場合を無職・学生とした。

以上の検証方法で，犯人の年代が分かっている場合の犯人属性推定における精度も同時に検証した。具体的には，上述の手続きに加えて，年代変数において確認された年代カテゴリーの確率分布を1に固定して確率推論を行った。

また，犯人属性以外の各変数がどの程度職業変数の推定に有効であるかを検討するために効果分析をした。効果分析では，まず犯人属性以外の1変数の確率を1に固定してサンプリング法によって確率推論を実施した。続いて，確率推論の結果得られた職業変数の事後確率とベースレートである事前確率の差を検討することにより，確率を1に固定した変数が職業推定に与える影響の程度について分析した。

結果と考察

構築したモデルを Fig. 1 に示す。モデル全体の MDL は150,381.9，ノード平均の MDL は2,837.4であった。モデルを概観すると，行動変数のうちでは，飲酒・薬物等摂取，潜伏，誘い出し，自称，会話，首を絞める，手足の暴力，火薬物による威嚇，指挿入，口内射精，被害者を全裸，性具等の利用，異常体位，肛門挿入，複数回の性的行為，金品を渡すという16変数が職業変数に直接リンクしたことが分かる。また，乗物変数では，自動車使用，自転車使用の2変数が，被害者変数では，女子学生対象，カップル対象，幼児・年少者対象の3変数が職業変数に直接リンクした。

次に，職業変数の推定に対する行動・乗物・被害者変数の効果分析の結果を Table 2 に示す。効果分析の結果，職業変数に直接リンクしている変数のほとんどが，他の変数と比較して，職業変数の推定における影響力が高いことが示された。ただし，スプリー犯行（一昼夜で連続して犯行に及ぶこと），バイク使用，高齢者対象のように職業変数と直結していない変数でも，職業変数に影響することも示された。乗物・被害者変数の全てが，他の変数と比較して，職業変数の推定に対して高い値を示しているこ

Fig. 1 Bayesian network model of the indoor-sex-offence cases.

Table 2 Results of effect-analysis of criminal behavior/vehicle/victim variables in prediction of employment.

有職である確率を高くする変数		職業変数に影響のない変数		無職である確率を高くする変数	
幼児・年少者対象	0.0682	姦淫時の首絞め	0	カップル対象	-0.1822
自動車使用	0.0467	フェラチオ	0	自転車使用	-0.1621
飲酒・薬物等摂取	0.0403	凶器威嚇	0	火薬物による威嚇	-0.0650
首を絞める	0.0208	布等による威嚇・暴力	0	バイク使用	-0.0615
スプリー犯行	0.0169	燃焼具所持	0	女子学生対象	-0.0541
金品を渡す	0.0154			潜伏	-0.0338
指挿入	0.0123			複数回の性的行為	-0.0333
高齢者対象	0.0100			口内射精	-0.0320
会話	0.0081			自称	-0.0196
身体を舐める	0.0043			犯行後の連絡	-0.0126
追跡	0.0034			甘言接触	-0.0103
身分証等強取	0.0028			誘い出し	-0.0101
被害者を全裸	0.0023			肛門挿入	-0.0085
凶器所持を装う	0.0022			脅威による接触	-0.0071
性具等の利用	0.0021			異常体位	-0.0051
クニニリングス	0.0018			目を覆う	-0.0050
素手の暴力	0.0014			再会の約束	-0.0042
偽計侵入	0.0013			写真撮影等	-0.0029
犯人が全裸	0.0009			異物挿入	-0.0028
親切な言動	0.0007			下見	-0.0004
就寝中の被害者	0.0006			手足の暴力	-0.0002
無締り箇所侵入	0.0001			偽計	-0.0001
刃物類所持	0.0001				

※数値は有職の事後確率から事前確率を差し引いた値を示す。

とは非常に興味深い．また，Fig. 1によると，これらの変数のほとんどが犯人属性変数のみとリンクしており，わずかに自動車使用とスプリー犯行，女子学生対象と甘言接触が，それぞれリンクしているのみであった．この点に関しては，乗物・被害者変数は犯人が任意に選択することや，犯人と被害者の行動の相互作用の中で変化しない変数であるために，行動変数とは独立に犯人属性変数とリンクしたものと解釈できる．このように，ベイジアンネットワークでは，構築したモデルから変数間の関係を容易に理解できることが利点の1つといえよう．また，モデル構築によって，犯人属性と連関のある変数を探索するだけでなく，効果分析によって各変数が犯人属性の推定に寄与する程度を定量的に知ることも可能である．

仮想未解決事件50件でモデルの精度を検証した結果，職業の推定的中率は年代情報がない場合で74%，年代情報がある場合で88%であった．これらの推定的中率の内訳を Fig. 2 に示す．年代情報がない場合の38件の有職者を正しく有職と推定できた割合は81.6%，12件の無職者や学生を正しく無職・学生と推定できた割合は50.0%であった．一方，年代情報がある場合では，有職者を正しく有職と推定できた割合は94.7%，無職者や学生を正しく無職・学生と推定できた割合は66.7%であった．

