

テクニカルノート

負の相関ルール集合の極小生成子に基づく圧縮表現

岩沼 宏治^{1,a)} 佐生 隼一² 黒岩 健歩² 山本 泰生¹

受付日 2016年1月27日, 採録日 2016年4月5日

概要: 本論文では負の相関ルール集合の圧縮形式について考察する. まず初めに, 頻出アイテム集合の圧縮表現としてよく知られる飽和集合が, 負の相関ルール集合の圧縮技術としては不適切であることを示す. 次に, その解決策として極小生成子を用いた負ルール集合の圧縮表現を新しく提案する. 極小生成子は飽和アイテム集合と対をなす概念である. 提案した圧縮表現の無損失性などを理論的に証明する. また実証実験を行った結果, 稠密なデータセット上の負ルール集合の圧縮に有効であることが確認できたので報告する.

キーワード: 相関ルールマイニング, 負ルール, 極小生成子, 飽和アイテム集合, 圧縮

A New Compression Form for a Set of Negative Association Rules Based on Minimal Generators

KOJI IWANUMA^{1,a)} SHUNICHI SASHO² YASUHO KUROIWA² YOSHITAKA YAMAMOTO¹

Received: January 27, 2016, Accepted: April 5, 2016

Abstract: In this paper, we study a lossless compression for a set of negative association rules. First, we show a representation difficulty which occurs in applying the closed itemset technique for such a compression problem. Next we propose a new compression method for a set of negative rules, which is based on minimal generators. A minimal generator for an itemset is a dual concept of a closed itemset, and can solve the above representation problem. We prove that the proposed method based on minimal generators is a lossless compression, and also show, through experiments, that the method is effective for compressing a set of negative rules extracted from dense data sets.

Keywords: association rule mining, negative rule, minimal generator, lossless compression, closed itemset

1. はじめに

相関ルールとは, トランザクションデータベース中で頻繁に共起するアイテム集合の関係を記述したものである. X と Y をアイテム集合とすると, データベース中で X が出現するトランザクションの多くに Y も出現することを $X \Rightarrow Y$ と表し, 正の相関ルールと呼ぶ. これに対して本論文では, X と Y がほとんど同時には生起しない現象を表現する $X \Rightarrow \neg Y$ や $\neg X \Rightarrow Y$ なる形の負の相関ル

ルを考察する. 負の相関ルールは正のルールでは表現が困難な共起関係を簡潔に記述でき, データから有益な情報を抽出することが可能になる. ただし, 負のルールを抽出するためには, 非頻出なアイテム集合を扱う必要がある. 正の相関ルールの場合と比べて探索空間が格段に大きく, また抽出されるルールの数も非常に多くなる. これまでに抽出計算の効率化に関して多くの研究 [1], [2], [3], [5], [6] がなされてきている. 本研究では, 抽出した負の相関ルール集合の圧縮表現について考察し, 極小生成子 (minimal generators) [7] に基づく形式を新しく提案し, 抽出ルールの数の削減を図る.

頻出アイテム集合および正ルールの集合の圧縮には飽和集合がよく用いられるが, 負ルール集合の圧縮に用いた場合, 本来抽出すべき負ルールが表現できなくなる現象が生

¹ 山梨大学大学院総合研究部
University of Yamanashi, Kofu, Yamanashi 400-8511, Japan
² 山梨大学大学院医学工学総合教育部コンピュータ・メディア工学専攻
University of Yamanashi, Kofu, Yamanashi 400-8511, Japan
^{a)} iwanuma@yamanashi.ac.jp

じる。本論文では、飽和アイテム集合と対の概念である極小生成子を用いて、負ルール集合の圧縮表現を考案する。極小生成子を用いることにより、負ルール集合の完全な圧縮、すなわち無損失圧縮が可能となる。予備的な実証実験の結果、稠密なデータセット上の負ルール集合において有効な圧縮効果も確認できたので、とり急ぎ報告する。

