

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5519210号
(P5519210)

(45) 発行日 平成26年6月11日 (2014. 6. 11)

(24) 登録日 平成26年4月11日 (2014. 4. 11)

(51) Int. Cl. F I
G06N 5/04 (2006.01)
 G06N 5/04 580A
 G06N 5/04 550J

請求項の数 4 (全 13 頁)

(21) 出願番号	特願2009-185726 (P2009-185726)	(73) 特許権者	598001261
(22) 出願日	平成21年8月10日 (2009. 8. 10)		株式会社エイチアイ
(65) 公開番号	特開2011-39734 (P2011-39734A)		東京都目黒区東山 1 丁目 4 番 4 号
(43) 公開日	平成23年2月24日 (2011. 2. 24)	(74) 代理人	100079005
審査請求日	平成24年7月6日 (2012. 7. 6)		弁理士 宇高 克己
前置審査		(72) 発明者	竹内 俊治
			東京都目黒区東山 1 丁目 4 番 4 号 株式会
			社エイチアイ内
		(72) 発明者	末永 貴一
			東京都目黒区東山 1 丁目 4 番 4 号 株式会
			社エイチアイ内
		(72) 発明者	川端 一生
			東京都目黒区東山 1 丁目 4 番 4 号 株式会
			社エイチアイ内

最終頁に続く

(54) 【発明の名称】 モデル生成装置、モデル生成方法、情報処理システム、及びプログラム

(57) 【特許請求の範囲】

【請求項 1】

実測データの蓄積状況に応じて事前確率分布を算出する関数のパラメータを設定する設定手段と、

前記設定されたパラメータを用いて算出した事前確率分布に基づいて確率変数間の条件付依存性を決定し、この決定した条件付依存性に基づいて構築した前記確率変数間の相関関係に基づいてベイジアンネットワークのモデルを生成するモデル生成手段とを有し、

前記モデル生成手段は、前記実測データの蓄積量が前記条件付依存性を設定するには少ない量である場合、及び、実測データが所定の期間より古い場合の少なくとも一方である場合に、所定の条件付依存性に基づいて前記確率変数間の相関関係を構築することを特徴とするモデル生成装置。

【請求項 2】

実測データの蓄積状況に応じて事前確率分布を算出する関数のパラメータを設定する設定手段と、

前記設定されたパラメータを用いて算出した事前確率分布に基づいて確率変数間の条件付依存性を決定し、この決定した条件付依存性に基づいて構築した前記確率変数間の相関関係に基づいてベイジアンネットワークのモデルを生成するモデル生成手段とを有し、

前記モデル生成手段は、前記実測データの蓄積量が前記条件付依存性を設定するには少

ない量である場合、及び、実測データが所定の期間より古い場合の少なくとも一方である場合に、所定の条件付依存性に基づいて前記確率変数間の相関関係を構築することを特徴とする情報処理システム。

【請求項3】

実測データの蓄積状況に応じて事前確率分布を算出する関数のパラメータを設定する設定ステップと、

前記設定されたパラメータを用いて算出した事前確率分布に基づいて確率変数間の条件付依存性を決定し、この決定した条件付依存性に基づいて構築した前記確率変数間の相関関係に基づいてベイジアンネットワークのモデルを生成するモデル生成ステップとを有し、

前記モデル生成ステップは、前記実測データの蓄積量が前記条件付依存性を設定するには少ない量である場合、及び、実測データが所定の期間より古い場合の少なくとも一方である場合に、所定の条件付依存性に基づいて前記確率変数間の相関関係を構築することを特徴とするモデル生成方法。

【請求項4】

情報処理装置のプログラムであって、前記プログラムは前記情報処理装置に、実測データの蓄積状況に応じて事前確率分布を算出する関数のパラメータを設定する設定処理と、

前記設定されたパラメータを用いて算出した事前確率分布に基づいて確率変数間の条件付依存性を決定し、この決定した条件付依存性に基づいて構築した前記確率変数間の相関関係に基づいてベイジアンネットワークのモデルを生成するモデル生成処理とを実行させ、

前記モデル生成処理は、前記実測データの蓄積量が前記条件付依存性を設定するには少ない量である場合、及び、実測データが所定の期間より古い場合の少なくとも一方である場合に、所定の条件付依存性に基づいて前記確率変数間の相関関係を構築することを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ベイジアンネットワークのモデルを生成する装置、方法、及びそのプログラムに関する。

