Vol.2016-DPS-168 No.4
Vol.2016-SPT-21 No.4
Vol.2016-EIP-74 No.4
2016/11/17

プライバシー保護した分散データマイニングの通信コスト削減手法

小林 由真[1,a)]　王 家宏[1]　児玉 英一郎[1]　高田 豊雄[1]

概要：分散データマイニングにおいて，取り扱うデータの保護は重要な課題である．秘密分散法を用いることにより，データの暗号化と比べて，低い通信コストでデータを保護しながら頻出パターンの抽出ができる．しかし，多数の通信が必要となり，インターネットのような帯域に制限のある環境では，処理時間が増加する恐れがある．本研究は，通信コストを削減した分散データマイニング手法の提案とその性能評価を行う．

An Approach to Reducing Communication Cost of Secure Data Mining

Yuma Kobayashi[1,a)]　Jiahong Wang[1]　Eiichiro Kodama[1]　Toyoo Takata[1]

1. 背景

近年，企業では取り扱うデータをデータベースに蓄積し，データマイニング技術によって有益な情報を抽出することが一般的となっている．医療機関を例とした場合，『その医療機関では，頭痛と肩こりの症状を持った多くの患者は，風邪をひいている』といった関係性のあるデータを抽出する．このデータは相関ルールと呼ばれる．また，相関ルールを出力する技術を相関ルールマイニングと呼ぶ．本研究では，相関ルールマイニングを対象とする．

医療機関は、しばしば複数の地域にまたがって点在している．例えば，盛岡病院や滝沢医院のように，各地域に数多くの医療機関が存在する．このような環境では，複数のデータベースを対象とした相関ルールマイニング技術である分散データマイニングを用いて，共通して頻出するパターンの抽出を行う．このためには，各医院の秘密データを他の医院へ送信して，集計する必要がある．

しかし，秘密データをそのまま送信してしまうと，プライバシー漏洩の危険性が存在する．例として，医療患者のよく併発する症状を発見するために患者のデータを用いる場合を考える．患者のデータをそのまま渡してしまうと，その患者個人が特定される可能性が存在する．

そのため，分散データマイニングのプライバシー保護を課題とした研究が存在する．研究例として，CRDM[1] が挙げられる．この研究では，秘密分散法を用いたプライバシー保護を行っている．具体的には，各医院の秘密データをいくつかの値に分割し，他の医院を経由させてから集計を行う．このプライバシー保護手法を用いることで，各医院が秘密データの断片しか見ることができない状態でデータマイニングを行うことができる．秘密分散法を用いてデータマイニングを行うことで，暗号化などの負荷の重い処理を行わずにプライバシー保護を行うことができる [1].

しかし，秘密分散法には多数の通信が必要となる．このため，帯域の狭い環境で CRDM を用いたデータマイニングを行った場合，処理時間が増加する恐れがある．

本研究では，秘密分散法を用いた分散データマイニングの通信コストの削減を目的とする．提案手法として，データマイニングを行う環境に合わせて，適切な通信数になるように事前に通信相手を調整することで，通信コストの削減を行う．性能評価の結果，通信コストの削減が確認できた．特にデータマイニングに参加するノードが多数存在する場合，通信コストを大きく削減できる．

本研究により，ネットワーク帯域の狭い環境にあるデータベースを対象としたデータマイニングでも，処理時間の

[1] 岩手県立大学ソフトウェア情報学研究科
[a)] g231o010@s.iwate-pu.ac.jp

Vol.2016-DPS-168 No.4
Vol.2016-SPT-21 No.4
Vol.2016-EIP-74 No.4
2016/11/17

増加を抑えることができる．これにより，分散データマイニングができる対象の拡大と分散データマイニングの性能の向上に貢献が期待できる．

第 2 節では，分散データマイニングに関する関連研究の説明と問題点について述べる．第 3 節では，提案手法の説明と提案アルゴリズムについて述べる．第 4 節では，提案アルゴリズムの性能評価の結果を述べ，最後にまとめとして結論を述べる．

2. 関連研究

基本的な分散データマイニングアルゴリズムとして，FDM[2] が挙げられる．この研究は複数のデータベースを対象とした環境で Apriori アルゴリズム [3] を用いて共通する相関ルールを抽出する．しかし，FDM はデータマイニングに用いる自ノードのプライベートデータであるアイテムセットとサポート値そのものを他ノードへ送信してしまうため，情報流出の危険性が存在する．

FDM のプライバシー保護を課題とした研究として SFDM[4] が挙げられる．この研究では，データマイニングに用いるアイテムセットとサポート値を暗号化することでプライバシー保護を行っている．しかし，暗号化を行うために高い計算コストが必要となるため，処理時間が増加してしまう．

