

ユビキタスシステムにおけるリコメンデーションの応用研究について

石野 正彦*

The Research of the Applications about the Recommendation in the Ubiquitous Systems

Masahiko Ishino

Abstract

This paper is a research of technique about the recommendation technology evaluates the palatableness of the person in the ubiquitous systems by data-mining. We study application of discovery of the knowledge that is new from many history data to the base in this recommendation technology and the development of the new product and the marketing support. We repeat recommendation to be able to get closer to user's needs for behavior of user each from the ubiquitous systems by the analysis of the acquisition data of the ubiquitous systems. Thereby the potential taste and the user's "notice" are provided.

Keyword: Ubiquitous system, Recommendation, Data-mining, Marketing, Information security

1. はじめに

インターネットではユーザのウェブサイトへの訪問履歴が把握でき、ネットショップにおいて顧客のアクセスログや購買履歴を把握することができる。また、ユビキタスシステムの電子タグなどで得られたデータの分析によって、システムのユーザの様々な振る舞いに対して、ユーザのニーズに近づけるための推奨をおこない、ユーザの潜在的な嗜好の推定や気づきを与える。

今まで、商品推奨のモデリングによる顧客向け商品推奨や情報セキュリティのリスク推定などについてリコメンデーションシステムの応用研究を進めてきた^{1),2),3)}。また、ライブメモの情報タグの分析による先行研究も参考とした⁴⁾。データの属性を抽出し、複数の属性間の重み付けを行列で表現し、推奨すべき要素を優先付けする評価法である、「属性関係行列」による数理的モデリング手法を適用した各種の応用事例について紹介する。

2. 商品推奨モデル

属性関係行列によるモデリング手法は商品属性や顧客属性に対し、コンジョイント分析を用いてウエイト付けをおこなう。商品属性をカテゴリ、ブランド、デザイン、カラー、スペック、使い易さ、丈夫さ、プライスなどを設定し顧客別の購買データから商品の属性値ベクトル x^m へ変換する。商品属性ベクトルからなる商品属性行列を表1のように該当する欄に"1"を挿入し、商品属性行列 X を生成する。顧客は、年代、性別、業種、職種、スキル、趣味、年収、地域などの属性を設定し、商品データを顧客属性値ベクトル y^n へ変換し、顧客属性行列 Y を生成する。また、購買データから顧客ごとの商品属性の嗜好性をコンジョイント分析で商品属性に対するウエイト付けをする。このウエイト付けで顧客の購入頻度が多い商品属性の推奨度が高いとみなす。

* 経営情報学科

$$x_j, j = 1 \sim m \in B^m \quad (1)$$

$$y_i, i = 1 \sim n \in B^n \quad (2)$$

表 1 に商品属性行列 X を示す。

表 1 商品属性行列 X

属性 ・ 商品	カ テ ゴ リ	ブ ラ ン ド	デ ザ イ ン	カ ラ ー	ス ペ ッ ク	使 い 易 さ	丈 夫 さ	プ ラ イ ス
g_1	1	0	0	1	0	0	1	0
g_2	0	1	0	0	0	1	0	1
g_3	0	0	1	0	1	0	1	0
\vdots								
g_k	0	1	0	0	0	1	0	1

3. 属性関係行列

商品属性行列 X と顧客属性行列 Y から属性関係行列 W を定義して表 2 に属性間のウエイト値を設定する。この属性間ウエイト値の大きさは、商品と顧客の属性の関係の嗜好性の強さを表す。

表 2 属性関係行列 W

	カ テ ゴ リ	ブ ラ ン ド	デ ザ イ ン	カ ラ ー	ス ペ ッ ク	使 い 易 さ	丈 夫 さ	プ ラ イ ス
商品								
顧客								
属性								
年代	4	3	5	2	1	2	1	2
性別	1	2	1	1	4	3	1	1
業種	2	1	2	1	1	2	1	3
職種	3	4	1	2	1	2	3	1
スキル	2	1	3	1	5	2	4	1
趣味	4	1	4	3	1	1	1	1
年収	1	4	2	2	3	1	2	1
地域	2	2	3	1	1	2	4	1

