

D-026

ラフ集合を用いた対象者との関係に基づく開示情報選定

Rough Sets to Control Privacy Release Based on Relationships with Receivers

小林 勇人†
Hayato Kobayashi

島川 博光†
Hiromitsu Shimakawa

1. はじめに

2005年4月に個人情報保護法が施行されたようにプライバシー保護に対する関心が高まっている。個人情報として自身の情報はもちろんであるが、購買履歴や行動履歴などの履歴情報、位置情報やスケジュール情報のように行動を推測される情報も存在する。個人情報は極力自ら開示すべきではないが、情報開示せざるを得ない状況も存在する。また個人の情報はその種類によって開示する内容、もしくは対象が異なり、誰にでも公開すべきではない。このような状況において、プライバシー保護に配慮しながら情報を開示することが必要となる。

現状では状況に応じた情報開示を的確にかつ自動的に行うことができない。状況に応じた開示を行うさいには、部分的な情報開示、または情報開示しないことも必要となる。だが、状況に適した情報の詳細さを手動で設定することは手間がかかり、大きな労力も必要となる。

既存研究として、プライバシーを保護するための研究[1]が行われており、空間の概念を用いて情報開示を設定する研究[2]もなされている。この研究においては情報の開示に対して、空間の状況によって開示する段階を決定している。一般的に人間は情報提示者と開示対象者の間にある関係により、情報開示するさいの詳細さを決定する。本研究においては、この関係を用いて対象者を集合化し、各集合にあわせた詳細度を指定する。

2. 対象者との関係に基づく開示

2.1 情報変換

本研究においては情報開示のなかからスケジュール開示に焦点を当てる。スケジュール開示に焦点を当てた理由として、情報提示者の行動やプライバシーに関する内容を他人に推測されることが挙げられる。さらにスケジュール情報は他者に開示する機会も多い。このようにプライバシーを保護しながら情報開示せざるを得ない状況においては、開示情報の詳細さを変更する必要が生じる。本手法を用いることで、自動的に開示対象者ごとに適した開示情報を提示する。

個々の対象者に対応した開示情報をそれぞれ設定することは大きな労力となる。そこで情報提示者と開示対象者との関係を設定して開示対象者の特徴を掴むことにより開示情報の条件をルール化する。開示情報のルールを求めるためにラフ集合理論[3]を用いる。ラフ集合理論を用いることで、開示情報の詳細さを決定するために必要となる特徴的な条件を抽出する。

2.2 開示の条件

情報を変換するときと同様な状況でどのように情報を開示するかが分かれば、他の状況における変換に利用す

コンテンツ	ルール										開示意志
	家族	大学	研究室	サークル	アルバイト	上司(教員)	先輩	友人	後輩		
皮膚科通院	1	0	0	0	0	0	0	0	0	0	皮膚科通院
皮膚科通院	0	1	0	0	0	0	0	0	1	0	通院
皮膚科通院	0	0	0	0	0	1	1	0	0	0	用事
研究室定例ゼミ	0	1	1	0	0	0	1	0	0	0	研究室定例ゼミ
研究室定例ゼミ	1	0	0	0	0	0	1	0	0	0	大学
研究室定例ゼミ	0	0	0	0	1	1	0	0	0	用事	
アルバイト	1	0	0	0	0	0	0	0	0	0	アルバイト
アルバイト	0	1	0	1	0	0	0	0	1	0	私用
アルバイト	0	0	0	0	0	0	0	0	0	0	用事

図1: 開示の決定表

ることができる。これは同様な状況では同様な情報変換法が適用できるためである。変換するために必要となる情報を示した表を決定表と呼び、これを図1に示す。決定表により、開示情報の変換を行う。

開示するためのルールを適用することで、多数の対象者に対して情報開示することが可能になる。開示情報を対象者に適した形に変換するためのルールを開示ルールと呼ぶ。開示ルールは、提示情報について対象者に開示するための条件を示している。

決定表は情報提示者と開示対象者との関係を表すルール、変換前の情報であるコンテンツ、コンテンツの情報をいかなる開示内容に変換するかを示す開示意志から構成される。開示意志は対象者に示される情報であり、対象者に適した情報の詳細さを示している。ルールが開示意志を決定する条件部となる。

ルールは複数の属性により構成される。属性は、家族、友人のように、ある対象者を特定するための要素である。ルールの属性要素は情報提示者により必要な項目が異なるため、情報提示者それぞれが設定する。個々の属性について満たす場合は1を、満たさない場合は0を記述する。すべての属性について記述することにより、対象者が情報提示者と、どのような関係であるかを示す。このルールの属性要素を用いて特徴を判別し、開示情報を決定する。

2.3 スケジュールの詳細度階層

決定表を用いて開示意志を決定する。この情報を開示ルールと呼ぶ。開示ルールを生成するさいに、同一コン

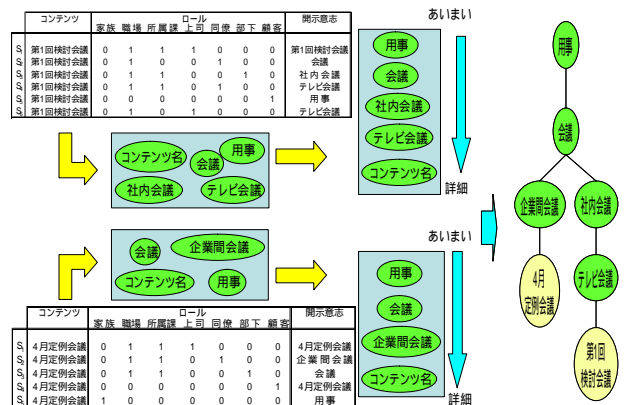


図2: スケジュール木の生成手順

†立命館大学大学院 理工学研究科

コンテンツ内の開示意志における詳細さの順序を決定する必要がある。決定した詳細さを階層的に表し、すべてのコンテンツで結合した木をスケジュール木と呼ぶ。決定表から開示意志の詳細さを決定して階層構造で表す流れを図2に示す。スケジュール木の根は用事というノードで固定され、葉がコンテンツ名である。根と葉を結ぶ枝上のノードは開示意志階層中のノードを利用する。結合するさいに、ノードが同一名の場合は、ひとつにまとめる。