【背景技術】

【0002】

近年、ユーザの嗜好や意図を推定して、ユーザが所望する動作やサービスを提供する機器のニーズが高まっている。

【0003】

このように、ユーザが所望する動作やサービスを推定するために、ベイジアンネットワーク（Bayesian Network）が用いられている（例えば、特許文献1）。ベイジアンネットワークとは、統計データに基づいて様々な事象の因果関係や依存関係をベイズの定理を用いて表現するものであり、複雑な因果関係や依存関係を有向グラフによって表すとともに、個々の定性的な関係を条件付確率で表す確率推論のモデルである。特許文献1には、被験者に対してアンケート等の聞き取り調査を実施して集めたデータに基づいてベイジアンネットワークのモデルを作成し、作成したモデルにユーザの状況を照らし合わせることでユーザの嗜好を推定し、コンテンツを推薦する技術が記載されている。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2007-058398号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、特許文献1では、一般的な統計データに基づいて作成したベイジアンネットワークモデルを用いてユーザの嗜好を推定しているため、実際にはユーザ個人の嗜好とずれが生じてしまうという問題がある。

【0006】

また、ベイジアンネットワークは、多種多様に富んだ離散的なデータを定性的に捉えることにより各事象の複雑な関係を導き出し、不確実な事象の起こりうる可能性を推定するものである。そのため、ユーザが所望する動作やサービスをユーザ毎に推定するには、該ユーザに関する情報を十分な量まで蓄積させなければならない。従って、ベイジアンネットワークを個人に特化させて用いるのには向いていないことになる。

【0007】

そこで、本願発明が解決しようとする課題は、上記問題点を解決することにより、ベイジアンネットワークを個人に特化させて用いることができる技術を提供することにある。

【課題を解決するための手段】

【0008】

上記課題を解決する本発明のモデル生成装置は、実測データの蓄積状況に応じて事前確率分布を算出する関数のパラメータを設定する設定手段と、前記設定されたパラメータを用いて算出した事前確率分布に基づいて確率変数間の条件付依存性を決定し、この決定した条件付依存性に基づいて構築した前記確率変数間の相関関係に基づいてベイジアンネットワークのモデルを生成するモデル生成手段とを有し、前記モデル生成手段は、前記実測データの蓄積量が前記条件付依存性を設定するには少ない量である場合、及び、実測データが所定の期間より古い場合の少なくとも一方である場合に、所定の条件付依存性に基づいて前記確率変数間の相関関係を構築することを特徴とする。

【0009】

上記課題を解決する本発明の情報処理装置は、実測データの蓄積状況に応じて事前確率分布を算出する関数のパラメータを設定する設定手段と、前記設定されたパラメータを用いて算出した事前確率分布に基づいて確率変数間の条件付依存性を決定し、この決定した条件付依存性に基づいて構築した前記確率変数間の相関関係に基づいてベイジアンネットワークのモデルを生成するモデル生成手段とを有し、前記モデル生成手段は、前記実測データの蓄積量が前記条件付依存性を設定するには少ない量である場合、及び、実測データが所定の期間より古い場合の少なくとも一方である場合に、所定の条件付依存性に基づいて前記確率変数間の相関関係を構築することを特徴とする。

【0010】

上記課題を解決する本発明のモデル生成方法は、実測データの蓄積状況に応じて事前確率分布を算出する関数のパラメータを設定する設定ステップと、前記設定されたパラメータを用いて算出した事前確率分布に基づいて確率変数間の条件付依存性を決定し、この決定した条件付依存性に基づいて構築した前記確率変数間の相関関係に基づいてベイジアンネットワークのモデルを生成するモデル生成ステップとを有し、前記モデル生成ステップは、前記実測データの蓄積量が前記条件付依存性を設定するには少ない量である場合、及び、実測データが所定の期間より古い場合の少なくとも一方である場合に、所定の条件付依存性に基づいて前記確率変数間の相関関係を構築することを特徴とする。

【0011】

上記課題を解決する本発明は、情報処理装置のプログラムであって、前記プログラムは前記情報処理装置に、実測データの蓄積状況に応じて事前確率分布を算出する関数のパラメータを設定する設定処理と、前記設定されたパラメータを用いて算出した事前確率分布に基づいて確率変数間の条件付依存性を決定し、この決定した条件付依存性に基づいて構築した前記確率変数間の相関関係に基づいてベイジアンネットワークのモデルを生成するモデル生成処理とを実行させ、前記モデル生成処理は、前記実測データの蓄積量が前記条件付依存性を設定するには少ない量である場合、及び、実測データが所定の期間より古い場合の少なくとも一方である場合に、所定の条件付依存性に基づいて前記確率変数間の相関関係を構築することを特徴とする。