以下、2章は準備である。3章では、まず負ルール集合の圧縮に関する飽和集合の限界を示す。次に極小生成子を用いた新しい圧縮表現を提案し、その無損失性などを証明する。4章では圧縮率の実証実験を行う。5章はまとめである。

2. 準備

$\mathcal{I} = \{a_1, a_2, \dots, a_n\}$ をアイテムの全体集合とし、トランザクション T をアイテム集合 $T \subseteq \mathcal{I}$ と定める。トランザクションデータベース（以下、データベースと略） $\mathcal{D}$ をトランザクションの多重集合とする。 X をアイテム集合とすると、多重集合 $\mathcal{D}(X)$ を $\mathcal{D}(X) = \{T \in \mathcal{D} | X \subseteq T\}$ と定め、 $\mathcal{D}$ の各要素を X の出現と呼ぶ。以下では多重集合 A の大きさを $|A|$ と表記し、 $|\mathcal{D}(X)|$ を X の $\mathcal{D}$ 中の出現頻度と呼ぶ。 X の $\mathcal{D}$ 中の支持度 $\sup(X)$ を $\sup(X) = \frac{|\mathcal{D}(X)|}{|\mathcal{D}|}$ と定義する。最小支持度 ms をユーザが与える閾値とすると、 $\sup(X) \geq ms$ を満たす X を頻出アイテム集合と呼ぶ。

X と Y を互いに素なアイテム集合とすると、 $X \Rightarrow Y$ なる表現を正の相関ルール（以下、正ルールと略）と呼ぶ。また負の相関ルール（以下、負ルールと略）とは

$$\neg X \Rightarrow Y \text{ (左否定形)}, \quad Y \Rightarrow \neg X \text{ (右否定形)}$$

なる形の表現^{*1}である。相関ルールの $\Rightarrow$ の左右のアイテム集合をそれぞれ前件、後件と呼ぶ。上記の $\neg X$ は“ X が出現しない事象”を表現しており、負のアイテム集合と呼ばれる。以下、 C_X は X または $\neg X$ を表すものとする。また $C_X \Rightarrow C_Y$ で、左否定形 $\neg X \Rightarrow Y$ と右否定形 $X \Rightarrow \neg Y$ のいずれかを表すものとする。負アイテム集合および正負のルールの支持度 $\sup$ と確信度 conf を以下のように定める [1], [5], [6]。

$$\sup(\neg X) = 1 - \sup(X)$$

$$\sup(X \Rightarrow Y) = \sup(X \cup Y)$$

$$\sup(X \Rightarrow \neg Y) = \sup(X) - \sup(X \cup Y)$$

$$\sup(\neg X \Rightarrow Y) = \sup(Y) - \sup(X \cup Y)$$

$$\text{conf}(C_X \Rightarrow C_Y) = \frac{\sup(C_X \Rightarrow C_Y)}{\sup(C_X)}$$

上記の尺度のほかに、文献 [3], [6] では負ルールを評価する統計的な評価尺度（以下、関連度と略） rel を導入してい

る。通常に関連度は支持度だけを用いて定義されており、本論文でもそのようなものだけを考える。最小支持度、最小確信度、最小関連度とは、それぞれユーザが与える支持度、確信度、関連度に関する閾値である。このとき妥当な負ルールは以下のように定義される。

定義 1 (文献 [3]) 最小支持度 ms , 最小確信度 mc , 最小関連度 mr に対して、妥当 (valid) な負ルール $C_X \Rightarrow C_Y$ とは以下の条件を満たすものである。

(1) $X \cap Y = \emptyset$ (独立性条件)

(2) $\sup(X) \geq ms$ かつ $\sup(Y) \geq ms$

(3) $\sup(X \cup Y) < ms$

(4) $\sup(C_X \Rightarrow C_Y) \geq ms$

(5) $\text{conf}(C_X \Rightarrow C_Y) \geq mc$

(6) $\text{rel}(C_X \Rightarrow C_Y) \geq mr$

先行研究 [1], [6] では、妥当な負ルールを抽出するために、Apriori 流のボトムアップ型アルゴリズムを提案している。定義 1 の条件 (3) より、妥当な負ルール $C_X \Rightarrow C_Y$ の台集合 $X \cup Y$ は必ず非頻出となるため、Apriori 流の手法 [1], [6] は非常に効率が悪い。そのため文献 [5] では、頻出集合 X と Y を組み合わせた負ルールをボトムアップ的に生成と検査をする手法が提案されている。さらに文献 [2] では、 X と Y の組合せ探索空間をさらに効果的に削減するトップダウン型アルゴリズムが提案されている。