分散データマイニングにおけるプライバシー保護と計算コストを課題とした研究として，CRDM[1] がある．この研究では、アイテムセットのサポート値をランダムな値で分割し，他ノードを経由させてから集計を行う秘密分散法を用いる．これにより，暗号化などの負荷の重い処理を行わずにプライバシー保護を行うことができる．また，ノード間の結託による情報漏洩に対する耐性 (結託耐性) が高く，情報漏洩のためには攻撃対象のノード以外のノードすべてと結託する必要がある．しかし，プライバシー保護のために多数の通信が必要となる．このため，帯域の狭い環境でデータマイニングを行った場合，処理時間が増大する可能性が存在する．

3. プライバシー保護した分散データマイニングにおける通信コストの削減手法

CRDM を用いた分散データマイニングには，多数の通信が必要となる．本研究では，CRDM はデータマイニングに参加するノード数の増加によって，過剰なプライバシー保護性能を提供してしまうことに注目し，利用者の要求に合わせて結託耐性数を設定できるようにアルゴリズムを変更することによって，通信コストの削減を図る．

3.1 システムモデル

全国に点在する医療機関が協力してこのシステムを用いて，患者の訴えた症状の履歴から併発しやすい症状の組み合わせを相関ルールとして抽出することを考える．症状やその ID などのデータマイニングに用いるデータの要素をアイテムと呼び，1 つ以上のアイテムの集合をアイテムセットと呼ぶ．患者が訴えた症状の履歴 (アイテムセット) をトランザクションと呼ぶ．各医院はトランザクションデータベースを保持しており，患者の症状の履歴 (アイテムセット) をトランザクションとして大量に保存している．各医院は同じ種類の症状を扱うものとし，症状の ID は共通しているものとする．

このデータマイニングでは，プライバシー保護のため，ある医院内で症状を訴えた患者の数を他の医院に知られないように相関ルールを抽出する．プライバシーの定義は節 3.2 に後述する．

システムモデルを図 **1** に示す．システムに参加する医療機関をノードとして扱う．ノードの総数を M とする．各ノードには 0 から $M-1$ までの ID が割り振られており，ノードの ID を i，ノードを N_i $(0 \leq i < M)$ と記す．また，N_0 を管理者ノードとして扱い，他を参加者ノードとする．ノード N_i が持つデータベースを DB_i と表す．また，DB_i に格納される各トランザクションを $T_{i,j}$ とし，$DB_i = \{T_{i,1}, T_{i,2}, \cdots\}$，$DB = \cup_{i=0}^{M-1} DB_i$ とする．ノード N_i が持つ，要素の数が k の任意のアイテムセットを $X_{i,k}$，$X_{i,k}$ に対するサポート値を $V_{i,k}$ と表す．$X_{i,k}$ に含まれるアイテムの数を $X_{i,k}$ の長さと呼び，現在求めている頻出アイテムセットの長さを len と表す．アイテムセットが頻出とされる基準を表す最低サポート値を min_sup と表す．min_sup 以上のサポート値を持つアイテムセットを頻出とする．

医療機関のデータマイニングを例とする．各医院は，盛岡病院を N_0，滝沢医院を N_1 のように表記される．また，N_0 はデータベース DB_0，N_1 はデータベース DB_1 を持ち，データベースには患者が訴えた症状の履歴が格納されている．アイテムは “頭痛” や “肩こり” のように表し，アイテムセットは $X_{1,2}$ ={ 頭痛，肩こり } のように表記する．このとき，アイテムセット $X_{1,2}$ の長さは 2 となる．$min\_sup = 0.05$ としたとき，データベース内で 5%以上の割合で出現するアイテムセットを，頻出アイテムセットとしてすべて抽出することが目的となる．

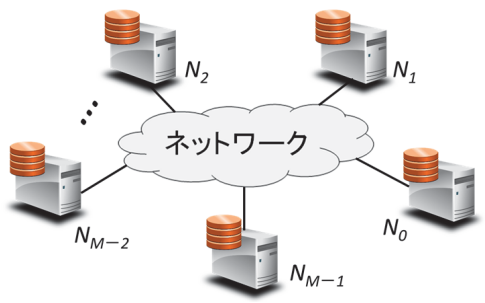

図 1　分散データマイニングのシステムモデル

Vol.2016-DPS-168 No.4
Vol.2016-SPT-21 No.4
Vol.2016-EIP-74 No.4
2016/11/17

3.2 プライバシーの定義

上述したシステムモデルのもとで，各ノードで計算したアイテムセットのサポート値を秘密データとして，プライバシー保護の対象する．

プライバシー保護手法として秘密分散法を用いる．秘密分散法に対する攻撃手法の 1 つとして，各ノードが結託して攻撃対象ノードの分割した秘密データをすべて入手することが考えられる．この結託による攻撃に対する耐性を評価するために，結託耐性数をプライバシー保護性能の指標として用いる．結託耐性数とは，攻撃者が攻撃ノードの秘密データを入手するために，結託が必要なノードの数のことである．