4. 商品の推奨

顧客 c_i から見た特定商品の推奨評価値 p_i は式(3)で求められる。

$$p_i = W^T y_i \quad (3)$$

任意の年代、職種、地域などの顧客 c_i の属性ベクトルを式(4)とする。

$$y_i = (1, 0, 0, 0, 1, 0, 1, 0)^T \quad (4)$$

顧客 c_i の商品推奨評価 p_i は、式(5)となる。

$$p_i = W^T y_i = (7, 8, 10, 5, 9, 5, 7, 4)^T \quad (5)$$

式(5)で、ブランド、デザイン、スペックなどの商品属性のウエイト値が高い商品から推奨する。本モデルは、購買データを属性へ変換し、商品を推奨する。商品属性と顧客属性からの推奨法であるため、商品属性による類似商品も推奨でき、販売支援システムの付加機能として活用できる。

5. 考察

顧客の嗜好に関する情報や、購買履歴などを参照して、それぞれの好みに合致すると思われる商品やサービスを推奨する。過去に同じ商品を購入したことがある他の顧客を似たような嗜好の持ち主とみなし、その興味対象を紹介するサービスやオペレータが顧客の希望に則して個別に情報収集し配信するなど、様々なサービスが提供できる。顧客にとっては自分の好む情報に効率よくアクセスでき、情報を提供する側にとっては、顧客の購入率を高めることができる。顧客満足度の向上と販売促進の両方を実現するサービスの一手法としてさらに研究を進める。

6. 情報セキュリティのリスク推定モデル

社会でのインターネットの普及により、さまざまな情報セキュリティ対策が必要となってきた。企業では情報セキュリティのインシデントが日常的に頻発し、その対策に注力している。インシデントの推定方法に IT 資産や部門属性の重要度を付加した数理モデルに改良し、インシデントの発生度を推定する方法を提案した^{5),6)}。情報セキュリティにおけるインシデントの推定方法について数理モデルの適用と運用方法について述べる。

6. 1 推定目的

種々のインシデントの属性と対象部門の属性との関係の大きさをウエイト付けした「属性関係行列」の数理モデルを定義して、或る特性を持つ部門において今後、想定されるインシデントの推定順位を設定する。この順位付けしたインシデントと ISMS 資産評価基準をもとに AHP によって評価を行なう⁷⁾。推定したインシデントを ISMS 担当者が絞り込むことができ、目標にあった的確な情報セキュリティ対策が組める。IT 資産の重要度や特性によって、インシデントの発生を推定する。発生した警告ログ情報を属性変換し、発生していない類似のインシデントを推定する。

この手法を適用することにより、今後、発生率が高いと推定されたインシデントに対して ISMS における情報セキュリティ対策が立てられる。情報セキュリティに関するインシデントの管理や検出に関する先行研究では、異常検知によるリスク回避策が多い^{8),9)}。表 3 に示す通り、本手法ではインシデント属性と対象部門属性に着目して脆弱性を推定した。ログデータ分析手順の中で分類属性へ変換し、優先度の高いインシデントを推定した。

表 3 先行研究との比較

	先 行 研 究	本 手 法
検 出	不正行動予知 異常検知	インシデントの推定
手 法	統計分析 例外値検出	属性行列 データマイニング
評 価	危険予知	脆弱性の推定
イ ン シ デ ン ト	データ検知	属性の推定
リ ス ク 管 理	リスク回避	リスク回避
対 策 優 先 順 位	現象把握	優先順位付け

6. 2 属性値ベクトル

属性行列のインシデント属性や部門属性に対してはコンジョイント分析を用いてウエイト付けを行い、それらのウエイトを用いて部門へのインシデントの推定をおこなう。インシデントを種別(不正侵入、情報漏洩、不正コピー、操作ミス)や時間帯の属性を設定し、部門ごとの警告ログからインシデントの属性値ベクトル x^m へ変換した。特定期間の警告ログの発生件数によって部門ごとのインシデント属性の傾向についてコンジョイント分析におこない、インシデント属性に対するウエイト付けをする。各部門は、業務、IT 資産の属性を設定し、インシデント別のログデータを部門の属性値ベクトル y^n へ変換する。警告ログの発生頻度が多い部門が今後、インシデント発生する可能性が高いとみなし、コンジョイント分析により、部門属性のウエイト付けをおこなう。

$$a_j, j = 1 \sim m \in B \quad (6)$$

$$d_i, i = 1 \sim n \in B \quad (7)$$

インシデント属性ベクトルから成るインシデント属性行列を表 4 に例示する。インシデント属性の不正アクセス、不正コピー、情報持ち出し、暗号化未処理や時間帯属性の定時間内、深夜、休日の該当する欄に“1”を挿入し、インシデント属性行列 A を生成する。部門属性ベクトルから成る部門属性行列も部門属性の総務、営業、製造、研究開発や IT 資産属性の個人情報、設計情報、販売情報、知的財産の該当する欄に“1”を挿入し、部門属性行列 D を生成する。