3. 負ルール集合の極小生成子に基づく圧縮表現

妥当な負ルールは正ルールに比べて本質的に大量に存在 [1] する。よって抽出計算の高速化とは別の問題として、抽出される負ルールの数の抑制問題がある。正のルール $X \Rightarrow Y$ では、台集合 $X \cup Y$ を飽和集合に限定することによって、抽出ルールの数を抑える手法がよく知られている。しかし、負ルールの抽出では前章で述べたように台集合を直接的に生成することは望ましくない。組合せ検査を行う X と Y をどのようにして限定するかが問題となる。本章ではこの抑制法として、極小生成子に基づく負ルール集合の圧縮表現法を提案する。

定義 2 (文献 [7]) アイテム集合 X が飽和しているとは、以下の条件を満たすアイテム集合 X' が存在しない場合をいう。

$$X \subseteq X', \quad X \neq X' \text{ かつ } \sup(X) = \sup(X')$$

アイテム集合 Y の閉包 $\text{clo}(Y)$ を、 $Y \subseteq X$ かつ $\sup(Y) = \sup(X)$ を満たす飽和アイテム集合 X と定める。 Y の生成子とは以下の条件を満たす Z をいう。

$$Z \subseteq Y \text{ かつ } \sup(Z) = \sup(Y)$$

Y の生成子 Z が極小とは、 $Z' \subseteq Z$ かつ $Z' \neq Z$ となる Y の生成子 Z' が存在しない場合をいう。

^{*1} 本論文では両否定形の負ルール $\neg X \Rightarrow \neg Y$ は取り扱わない。両否定形の負ルールに関する議論は文献 [2] を参照していただきたい。

極小生成子は閉包の対になる表現である。アイテム集合 X の閉包 $\text{clo}(X)$ は一意に定まるが、極小生成子は一般には複数存在する [7]。以下では X の極小生成子の集合を $\mathcal{MG}(X)$ と表記する。

X' を X の閉包または極小生成子とすれば、 X と X' はデータベース中で必ず同じトランザクションに出現する。このとき、 X が出現する相関ルール $R[X]$ に対して、 X を X' に置き換えた相関ルール $R[X/X']$ を考えれば、 $R[X]$ と $R[X/X']$ の支持度、確信度および関連尺度はまったく同じになる。よって、 $R[X]$ と $R[X/X']$ の両方を抽出することは冗長であり、 $R[X/X']$ で代表させることが適切と考えられる。しかしながら、 X' として閉包（飽和アイテム集合）を考えた場合、本来抽出すべき $R[X]$ が復元できなくなる場合がある。

例 1 データベース $\mathcal{D}$ として表 1 を考え、最小支持度 $ms = 0.3$ 、最小確信度 $mc = 0.5$ と仮定する。本例では議論の簡単化のために関連度は省略する。アイテム集合を要素列で表記し、出現頻度を適宜付記する。すなわち、出現頻度 3 のアイテム集合 $\{a, b\}$ を $ab : 3$ のように表記する。このとき $\mathcal{D}$ の頻出アイテム集合は以下のとおりである。