本研究で扱った職業変数では，評価データの有職者のベースレートが76%であったため，仮に仮想未解決事件全てで有職と推定すると，仮想未検挙事件50件の推定的中率は76%となり，本研究における年

代情報がなかった場合に比べて推定的中率が高くなってしまう．このように，ベースレートの確率分布が偏っている変数を推定する場合，確率分布が偏って高い方のカテゴリを選択することでの的中率が高くなる．そのため，今後は推定対象となる属性のベースレート（事前確率）の偏りの程度と，犯人属性推定の精度の関係について検討を進める必要がある．また，モデルの精度に関しては，モデル構築の際に採用する探索アルゴリズムや情報量基準によって構築されるモデルが異なることや，確率推論の際に採用するアルゴリズムによって推定結果が異なることも予想される．例えば，Baumgartner et al. (2005)²⁷⁾は，K2を改良したK'2を犯罪者プロファイリングに適用し，精度が15%向上したと報告している．確率推論アルゴリズムについては，本研究で採用したサンプリング法は，データ数が多ければ真の解に漸近するが，ノード数やノードのカテゴリが増加すると，必要なデータ数が指数関数的に増大するとされている²⁵⁾．そのため，本研究ではノード数もしくはノードのカテゴリ数が多かったため，必要なデータ数が足りず，精度が低かったことが考えられる．このことから，Loppy belief propagation²⁶⁾のように大規模ネットワークに対しても高速に計算し，真の解への収束が保証されるようなアルゴリズムを採用しての検討も求められる．

本研究では，有職者を正しく有職と推定できた割合，無職者・学生を正しく無職・学生と推定できた割合は，ともに年代情報がない場合より年代情報がある場合の方が10%以上高かった．この結果は，被

Fig. 2 Correct rates of predict employment in the validation data (50 cases).

害者などの目撃証言から年代情報を得ることによって職業に関する予測精度が向上することを意味しており、犯罪者プロファイリングにおいても正確な目撃情報を得ることの重要性を示唆している。

また、本研究では事前確率からの変化の方向性を検討することで、有職と無職・学生を推定した。このような手続きを行った理由は、学習データのベースレートである事前確率から事後確率への変動が小さく、確率推論後も有職者である確率が高い値に固定されてしまい、事後確率を直接評価したのでは、職業の推定がうまくいかなかったためである。このような現象が生じた原因については、前述のように、仮想未解決事件全てを有職と推定しても高い推定的中率を維持できるため、有職の事後確率が高い値に固定されてしまった可能性が考えられる。犯罪者プロファイリングでは、有職者は有職と、無職者・学生であるならば無職・学生と正確に推定することが目的であることから、本研究では仮想未解決事件において確認された情報が示す方向性によって職業の推定を行った。しかし、本研究のような推定方法では、事後確率を予測の妥当性指標とすることは困難である。そのため、今後は事後確率そのものを推定結果とできる条件（事前確率の偏りの程度など）について検討する必要がある。

確率推論の方法として、本研究では確率分布を1または0に固定するハードエビデンスで確率推論を行った。これは、犯行が1件の場合に適用される手法である。しかし、現実の犯罪では事件が連続発生することもあり、連続犯行において犯人は常に同じ犯罪行動をとるとは限らない。そのような場合には、任意の確率分布に固定するソフトエビデンスを用いることで、より精度の高い確率推論ができるものと思われる。具体的には、同一犯によるとと思われる犯行のうち4件中3件で凶器が使用されている場合、確率を3/4に固定し確率推論することで、連続発生する当該事件の情報をより正確に捉え予測精度の高い分析ができると考えられる。

本研究で構築したモデルは、精度について多くの課題がある。しかし、ベイジアンネットワークではモデルを柔軟に変容させることで、精度の高いモデルを構築することができることから、K'2を含め新

しく提案される探索アルゴリズムを用いてモデル構築を行ったり、確率推論についても **Loppy belief propagation** やソフトエビデンスを用いることで精度の向上が期待できる。ただしその際には、前述のように、推定の対象となる変数のベースレートや、今回用いたデータベースにない変数についても併せて検討する必要がある。

また、ベイジアンネットワークの特徴の1つであるデータベースを用いないモデル構築を利用して、経験豊富な捜査員の意思決定モデルを独立的に構築したり、あるいはそれをデータベースによるモデルと比較することで、両者を組み込んだモデルを構築することも可能である。さらに、情報が蓄積され膨大になっていく事件データや、当該事件で随時得られる情報も、ベイジアンネットワークではモデルさえ構築すれば、即座に分析に用いることができる。そのため、従来の多変量解析を用いた手法に比べ、迅速かつ効率的な犯罪者プロファイリングの実施が期待できる。加えて、ベイジアンネットワークは犯行予測の手法としても注目されている²⁸⁾。例えば、連続して犯行に及んだ犯罪者のデータから得られる1件目と2件目の犯行の間隔という時間情報を、犯罪行動や犯人属性の変数とともに組込んだモデルを構築することで、当該事件における1件目の行動形態から2件目の犯行発生期間をある程度推定することができる。このように、ベイジアンネットワークは、従来の多変量解析による方法に比べて汎用性が高いことから、今後ますます期待される分析手法といえよう。