表4 インシデント属性行列 A

インシデント	不正アクセス	不正コピー	情報持ち出し	暗号化未処理	操作ミス	定時間内	深夜	休日
a_1	1	0	0	0	0	0	1	0
a_2	0	0	1	0	0	1	0	0
a_3	0	0	0	0	1	0	1	0
\vdots								
a_k	0	1	0	0	0	1	0	0

6. 3. 属性のウェイト付け

特定部門におけるインシデント属性の発生順位のウェイト付け（パートワース値）を表すインシデント属性ベクトルと特定インシデントにおける部門属性の発生順位のウェイト付けのそれぞれをコンジョイント分析によって求めた。対象部門 $\{d_i\}$ に対するインシデント属性のパートワース行列 U を式(8)で表す。

$$U \in R^{n \times m} \tag{8}$$

部門 d_i に対するインシデントのウェイト行列成分 U_i は、式(9)となる。

$$U_i = (u_{i1}, u_{i2}, \dots, u_{im}) \tag{9}$$

部門 d_i に対するインシデント a_j の推定評価値は、式(10)で表される。

$$E_{aj} = U^T y_i \tag{10}$$

ウェイト行列 U を用いて、インシデント a_j の発生評価値 E_{aj} を求め、インシデント集合の中から推定評価値の高いインシデントを推定インシデントの選定候補とする。特定インシデントに対する部門属性のウェイト行列 V を式(11)で表す。

$$V \in R^{m \times n} \tag{11}$$

インシデント a_j に対する部門のウェイト行列は、式(12)となる。

$$V_j = (v_{j1}, v_{j2}, \dots, v_{jn}) \tag{12}$$

インシデント a_j に対する部門 d_i の推定評価値は、式(13)で表される。

$$E_{di} = V^T x_j \tag{13}$$

次にインシデントと部門の2つの属性の関係をウエイト付けした属性関係行列 $W^{m \times n}$ を式(14)に定義する。各部門の警告ログデータの属性から特定部門において、これから発生し得るインシデント発生を推定する。インシデントと部門の属性関係行列 $W^{m \times n}$ を示す。

$$W = \begin{array}{c} \boxed{\text{インシデント属性}} \\ \left[\begin{array}{cccc} w_{11}, w_{12}, w_{13}, \dots, w_{1n} \\ w_{21}, w_{22}, w_{23}, \dots, w_{2n} \\ \vdots & & \vdots & \\ w_{m1}, w_{m2}, w_{m3}, \dots, w_{mn} \end{array} \right] \begin{array}{c} \boxed{\text{部}} \\ \boxed{\text{門}} \\ \boxed{\text{属}} \\ \boxed{\text{性}} \end{array} \end{array} \quad (14)$$

同様にインシデント別部門のウエイト行列は式(15)となる。

$$V_j = W \quad x_j \quad (15)$$

6. 4 インシデントのリスクの推定評価

表5に インシデント属性のウエイト行列と推定評価値を示す。

表5 インシデント属性のウエイトと推定評価値

インシデント	不正アクセス	不正コピー	情報持ち出し	暗号化未処理	操作ミス	定時間内	深夜	休日	推定評価値
a_{r1}	0	6	0	0	0	4	0	0	10
a_{r2}	0	0	5	0	0	0	3	0	8
a_{r3}	0	0	0	4	0	0	0	3	7
\vdots									
a_{rk}	2	0	0	0	0	1	0	0	3

表6 属性関係行列 W

属性	不正アクセス	不正コピー	情報持ち出し	暗号化未処理	操作ミス	定時間内	深夜	休日
総務	4	3	5	2	1	2	1	1
営業	1	2	1	1	4	3	1	1
製造	2	1	2	1	1	2	1	3
研究開発	3	4	1	2	1	2	3	1
個人情報	2	1	3	1	5	2	4	1
設計情報	4	1	4	3	1	1	1	1
販売情報	1	4	2	2	3	1	2	1
知的財産	2	3	3	1	1	3	4	1