$$a : 3, b : 5, c : 3, ab : 3, bc : 3$$

この中で飽和集合は b, ab, bc である。このとき、負ルール R_0 として $ab \Rightarrow \neg c$ を考えてみる。 $\sup(R_0) = \frac{1}{3} \geq ms$ 、 $\text{conf}(R_0) = \frac{2}{3} \geq mc$ であり、後件の否定をはずした正ルールも $\sup(ab \Rightarrow c) = \frac{1}{6} < ms$ なので、 R_0 は妥当な負ルールである。一方で頻出アイテム集合 c の閉包は bc であることから、 R_0 の c を飽和集合 bc で置き換えた表現 $R_1 (= R_0[c/bc])$ は、 $ab \Rightarrow \neg bc$ となる。しかし、 R_1 は定義 1 の (1) に示した前件と後件の独立性条件に違反するので妥当なルールとして生成することはできない。 R_0 を表現する飽和集合を用いた負ルールは R_1 以外には存在しない。そのため、前件と後件に飽和集合を持つ妥当な負ルールだけを代表ルールとして抽出すると、 R_0 は復元できない。 R_0 は妥当なルールであるので、これは飽和集合に基づいて負ルール集合を圧縮した場合の不完全性（情報損失）を示している。これに対して、 $\mathcal{D}$ 上の頻出アイテム集合に対する極小生成子は a, b, c である。 R_0 の ab を極小生成子 a で置き換えた表現 $R_2 (= R_0[ab/a])$ は、 $a \Rightarrow \neg c$ となる。

表 1 トランザクションデータベース $\mathcal{D}$
Table 1 A transaction database $\mathcal{D}$.

TID	アイテム集合
1	abc
2	ab
3	ab
4	bc
5	bc
6	d

る。 R_2 は独立性の条件を満足するので妥当な負ルールとして生成できる。

以下に示すように、任意の妥当な負ルールは極小生成子を用いた負ルールで表現できる。以下では $C_{[X/X']} \Rightarrow C_{[Y/Y']}$ で、負ルール $C_X \Rightarrow C_Y$ の前件部分の X と後件部分の Y を、それぞれ X' と Y' で置き換えたものを表すものとする。

補題 1 X と Y をアイテム集合とし、 X' と Y' を $X' \in \mathcal{MG}(X)$ と $Y' \in \mathcal{MG}(Y)$ なる任意の極小生成子とする。このとき負ルール $C_X \Rightarrow C_Y$ が妥当であるならば、負ルール $C_{[X/X']} \Rightarrow C_{[Y/Y']}$ も必ず妥当な負ルールとなる。

証明 X と X' はデータベース中の出現がまったく同一、すなわち $\mathcal{D}(X) = \mathcal{D}(X')$ である。 Y と Y' も同様である。よって 2 つの負ルール $C_X \Rightarrow C_Y$ と $C_{[X/X']} \Rightarrow C_{[Y/Y']}$ の支持度と確信度、および関連度はまったく同一の値となる。また正ルール $X \Rightarrow Y$ と $X' \Rightarrow Y'$ の支持度もまったく同一である。よって定理の仮定より、負ルール $C_{[X/X']} \Rightarrow C_{[Y/Y']}$ は定義 1 の (2) から (5) の条件を満たす。さらに $X' \subseteq X$ かつ $Y' \subseteq Y$ より、(1) の独立性条件も明らかに満たす。以上から本定理は証明された。□

極小生成子を用いた負ルール $C_X \Rightarrow C_Y$ から、一般のルールを復元する場合には、 X と Y の拡大作業の適否を考える必要がある。たとえば、例 1 の負ルール $a \Rightarrow \neg c$ は、前件と後件を拡大した $ab \Rightarrow \neg c$ や $a \Rightarrow \neg bc$ は代表するが、 $ad \Rightarrow \neg c$ や $a \Rightarrow \neg bd$ などは代表していない。前件と後件の拡大は極小生成子の閉包の範囲に限定する必要がある。

補題 2 MX と MY を極小生成子とする。このとき、 $MX \subseteq X \subseteq \text{clo}(MX)$ と $MY \subseteq Y \subseteq \text{clo}(MY)$ なる任意の X と Y に対して、2 つの負ルール $C_{MX} \Rightarrow C_{MY}$ と $C_{[MX/X]} \Rightarrow C_{[MY/Y]}$ の支持度、確信度、関連度の値はそれぞれ等しい。