10

20

30

40

50

【発明の効果】

【0012】

本発明によると、十分な量のデータが蓄積されていなくとも、ベイジアンネットワークのモデルを生成することができる。

【図面の簡単な説明】

【0013】

【図1】本発明の第1の実施の形態のブロック図である。

【図2】本発明のアルゴリズムを説明するための図である。

【図3】本発明のフローチャートである。

【図4】パラメータと事前確率分布との関係を示す図である。

10

【図5】本発明の第2の実施の形態のブロック図である。

【発明を実施するための形態】

【0014】

一般的に、発生する様々な事象は単独で発生することもあるが、原因となる事象があり、その原因となる事象の影響で発生することが多い。

【0015】

例えば、事象 X^1 が「午前／午後」であり、事象 X^2 が「今日の天気予報のコンテンツを見る／見ない」であり、事象 X^3 が「音楽ツールを使用する／しない」であったとする。

【0016】

20

通常のベイジアンネットワークのモデルの作成では、

- ・ 900回、午前に今日の天気予報を見た
- ・ 100回、午後に今日の天気予報を見た
- ・ 500回、午前に音楽ツールを使用した
- ・ 500回、午後に音楽ツールを使用した

というような実測値（サンプルデータ）から算出した条件付確率変数に基づいて、事象 X^1 の事象の変化が事象 X^2 の事象の発生及び事象 X^3 の事象の発生に関係しているかを判断して、事象 X^1 と X^2 と X^3 との依存関係を作成する。因みに、この一例の場合、今日の天気予報のコンテンツの視聴が時間帯によって大きく変動しているため事象 X^1 と X^2 との依存関係が強いとし、音楽ツールの使用は時間帯によっては変動していないため事象 X^1 と X^3 との依存関係は無いとなる。

30

【0017】

しかしながら、例えば、

- ・ 3回、午前に今日の天気予報を見た
- ・ 2回、午後に今日の天気予報を見た
- ・ 1回、午前に音楽ツールを使用した
- ・ 4回、午後に音楽ツールを使用した

というように実測値の数が少ない場合、実際は上記一例のように事象 X^1 と X^2 との依存関係が強く、事象 X^1 と X^3 との依存関係は無いにもかかわらず、本例では事象 X^1 と X^2 との依存関係は無く、事象 X^1 と X^3 との依存関係はあるというように実際とは異なる依存関係できてしまう可能性がある。

40

【0018】

また、

- ・ 先月は900回、午前に今日の天気予報を見た
- ・ 先月は100回、午後に今日の天気予報を見た
- ・ 今月は10回、午前に今日の天気予報を見た

というように実測値の分布が時間の経過によって急激に変化し、現在では事象 X^1 と X^2 との依存関係が弱くなったにも関わらず、本例では事象 X^1 と X^2 との依存関係が強いとなってしまう、やはり実際とは異なる依存関係できてしまう可能性がある。

【0019】

50

そこで、本発明は、ベイジアンネットワークのモデルの作成対象である各事象（確率変数）間の依存関係を、各事象の実測値の蓄積状況に応じて設定する。

【0020】

本発明の特徴を説明するために、以下において、図面を参照して具体的に述べる。

【0021】

〈第1の実施の形態〉

本発明を実施するための第1の実施の形態について図面を参照して詳細に説明する。

【0022】

図1は、本願発明のモデル生成装置100のブロック図である。

【0023】

情報収集部101は、ベイジアンネットワークのモデルを作成するための各種事象の観測値を収集する。収集した情報は、事象毎に統計データ102に記憶するが、このとき、連続値データを意味を持ったまとまりにするために、標本化又は量子化等を行ってクラスタリング（離散化）して記憶する。情報収集部101が収集する観測値は、例えば、ユーザの位置情報、時刻情報、気象情報、ユーザの動作又はユーザが装置に対して行った操作のログ情報等、変数又は変動を有する情報である。情報の収集は、新規データが観測された時に行っても、間欠的に行っても良く、収集する情報の性質に合わせて適宜行われる。