属性関係行列の属性間ウエイト値を求めた。ウエイト値が大きいほど関係性が強い。部門 d_i から見たターゲットインシデントの評価推定値 p_i は式(16)で求められる。

$$p_i = W^T y_i \quad (16)$$

総務部門、個人情報の部門 d_i の属性ベクトルを式(17)とした場合、

$$y_i = (1, 0, 0, 0, 1, 0, 0, 0)^T \quad (17)$$

部門 d_i におけるインシデントの推定評価値ベクトル p_i は、式(18)となる。

$$p_i = W^T y_i = (6, 4, 8, 3, 6, 4, 5, 2)^T \quad (18)$$

式(18)から、情報持ち出し、深夜のインシデント属性の発生推定ウエイト値が高い。インシデント a_j に対する部門の評価推定値ベクトル q_j は、式(19)で求められる。

$$q_j = W^T x_j \quad (19)$$

不正コピー、定時間内のインシデント a_j の属性ベクトルを式(20)とした場合、

$$x_j = (0, 1, 0, 0, 0, 1, 0, 0)^T \quad (20)$$

であるので、インシデント a_j に対する部門の推定評価値ベクトル q_j は、式(21)となる。

$$q_j = W x_j = (5, 5, 3, 6, 3, 2, 5, 6)^T \quad (21)$$

式(21)から、研究開発、知的財産の部門属性の発生推定ウエイト値が高い。このようなウエイト値の高い部門属性から推定インシデントの選定候補とする。

7. 情報セキュリティのリスク推定モデルの考察

情報セキュリティに関するインシデントの推定方法について考察をおこなった。

- (1) IT 資産の重要度や特性によってインシデントの発生を推定した。発生した警告ログ情報を属性変換し、発生していない類似のインシデントを推定した。さらに、インシデント属性と部門属性から推定であるため、ログデータの事後分析と比べてインシデント属性による他の類似インシデントが推定できた。
- (2)属性関係行列を用いて、或る属性を持つ部門に対するインシデントの推定と或る属性のインシデントに対する部門を推定する一つの方法を提案した。
- (3)実際の情報セキュリティ対策では、一定期間のログデータを適用したインシデントの推定後に、他の期間のログデータを繰り返して適用し、インシデント推定を実施するような PDCA サイクル型の情報セキュリティ対策が必要である。

「属性関係行列」の数理モデルを使い、特定部門において、インシデントの推定順位を設定し、順位付けしたインシデントと ISMS 資産評価基準をもとに優先順位を決める。ISMS の管理者が推定したインシデントの候補を絞り込み管理基準にあった情報セキュリティ対策が組める。

参考文献

- 1) 石野正彦,五月女健治,工藤司: "リコメンデーションシステムについてのモデリング手法と応用研究",電子情報通信学会技術研究報告,ソフトウェアインタプライズモデリング SWIM ,pp.37-pp.42,2009.9.
- 2) 石野正彦, 八巻直一, 市川照久, 水野忠則: "顧客の嗜好変化に対する学習機能を持った商品推奨法", 経営情報学会 2006 年度 春季全国研究発表大会予稿集, pp.554-557,2006.6.
- 3) 石野正彦, 八巻直一, 市川照久, 水野忠則: "セールスプロモーション現場での数理的手法による商品推奨方法の提案", 経営情報学会 2005 年度秋季全国研究発表大会予稿集,pp.440-443,2005.11.
- 4) 矢野絵美,飯島貴広,篠原勲,加藤俊一: "ユビキタスな情報タグのパーソナルライブメモ化と行動支援への応用",情報処理学会シンポジウム論文集,2003.
- 5) 石野正彦, 八巻直一, 工藤司, 五月女健治: "データマイニングによる情報セキュリティ・インシデントの推定方法に関する研究", 経営情報学会 2008 年度秋季全国研究発表大会要旨集,pp.5-8,2008.11.
- 6) 石野 正彦, 八巻 直一, 井上 春樹: "情報セキュリティにおける事故の推定方法に関する研究",2007 年度 経営情報学会 春季全国研究発表大会要旨集, pp.326-329, 2007.11.
- 7) 荒井良太,八巻直一:"ISMS における AHP を用いたリスクアセスメントと最適対策案決定",経営情報学会 2005 年度秋季全国研究発表大会予稿集,pp.26-29,2005.11.
- 8) 山西健司:"データマイニングの情報セキュリティへの応用",人工知能学会誌,2006.9.
- 9) 石黒正揮,鈴木裕信,村瀬一郎,篠田陽一: "インターネット上の脅威分析を支援する空間および時間的な特徴量に基づく分析手法", 情報処理学会論文誌, Vol.48, No.9, pp.3148-3162,2007.9.

(平成 22 年 3 月 31 日受理)