証明 閉包、極小生成子および支持度、確信度、関連度の定義より明らかである。□

例 1 で示したように、 MX と MY が互いに素であっても、 $MX \subseteq X$ と $MY \subseteq Y$ なる X と Y が互いに素とは限らないので、 $C_{MX} \Rightarrow C_{MY}$ が妥当であっても、 $C_{[MX/X]} \Rightarrow C_{[MY/Y]}$ が妥当とは限らないことに注意していただきたい。

データベース $\mathcal{D}$ 上の妥当な負ルールの集合を $\mathcal{R}$ とするとき、 $\mathcal{R}$ の圧縮表現としての極小生成子からなる負ルール集合を求める関数 $\text{MI}(\mathcal{R})$ を以下のように定める。

$$\text{MI}(\mathcal{R}) = \{C_{[X/MX]} \Rightarrow C_{[Y/MY]} \mid \exists (C_X \Rightarrow C_Y) \in \mathcal{R} : MX \in \mathcal{MG}(X), MY \in \mathcal{MG}(Y)\}$$

補題 1 より、 $\text{MI}(\mathcal{R})$ の各要素は必ず妥当な負ルールとなる。次に、極小生成子からなる妥当な負ルールの集合 $\mathcal{MR}$ が与えられたとき、 $\mathcal{MR}$ が表現する一般の負ルールの集合を求める関数 $\text{EX}(\mathcal{MR})$ を以下のように定める。

$$\begin{aligned}
EX(\mathcal{MR}) &= \{C_{[MX/X]} \Rightarrow C_{[MY/Y]} \mid \\
&\quad \exists (C_{MX} \Rightarrow C_{MY}) \in \mathcal{MR} : \\
&\quad MX \subseteq X \subseteq \text{clo}(MX), \\
&\quad MY \subseteq Y \subseteq \text{clo}(MY), X \cap Y = \emptyset\}
\end{aligned}$$

補題 1 と 2 および独立性の定義から、以下が明らかに成り立つ。

定理 1 データベース $\mathcal{D}$ 上の妥当な負ルールすべての集合を $\mathcal{R}$ とするとき、 $\mathcal{R} = EX(MI(\mathcal{R}))$ が成り立つ。

定理 1 より、妥当な負ルールの集合 $\mathcal{R}$ を関数 MI により圧縮し、 EX で完全に復元できることが分かる。よって $MI(\mathcal{R})$ は $\mathcal{R}$ の極小生成子を用いた無損失な圧縮表現と見なすことができる。また MI と EX の定義から、負ルールの支持度、確信度、関連度の値の正確な復元も容易であることは明らかである。

EX はその内部で各極小生成子の閉包を求める関数 clo を利用している。残念ながら、閉包の計算にはデータベース $\mathcal{D}$ が必要であることは注意を要する。実際の応用では $\mathcal{D}$ が保存されていない場面がほとんどであるので、圧縮表現の復元を行うためには、極小生成子を用いた負ルール集合だけではなく、極小生成子とその閉包の対応関係も同時に計算して記録しておく必要がある。実際上は、以下の対応関係 $\mathcal{CR}$ を覚えておけば十分である。 $\mathcal{CR}$ より clo 関数は容易に実現できる。

$$\begin{aligned}
\mathcal{CR} &= \{\langle MX, FX \rangle \mid \exists FX \in \text{FCS}(\mathcal{D}) : \\
&\quad MX \in \text{MG}(FX), MX \neq FX\},
\end{aligned}$$

ただし、 $\text{FCS}(\mathcal{D})$ はデータベース $\mathcal{D}$ 中の頻出飽和アイテム集合の集合である。 $\mathcal{CR}$ の定義より、極小生成子 MX を含む対が $\mathcal{CR}$ に含まれない場合は、その閉包 $\text{clo}(MX)$ は MX 自身となることに注意していただきたい。

以上の考察から、負ルール集合の圧縮表現 $MI(\mathcal{R})$ を $\mathcal{R}$ を明示的に求めず、データベース $\mathcal{D}$ から直接求めることも可能である。すなわち、まず最初に $\text{FCS}(\mathcal{D})$ を求め、次に $\mathcal{CR}$ を計算し、その過程で求めた極小生成子を組み合わせ、妥当な負ルールを抽出すればよい。それぞれのステップでアルゴリズム論的な工夫が可能であるが、紙面の都合上、本論文では省略する。