【0024】

統計データ102は、情報収集部101にて収集された各種データが事象毎に記憶される。

【0025】

基準データ103には、予め設定された各事象間の相関関係（依存性）が基準データとして記憶されている。この基準データは、マーケティングデータから生成されたベイジアンネットワークのモデルであっても良いし、モデル生成装置提供者が生成したベイジアンネットワークのモデルであっても良い。

【0026】

ベイジアンネットワークモデル105には、モデル生成部104にて生成されたベイジアンネットワークのモデルが記憶される。

【0027】

モデル生成部104は、設定部1041及び構築部1042を有する。設定部1041は、統計データ102に記憶されているデータの蓄積状態に応じて各事象間の依存性を設定する。構築部1042は、設定部1041が設定した依存性を用いて各事象間の相関関係を構築してベイジアンネットワークのモデルを生成する。そして生成したモデルをベイジアンネットワークモデルに記憶する。

【0028】

本発明におけるモデル作成のアルゴリズムについて説明する。

【0029】

モデル生成のアルゴリズムとしては、図2に示すアルゴリズムになる。このアルゴリズム内のcurrentGraphは、現在のベイジアンネットワークのモデルを示す。一方、TemporaryGraphは、直前のベイジアンネットワークのモデルを示す。図2に示したアルゴリズムは、統計データ102に格納された新規のデータをセットすると、

A1：ある1つの事象に注目する。

A2：新規のデータを含めた状態で、その事象と相関関係を有している事象を削除又は追加することにより相関関係を構築する。

A3：n個ある全事象各々に対して、A2を実行する。

A4：構築した相関関係を用いてモデルを生成する。

A5：生成したモデルと新規データを含める前のモデルとを比較し、これらが同じになるまで、即ち、新規のデータに対して生成する。

となる。

10

20

30

40

50

【0030】

図3は、このアルゴリズムをフローチャートで表したものである。尚、以下の動作は、新規データが所定量記憶される度に開始しても、新規データが記憶される度に開始しても、予め決められた時間に開始してもよい。また、設定部1041を複数用意し、異なる事象に対して並列に動作してもよい。

【0031】

まず、統計データ102に記憶された新規データをモデルに追加するかを判定する（ステップS1）。新規データの追加は、生成するモデルや測定する情報の性質又は特徴によって、例えば事象毎に追加しても、相関関係がある複数の事象をグループ化し、そのグループ毎にデータを追加してもよいし、全新規データを追加してもよい。

10

【0032】

変数gに、ベイジアンモデル105に記憶されているモデルを現在のベイジアンネットワークのモデルcurrentGraphとして格納する（ステップS2）。

【0033】

変数gと、作成途中のモデルであるTemporaryGraphが格納される変数g'とが一致しているかを判定する（ステップS3）。尚、作成途中のモデルが無い場合は、変数g'にモデルが格納されていないので一致しないと判定される。

【0034】

ステップS3において一致していないと判定されると、変数gに変数g'を格納する（ステップS4）。

20

【0035】

注目している事象に相関する事象を追加及び削除して、注目している事象の相関関係を構築する（ステップS5）。

【0036】

構築した相関関係を元に、ベイジアンネットワークのモデルを生成し、これをTemporaryGraphとし、変数g'に、TemporaryGraphを格納する（ステップS6）。

【0037】

全事象に対して、ステップS2及びS3を実行する（ステップS7）。

【0038】

ステップS3で一致しないと判定された場合、即ち、追加した新規データを考慮したベイジアンネットワークのモデルが出来ると、ステップS1に戻り、新規データを追加するかを判定し、追加しない場合はTemporaryGraphをcurrentGraphとして格納してベイジアンモデル105に記憶させる（ステップS8）。

30

【0039】

ここで、ステップS5について、詳細に説明する。

【0040】

ベイジアンネットワークは、その名前が示すとおり、ベイズの定理を用いて各観測を表す事象間、即ち確率変数間の因果関係（依存関係）を有向辺からなるネットワークとして表現し、確率推論を行う技術である。