2.4 開示ルール生成

決定表から開示ルールを自動生成する。同一コンテンツごとにまとめた決定表について、開示意志の特徴をルール要素から抽出するために縮約を行う。縮約のさいには識別行列を用いる。識別行列においては、ルールの各行をベクトルとみなし、各ベクトルを比較してその値が異なっている属性を抽出する。図2中の4月定例会議のコンテンツについて、縮約後の開示条件を表1に示す。

縮約後の決定表を用いて開示ルールを生成する。必要最小な条件で説明できる開示ルールを生成するために決定行列を作成する。決定行列においては開示ルールを求める開示意志と、それ以外の開示意志におけるルール属性の相違を抽出する。これを表2に示す。表2より S_1 と S_2, S_3, S_5 を区別するには、

$$(上司1 \text{ 同僚}0) \quad (上司1 \text{ 部下}0) \quad (上司1 \text{ 家族}0)$$

の条件が必要となる。同様に S_4 と S_2, S_3, S_5 を区別する。 S_1 もしくは S_4 からの属性値を用いるため、求められた互いの属性値を or 結合する。論理演算式で計算すると

$$\begin{aligned} & (上司1 + 同僚0) \times (上司1 + 部下0) \times (上司1 + 家族0) \\ & + (同僚0 + 顧客1) \times (部下0 + 顧客1) \times (家族0 + 顧客1) \\ & = 上司1 + 顧客1 \end{aligned}$$

になる。この式が4月定例会議の開示ルールである。式の意味は、「上司の属性が1または顧客の属性が1」となる。同様にして他の開示意志についても縮約を行い、決定行列を作成して開示ルールを生成する。このようにして用事を除くそれぞれの開示意志における開示ルールが自動生成される。

2.5 開示ルールに従った開示情報設定

それぞれの開示意志についてのルールが生成されたうえで、どの開示意志が選択されるかを決定する。対応する開示意志のルールに適合した場合にコンテンツが変換され、照合には前述の開示意志階層を利用する。スケジュール木の開示意志階層において開示ルールは詳細な開示意志から順に開示対象者のルールと比較する。適合した場合は、選択された開示意志が開示対象者に提示される。適合しなかった場合は、次に詳細な開示意志のルールと比較する。どの開示ルールにも適合しない場合には

表1：縮約後の決定表

	コンテンツ	ルール					開示意志
		家族	上司	同僚	部下	顧客	
S_1	4月定例会議	0	1	0	0	0	4月定例会議
S_2	4月定例会議	0	0	1	0	0	企業間会議
S_3	4月定例会議	0	0	0	1	0	会議
S_4	4月定例会議	0	0	0	0	1	4月定例会議
S_5	4月定例会議	1	0	0	0	0	用事

表2：決定行列

	S_2	S_3	S_5
S_1	上司1, 同僚0	上司1, 部下0	上司1, 家族0
S_4	同僚0, 顧客1	部下0, 顧客1	家族0, 顧客1

表3：評価結果

コンテンツ系列名	ケース数	適合数	適合率(%)
研究室ゼミ	49	46	93%
休暇	49	43	87%
通院	49	43	87%
サークル	49	42	85%
アルバイト	49	42	85%

もっともあいまいな提示情報である用事が選択される。

それぞれの開示意志における開示ルールが決定された後も、対象者に提示する開示意志を訂正することで開示ルールは適宜、変更される。訂正を繰り返すなかで開示ルールは情報提示者の望むルールに近づいていく。

3. 評価

生成された開示ルールが情報提示者の考えに適合しているかを評価した。あらかじめコンテンツ、開示意志階層、ルールを指定し、ルール属性により15人の開示対象者を設定した。被験者は15人の開示対象者に対してどのように情報を提示するか決定する。提示する開示意志については詳細さが異なる開示意志3段階から選択する。収集した15人分のデータの中から8人分を選択して開示ルールを自動生成し、残りの7人分のデータをルールの精度を比較する対象とした。

結果を表3にまとめる。結果としてコンテンツにより適合率に差が見られた。差が見られた要因としてコンテンツにより必要となるルールの詳細さが異なる点が考えられる。全体の適合率は88%であった。

4. おわりに

本論文では、情報提示者がそれぞれの開示対象者に適した情報提示をすることで、プライバシーを保護しながら情報開示を行うことを可能とする手法の提案をした。評価実験として開示ルールの精度を検証した。開示ルールによる開示情報の精度は88%であった。今後、多くの実験データを収集してデータの量と精度の関係を調べていく。

参考文献

- [1] Nor Zehan Binti Ahmad, 楯岡孝道, 阿部公輝, 柔軟なプライバシー保護を考慮した分散型位置情報システムの提案, 情報処理学会研究報告 2004-MBL-30, pp.25-32, Sep. 2004.
- [2] Vassilis Kostakos, Eamonn O'Neill, Alan Penn, Designing Urban Pervasive Systems, IEEE Computer, Vol. 49, No. 9, 2006.
- [3] 森典彦, 田中英雄, 井上勝雄 編, ラフ集合と感性—データからの知識獲得と推論, 海文堂, 2004.