4. 圧縮率の評価実験

本章では、極小生成子を用いる圧縮表現の縮率について実証的評価を行う。実験には、FIMD Repository^{*2}の表 2 に示す 4 つのデータセットを用いた。mushroom, connect は稠密なデータであり、retail, kosarak は疎なデータである。表中の item はデータ中のアイテムの種類数を示し、trans. はトランザクションの総数、ave. は各トランザク

^{*2} Goethals, B.: Frequent Itemset Mining Dataset Repository, <http://fimi.ua.ac.be/>

表 2 実験に使用したデータセット

Table 2 A summary of data sets.

ID	データセット	item	trans.	ave.
#1	mushroom	120	8,124	23.0
#2	connect	130	67,557	43.0
#3	retail	16,470	88,162	10.3
#4	kosarak	41,270	990,002	7.1

表 3 実験結果

Table 3 Results of evaluation experiments.

ID	ms	FI	CFI	MG	N-comp	Comp
#1	0.30	2,735	427	558	814,754	38,180
	0.35	1,189	260	317	98,020	10,032
	0.40	565	140	170	18,944	2,430
#2	0.95	2,201	811	811	1,212,480	288,928
	0.96	1,031	481	481	264,608	88,308
	0.97	487	284	284	62,480	29,156
#3	0.007	315	315	315	74,561	74,561
	0.008	243	243	243	43,254	43,254
	0.009	193	193	193	27,011	27,011
#4	0.01	383	383	383	72,263	72,263
	0.02	121	121	121	6,961	6,961
	0.03	65	65	65	2,051	2,051

ション中に出現するアイテムの平均数である。

実証実験の結果を表 3 に示す。最小確信度 $mc = 0.4$ 、関連尺度は lift^{*3}を用いて $mr = 1.0$ に固定し、最小支持度 ms の値を変化させて負ルールを抽出した。FI, CFI, MG はそれぞれ、頻出アイテム集合、頻出飽和アイテム集合、頻出極小生成子の総数である。N-Comp は妥当な負ルールの総数であり、Comp は極小生成子を用いた妥当な負ルールの総数である。

実験結果より、密なデータセット mushroom, connect はルール数が約 5%～45%に圧縮されている。復元に必要な対応関係 $\mathcal{CR}$ の対の最大数は極小生成子の数 MG と一致するので、圧縮表現のサイズとしては Comp+MG を考えることが適切である。mushroom と connect の場合は Comp+MG は N-Comp の値よりも十分に小さく、圧縮効果が確認できる。これに対して、疎なデータセットでは抽出ルール数がまったく減少していない。これは、すべての頻出アイテム集合が飽和集合であり、また同時に極小生成子となっているためである。なおこのような場合、上に定義した対応関係 $\mathcal{CR}$ は（本質的に恒等関係なので）空集合となることに注意していただきたい。よって関係 $\mathcal{CR}$ を記録しておくための領域的なオーバーヘッドは生じない^{*4}。

^{*3} 詳細は文献 [3] を参照していただきたい。

^{*4} 時間的なオーバーヘッドは、基本となる負ルール抽出アルゴリズムに強く依存するので、本論文での詳細な議論は省略する。ここでは、頻出アイテム集合の極小性の判定計算は一般に容易であり、そのためオーバーヘッドも通常は軽微になることを記しておく。

5. まとめ

本研究では、負ルール集合の極小生成子を利用した圧縮表現を新しく提案した。提案手法は情報損失のない圧縮法であり、実証実験の結果、稠密なデータセットに対しては圧縮率の点でも有効であることが確認できた。稠密なデータに関しては、正ルール集合の圧縮は飽和性の概念が有効であることがよく知られている [4], [7]。しかし、負ルール集合の圧縮には、3章で示したように飽和性の利用は困難である。これまで稠密なデータ上の負ルール集合の圧縮法は知られていなかった。本論文で提案した極小生成子を用いる手法は新しい可能性を開くと考えられる。