【0041】

40

統計データ102に、事象がN個あるとすると、確率変数の集合である統計データ全体は、 $X^N = X^{(1)} X^{(2)} \cdots X^{(N)}$ と表すことができる。各確率変数は、データの集合である。各確率変数のデータの集合は、

$$X^{(i)} = \{X_1^{(i)}, \dots, X_n^{(i)}\} \in V, i \in \{1, \dots, N\}$$

と表すことができ、これを単一サンプルデータと呼ぶ。そして、nは、単一サンプルデータのサンプルサイズを表す。Vはデータの分散を示し、

$$V = V_1 \times \cdots \times V_n$$

で定義される集合である。

【0042】

ある確率変数を X^i , X^j とし、変数 X^j と X^i との相関関係が、変数 X^j は X^i によって影響を受けている、とすると、 X^i と X^j との間の条件付依存性は $X^i \rightarrow X^j$ と表される。このような関係が成り立っている時、 X^i は親ノードと呼ばれ、 X^j は子ノードと呼ばれている。一般に起こる事象というものは、一つの原因事象に対して一つの結果事象を得るという単一な因果関係を取っているわけではなく、様々な事象が複雑に絡み合った因果関係を取っている。従って、一つの子ノードに対して複数の親ノードが存在することになり、これを

10

と表す。ステップ S5 では、子ノードになりうる事象 1 つ 1 つに注目して、親ノードとなりうる事象を削除及び追加して検索している。

【0043】

上記のように、親ノードは

$$\Pi(X_j) = \{X_1, \dots, X_i\}$$

と表すが、これらは子ノードに相関関係があると決定した親ノードである。親ノードを決定するには、各事象を親ノードの候補として追加及び削除することにより、最適なモデルの構造を探索しなければならない。そのために、親ノードの候補を追加したり削除したりした際に、追加又は削除した状態のモデルを評価して構造学習を行い、ヒューリスティックにモデルを生成する。この構造学習には、バイズ情報量基準を近似した記述長基準 (Minimum description length: MDL) を用いる。ある確率変数のサンプルデータ X^N が与えられたときの、その確率変数のモデル g の記述長基準 $L(g | X^N)$ は、

20

$$L(g | X^N) = H(g | X^N) + \frac{k(g)}{2} \log N. \quad (1)$$

30

と定義づけられる。式 (1) の項目は、

$$H(g | X^N) = \sum_{j=1}^n H(j | x^N; g),$$

$$k(g) = \sum_{j=1}^n k(j; g),$$

$$H(j | x^N; g) = \sum_{s \in S(j; g)} \sum_{q \in V_j} -n[q, s, j; g] \log \frac{n[q, s, j; g]}{n[s, j; g]},$$

$$k(j; g) = (2pv_i - 1) \prod_{k \in \Phi(j)} v_k,$$

40

$$n[s, j; g] = \sum_{i=1}^N 1\{\pi_j^{(i)} = s\}$$

$$n[q, s, j; g] = \sum_{i=1}^N 1\{x_j^{(i)} = q \wedge \pi_j^{(i)} = s\}$$

と定義される量である。ここで、親ノード候補の集合 $\phi(j)$ は

$$\phi(j) = \{k : \forall X_k\}$$

50

と表され、これは $\{1, 2, \dots, j-1\}$ の部分集合である。

$$\pi_j^{(i)}$$

は、 i 番目のサンプル $X^{(i)}$ における親ノードの集合

$$\Pi(X_j) = \{x^1, \dots, x^i\}$$

の実現値の組となる。 $S(j; g)$ は、モデル g が与えられたときの変数 X^j に関する親ノードの集合 Π_j 実現値全体の集合を表す。 $1\{\cdot\}$ は、指示関数を表し、ある事象が真である場合に「1」を、偽の場合に「0」を返す関数である。 $n[s, j; g]$ は、モデル g の事象 s 及び j の周辺頻度関数を表し、 $n[q, s, j; g]$ はモデル g の事象 q 、 s 及び j の同時頻度関数を表す。周辺頻度関数と同時頻度関数との間には、