文 献

- 1) Douglas, J. E. and Burgess, A. E.: Criminal profiling: A viable investigative tool against violent crime. *FBI Law Enforcement Bulletin*, 55(12), 9-13 (1986).
- 2) 渡邊和美, 高村 茂, 桐生正幸: 行動科学と情報分析からの多様なアプローチ 犯罪者プロファイリング入門. 北大路書房, 京都 (2006).
- 3) Turvey, B. E.: Criminal profiling: An introduction to behavioral evidence analysis. Academic Press, London (1999).

- 4) FBI: Crime scene and profile characteristics of organized and disorganized murders. *FBI Law Enforcement Bulletin*, 54(8), 18–25 (1985).
- 5) Holmes, R. H. and Holmes, S. T.: Profiling violent crimes: An investigative tool. 2nd edition. Sage Publications, Thousand Oaks, CA (1996).
- 6) Godwin, M.: Criminal psychology and forensic technology: A collaborative approach to effective profiling. CRC Press, Boca Raton, FL (2001).
- 7) 木村通治, 真鍋一史, 安永幸子, 横田賀英子: ファセット理論と解析事例—行動科学における仮説検証・探索型分析手法—. ナカニシヤ出版, 東京 (2002).
- 8) Canter, D., Heritage, R. and Kovackit, M.: *A facet approach to offender profiling*. Final Report to the Home Office, University of Surrey, Psychology Department, 1 and 2 (1991).
- 9) Canter, D. and Heritage, R.: A multivariate model of sexual offence behavior: Developments in offender profiling. *Journal of Forensic Psychiatry*, 1, 185–212 (1990).
- 10) 高村茂, 徳山孝之: 多変量解析を用いた犯罪者プロファイリング研究—侵入窃盗犯の類型化に着目して—. *犯罪心理学研究*, 43(2), 29–43 (2005).
- 11) 田口真二, 荘島宏二郎: 犯行を繰り返す強姦犯罪者の特徴. 日本心理学会第69回大会発表論文集, 412 (2005).
- 12) Davies, A., Wittebrood, K. and Jackson, J. L.: Predicting the criminal antecedents of a stranger rapist from his offence behavior. *Science & Justice*, 37, 161–170 (1997).
- 13) Davies, A., Wittebrood, K. and Jackson, J. L.: Predicting the criminal record of a stranger rapist. (Special Interest Series: Paper 12). Police and Reducing Crime Unit, Home Office, London (1998).
- 14) Strano, M.: A neural network applied to criminal psychological profiling: An Italian initiative. *International Journal of Offender Therapy and Comparative Criminology*, 48(4), 495–503 (2004).
- 15) Neapolitan, R. E.: Learning bayesian networks. Pearson Education Inc, Pearson Prentice hall, Upper Saddle River (2003).
- 16) スチュワート・ラッセル, ピーターノーヴィグ, 古川康一(監): エージェントアプローチ: 人工知能. 共立出版, 東京 (1997).
- 17) 渡辺澄夫, 荻原克幸, 赤穂昭太郎, 本村陽一, 福水健次, 岡田真人, 青柳美輝: 学習システムの理論と実現. 森北出版, 東京 (2005).
- 18) 佐伯 胖, 松原 望: 実践としての統計学. 東京大学出版会, 東京 (2002).
- 19) 繁桝算男, 柳井晴夫, 森敏昭: Q & A で知る統計データ解析 DOs and DON'Ts. サイエンス社, 東京 (2004).
- 20) Congdon, P.: Bayesian statistical modelling. John Wiley & Sons, Chichester (2001).
- 21) 鈴木雪夫, 国友直人: ベイズ統計学とその応用. 東京大学出版会, 東京 (2003).
- 22) 藤本和則, 本村陽一, 松下光範, 庄司裕子: 意思決定支援とネットビジネス. オーム社, 東京 (2005).
- 23) 渡部 洋: ベイズ統計学入門. 福村出版, 東京 (1999).
- 24) 繁桝算男: ベイズ統計入門. 東京大学出版会, 東京 (1985).
- 25) 繁桝算男, 植野真臣, 本村陽一: ベイジアンネットワーク概説. 培風館, 東京 (2006).
- 26) 本村陽一, 岩崎弘利: ベイジアンネットワーク技術 ユーザ・顧客のモデル化と不確実性推論. 東京電機大学出版局, 東京 (2006).
- 27) Baumgartner, K. C., Ferrari, S. and Salfati, C. G.: Bayesian network modeling of offender behavior for criminal profiling. Proc. of the 44th IEEE Conference on Decision and Control, and the European Control Conference 2005 (2005).
- 28) Oatley, G. C. and Ewart, B. W.: Crimes analysis software: 'pins in maps', clustering and bayes net prediction. *Expert Systems with Applications*, 25, 569–588 (2003).