疎なデータセット上の頻出アイテム集合や正ルール集合の圧縮には、飽和集合や極小生成子はあまり効果がない場合が多い。この問題に対応するために、情報損失を許容したうえで飽和性の概念を拡張した圧縮法 [8] などが開発されているが、負ルール集合に関する研究は進んでおらず、今後の課題である。また本論文では、圧縮復元に関するアルゴリズム論的な考察は省略した。いくつかの工夫が可能であり、稿を改めて発表を行う予定である。

謝辞 本研究の一部は JSPS 科学研究費補助金 (25330256) の援助を受けている。

参考文献

- [1] Cornelis, C., Yan, P., Zhang, X. and Chen, G.: Mining Positive and Negative Association Rules from Large Databases, *Proc. CIS 2006*, LNCS, Vol.4456, pp.613–618 (2006).
- [2] 井出典子, 岩沼宏治, 山本泰生: 負の相関ルールを抽出する高速トップダウン型アルゴリズム, 人工知能学会論文誌, Vol.29, No.4, pp.406–415 (2014).
- [3] 黒岩健歩, 岩沼宏治, 山本泰生: 関連尺度に基づいた負の相関ルール抽出手法の高機能化, 人工知能学会全国大会 (第 28 回) 論文集, 3J3-3 (2014).
- [4] 宇野毅明, 有村博紀: 頻出パターン発見アルゴリズム入門—アイテム集合からグラフまで, 人工知能学会全国大会 (第 22 回) 論文集, 3M1-1 (2008).
- [5] Wang, H., Zhang, Z. and Chen, G.: Mining a Complete Set of Both Positive and Negative Association Rules from Large Databases, *Proc. PAKDD '08*, pp.777–784 (2008).
- [6] Wu, X., Zhang, C. and Zhang, S.: Efficient Mining of Both Positive and Negative Association Rules, *ACM Trans. Information Systems*, Vol.22, No.3, pp.381–405 (2004).
- [7] Zaki, J.M.: Mining Non-Redundant Association Rules, *Data Mining and Knowledge Discovery*, Vol.9, pp.223–248 (2004).
- [8] Xin, D., Han, J., Yan, X. and Cheng, H.: On Compressing Frequent Patterns, *Data & Knowl. Eng.*, Vol.60, pp.5–29 (2007).



岩沼 宏治 (正会員)

1985 年東北大学大学院電子及通信工学専攻修士課程修了。現在、山梨大学大学院総合研究部教授。工学博士。人工知能の基礎（非単調論理、定理自動証明等）や系列データマイニング等の研究に従事。1987, 1989, 1990, 1991 年人工知能学会全国大会優秀論文賞。2004 年 FIT 優秀論文賞。2014 年日本ソフトウェア科学会第 3 回ソフトウェア論文賞。2014 年度人工知能学会研究会優秀賞受賞。



佐生 隼一

2015 年山梨大学工学部コンピュータ・メディア工学科卒業。現在、山梨大学大学院医学工学総合教育部修士課程コンピュータ・メディア工学専攻に在学中。これまで負の相関ルールマイニングの研究と開発に従事。



黒岩 健歩

2014 年山梨大学工学部コンピュータ・メディア工学科卒業。現在、山梨大学大学院医学工学総合教育部修士課程コンピュータ・メディア工学専攻に在学中。これまで負の相関ルールマイニングの研究と開発に従事。



山本 泰生 (正会員)

2005 年神戸大学大学院自然科学研究科電気電子工学修士課程修了。2010 年総合研究大学院大学博士課程修了。博士（情報学）。2009 年山梨大学大学院総合研究部助教。2014 年より JST さきがけ研究員（ビッグデータ統合利活用のための次世代基盤技術の創出・体系化）。仮説推論、システム生物学、データマイニング等の研究に従事。2010 年帰納論理プログラミング国際会議 (ILP2010) 最優秀学生論文賞受賞。2014 年度人工知能学会研究会優秀賞受賞。