$$n[s, j; g] = \sum_{q \in V_j} n[q, s, j; g]$$

の関係が成り立つ。これをもとに条件付頻度関数 $n[q, j | s; g]$ を定義すると、

$$n[q, j | s; g] \triangleq \frac{n[q, s, j; g]}{n[s, j; g]}$$

と定義することが出来る。この条件付頻度関数は、

$$\sum_{q \in V_j} n[q, s, j; g] = 1$$

という性質を持つ。同様に、同時関数及び周辺頻度関数についても、

$$\sum_{s \in S(j; g)} \sum_{q \in V_j} n[q, s, j; g] = N$$

$$\sum_{s \in S(j; g)} n[s, j; g] = N$$

という性質を持つ。

【0044】

記述長基準 $L(g | X^N)$ の式 (1) は、経験エントロピーを表す項 (H で与えられる関数) とペナルティ項として働くベイジアンネットワークの自由度 (k で与えられる関数) とを評価することを示している。従って、この式は、注目する子ノードに親ノードの変数を削除又は追加して仮のモデルの構造を作成することにより、経験エントロピーを最小化しつつ、ベイジアンネットワークモデル全体の同時確率分布の自由度が小さいモデル構造を選ぶことを示している。さらに記述長基準 $L(g | X^N)$ を各変数に注目すると、

$$L(g | X^N) = \sum_{j=1}^n H(X_j | X^N; g) + \frac{k(X_j; g)}{2} \log N \quad (2)$$

と書き直すことが出来る。このことから、各変数に対して

$$H(X_j | X^N; g) + \frac{k(X_j; g)}{2} \log N$$

を最小化することとなる。

【0045】

ベイジアンネットワークモデル全体の同時確率分布は、多項分布となる。本発明では基準データ 103 に記憶されている基準データ、即ち、事前の仮定を用いつつ、統計データ

10

20

30

40

50

102に収集されている観測値からベイジアンネットワークのモデルを構築するが、この多項分布に対してベイズ情報量基準を用いるため、多項分布の共役事前分布であるディリクレ分布を事前確率分布として仮定する。このディリクレ分布は、多変量分布であり、親ノードと子ノードとの間の条件付確率を表にしたCPT (Conditional Probability Table: 条件付確率表) の事前確率分布全体を仮定するものとなる。ディリクレ分布は、 β 分布の混合分布であるため、CPTの1つのエントリ、即ち親ノードのある状態に注目するには、 β 分布を考えれば良い。CPTの各エントリの事前確率分布 $f(x)$ は以下の β 分布で与えられる。

$$f(x) = \frac{x^{p-1}(1-x)^{p-1}}{\int_0^1 x^{p-1}(1-p)^{p-1} dx} \quad (3)$$

10

【0046】

ここで、式(3)の p で与えられる項目は、事前確率分布に対するパラメータである。 p が0に十分近い時は注目している子ノードと親ノードとの間に依存関係がある事前確率分布を想定することになる。一方、 p に大きな値を入れることで、注目している子ノードと親ノードとの間に依存関係が無い事前確率分布を想定することになる。図4は、パラメータと事前確率分布との関係を表している。図4の横軸は事前確率の値を示し、縦軸はその値の分布を示す。0又は1に寄った分布を事前確率分布とするときは、注目している子ノードと親ノードとの間に因果関係を仮定したことになる(グラフA: $p=0.5$)。逆に、中央で凸となる分布を用いているときは、注目している子ノードと親ノードとの間に依存関係が無いことを仮定している(グラフB: $p=3$)。つまり、 p が小さいときは、0と1に偏ったU字型の分布となり、注目している子ノードと親ノードとの間に事前確率分布として依存関係を仮定したことになる。 $p=1$ のときは一様分布となり、何も仮定しないことを示す(グラフC)。 p の値が大きくなれば、形状が平均1/2の正規分布に近づき、注目している子ノードと親ノードとの間に依存関係が無いことを仮定したことになる。この p の値は、結果としてペナルティ項として作用するベイジアンネットワークのモデル全体の確率分布の自由度の影響を弱める方向に作用する。

20

【0047】

本発明では、設定部1041は、統計データ102に収集されているデータの蓄積状況に応じて図4に示したパラメータ p を調整して、注目している子ノードと親ノードとの間の依存関係を強くする。例えば、設定部1041は、統計データ102に収集されているデータの蓄積量が、統計データ102のデータを用いて依存性を設定するには少ない量であると判定すると、基準データ103に格納されている依存性を用いて設定する。また、設定部1041は、統計データ102に収集されているある事象のデータがある一定の期間より古いデータであると判定すると、基準データ103に格納されている依存性を用いて設定したり、統計データ102のデータから導出する依存性と基準データ103に格納されている依存性とを考慮して導出する依存性を設定したりする。また、統計データ102に収集されているある事象のデータの急激な変化や、ある一定の周期での繰り返し等のデータの推移の特徴を考慮して導出する依存性を設定する。もっと詳しく説明すると、初期状態では統計データ102に情報がまだ蓄積されていないため、ベイズ情報量基準をそのまま用いると、注目している子ノードと親ノードとの間の依存関係が無い結果を出力してしまうことが多い。そこで、図4に示した事前確率分布 p を調整して、注目している子ノードと親ノードとの間の依存関係を強くする。つまり、 p で表されるパラメータ項を0に近くする。このパラメータの変更は、注目している子ノードに関連がありそうな親ノードとの間に積極的に依存関係を強くさせることになる。これは、情報収集部101が観測する情報、即ち確率変数は予め分かるので、影響する確率変数同士には予め依存関係を強くさせるように設定する。また、これとは逆に、依存関係が弱くなるように設定することもある。

30

40

【0048】

50

本発明によると、動的に各事象間の依存性を設定することができるので、十分な量のデータが蓄積されていなくとも、ベイジアンネットワークのモデルを生成することができる。また、データの推移に応じて、動的に各事象間の依存性を設定することができるので、環境の変化による各事象間の依存性の変化に対応することが出来る。

【0049】

〈第2の実施の形態〉

次に、本発明の第2の実施の形態について説明する。尚、上記実施の形態と同様の構成については同一番号を付し、詳細な説明は省略する。

【0050】

本発明のモデル生成装置は、各種組込機器に用いることができる。例えば、
・ユーザのTPOや行動パターンに応じた携帯電話等のメニューの動的提示
・ユーザのTPOや行動パターンに応じた配信情報や広告の提供
・ユーザのTPOや行動パターンや利用パターンに応じたネットワークサービスの提示又は提供
・ユーザのTPOや行動パターンや利用パターンに応じた各種電気製品の動作やユーザアシスト動作
に使用することが出来る。

10

【0051】

図5は、本実施の形態の情報処理装置600のブロック図である。尚、以下の説明では、携帯電話を例にして説明するが、コンテンツ、各種情報、及びサービスの提供又は提示、若しくは自装置の各種動作を実行するというように、ある事象に対して実行すべき処理を決定することができる装置であればよい。

20

【0052】

情報収集部101は、新規測定値を収集すると、その測定値を推定部602に出力する。

【0053】

アプリケーション部601には、電話機能、音楽プレイヤー、メール、カメラ、ゲーム、インターネット等の各種機能及びコンテンツを実行するためのツールが記憶されている。

【0054】

推定部602は、情報収集部101が収集したデータと、ベイジアンモデル105に記憶されているベイジアンネットワークのモデルとに基づいて、アプリケーション部601の機能及びコンテンツ毎に適合性を検証する。そして、検証結果に基づいて、実行すべき動作を推定し、該当する機能又はコンテンツを実行する。検証方法は、本発明においては特定せず、一般的な技術を用いて実行すればよい。

30

【0055】

尚、上記説明では、情報処理装置内に、モデル生成装置を構成する場合を用いて説明したがこれに限らない。例えば、ユーザの自宅に設置されている複数の電機機器を管理するホームサーバにモデル生成装置及びを構成させ、ホームサーバにてユーザの動作やユーザの動作に影響を与える情報を測定又は各電器機器から収集してベイジアンネットワークのモデルを生成して記憶させる。そして、電気機器からの命令に応答してモデル又は一部のモデルを通知し、電気機器側に設けられた推定部で自装置にて実行すべき動作を推定して該当する動作を実行するようにしてもよい。

40

【0056】

本発明によると、動的に各事象間の依存性を設定して個人向けのベイジアンネットワークのモデルを作成し、このモデルに基づいて現在の状況から起こりうる事象を推定できるので、ユーザの生活パターンや環境に合致する機能、コンテンツ、及びサービスを提供することができる。

【0057】

尚、上述した本発明のモデル生成装置及び情報提供装置は、上述の説明からも明らかな

50

ように、ハードウェアで構成することも可能であるが、コンピュータプログラムにより実現することも可能である。この場合、プログラムメモリに格納されているプログラムで動作するプロセッサによって、上述した実施の形態と同様の機能、動作を実現させる。尚、上述した実施の形態の一部の機能をコンピュータプログラムにより実現することも可能である。

【0058】

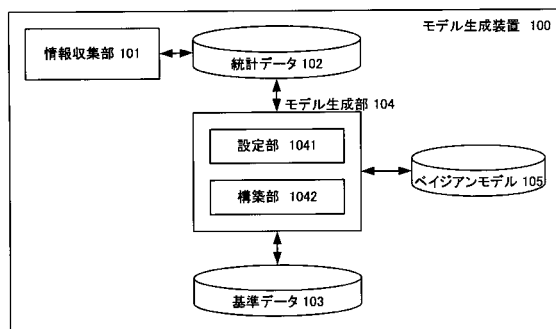
以上、実施の形態及び実施例をあげて本発明を説明したが、本発明は必ずしも上記実施の形態に限定されるものではなく、その技術的思想の範囲内において様々に変形し実施することが出来る。

【符号の説明】

【0059】

- 100 モデル生成装置
- 101 情報収集部
- 102 統計データ
- 103 基準データ
- 104 モデル生成部
- 1041 設定部
- 1042 構築部
- 105 ベイジアンモデル
- 600 情報処理装置
- 601 アプリケーション部
- 602 推定部

【図1】



【図2】

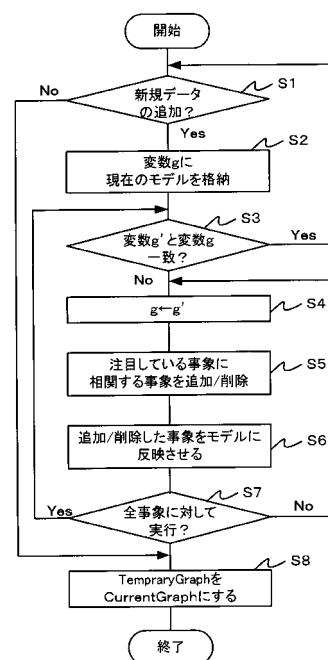
```

TemporaryGraph = g
While currentGraph ≠ TemporaryGraph do
  currentGraph ← TemporaryGraph
  i ← 1 { /* note that i ≠ 0. */ }
  for i < N do
    {delete, Xi, Π Xi} ← find_delete_dependency(Xi)
    {add, Xi, Π Xi} ← find_add_dependency(Xi)
    TemporaryGraph ← update_graph( {delete, Xi, Π Xi}, {add, Xi, Π Xi} )
  end for
end while

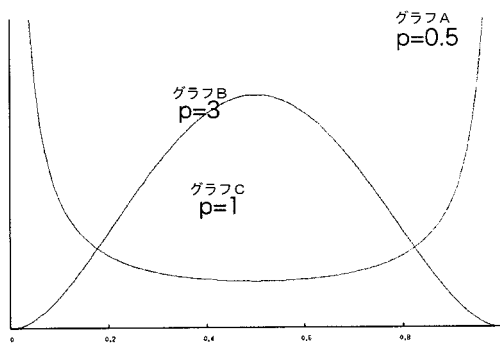
```

A1
A3
A2

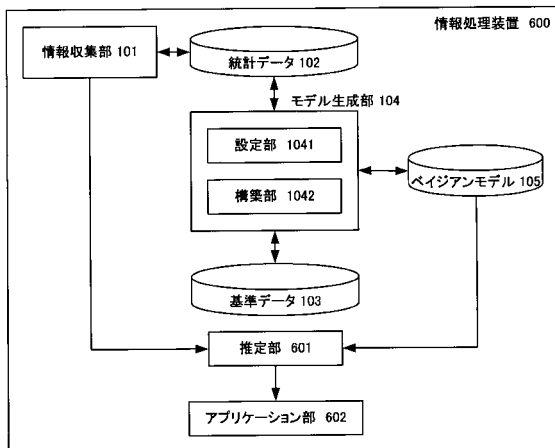
【図3】



【図 4】



【図 5】



フロントページの続き

(72)発明者 鈴木 啓高

東京都目黒区東山1丁目4番4号 株式会社エイチアイ内

審査官 相崎 裕恒

(56)参考文献 特開2006-127452 (JP, A)

特開2009-087318 (JP, A)

特開2009-093422 (JP, A)

特開2002-163434 (JP, A)

本村陽一, ベイズ推定における事前分布のグラフ構造モデリングと実生活行動理解, 情報処理学会論文誌, 日本, 社団法人情報処理学会 Information Processing Society of Japan, 2007年 6月15日, 第48巻, No. SIG9 (CVIM18), p.43-56

(58)調査した分野(Int.Cl., DB名)

G06F 15/18, 17/30

G06N 3/00 - 99/00