例として，攻撃ノード A がノード B のサポート値を手に入れるためにノード C とノード D との結託が必要である場合，ノード B の結託耐性数は 2 となる．

本提案は，Semi-Honest セキュリティモデル [5] を採用する．これは，ユーザがシステムを改ざんすることなく，システムプロトコルに則って利用することを示すモデルである．ただし，ユーザはシステムから入力と出力データを分析することによって，他ノードのプライベートデータの取得を試みることができる．

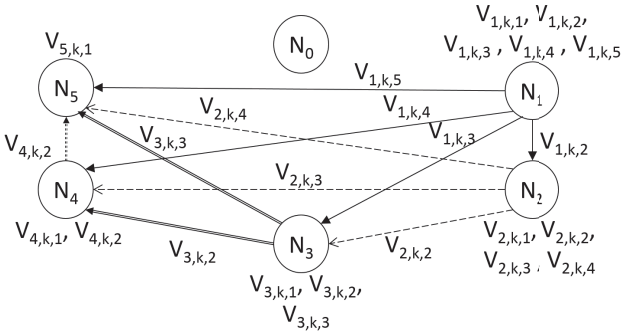

図 2 分散データマイニングの処理の流れ (秘密分散処理)

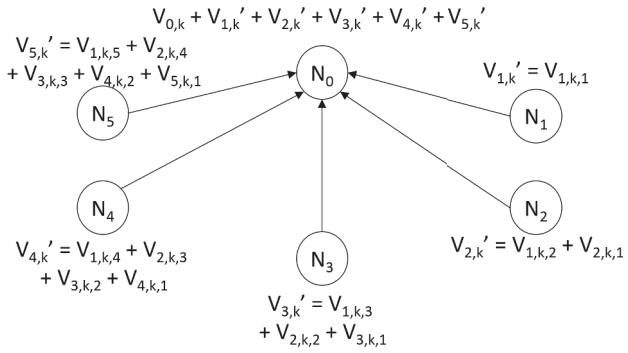

図 3 分散データマイニングの処理の流れ (集計処理)

3.3 CRDM 法の分散データマイニング

CRDM の秘密分散法を用いた分散データマイニングの処理の流れを**図 2** と**図 3** に示す．**図 2** では，はじめに各参加者ノード N_i は，サポート値を断片データとして $M-i$ 個のランダムな値で分割する．その後，分割した断片データをそれぞれ自身よりも ID の高いノードへ送信する．

図 3 では，送信されてきた断片データと自身のための断片データ $(V_{i,k,1})$ をすべて加算し，管理者ノードへ送信する．管理者ノードは，参加者ノードから送られてきた断片データの合計値と自身の秘密データをすべて加算することで，すべてのノードのサポート値の合計を計算し，頻出アイテムセットの計算を行う．その後，頻出アイテムセットから，次の頻出アイテムセットの候補を生成する．上記の処理を長さ 1 のアイテムセット候補から繰り返し実行する．

例として，医療機関のデータマイニングに当てはめた処理の流れを述べる．はじめに，長さ 1 のアイテムセットである { 頭痛 }，{ 肩こり }，{ 風邪 } のサポート値を各ノードで計算する．その後，秘密分散法を用いて各ノードの各アイテムセットのサポート値を管理者ノードで集計する．集計の結果として，min_sup を越えるサポート値を持つアイテムセットを頻出のアイテムセットとして抽出する．次に長さ 1 の頻出アイテムセットから長さ 2 の頻出アイテムセットの候補を生成する．例として，長さ 1 の頻出アイテムセットの { 頭痛 }，{ 肩こり }，{ 風邪 } からは，{ 頭痛，肩こり } や { 頭痛，風邪 }，{ 肩こり，風邪 } が生成される．その後は，現在計算している頻出アイテムセット候補の長さを k とすると，頻出アイテムセット候補のサポート値の計算，秘密分散法を用いた集計，長さ k の頻出アイテムセットの生成，長さ $k+1$ の頻出アイテムセット候補の生成を繰り返すことですべての頻出アイテムセットを生成する．

上記のアルゴリズムによって，プライバシー保護した分散データマイニングを行うことができる．しかし，秘密分散処理には多数の通信が必要となる．特に，ノード数が増加したとき，通信数も大きく増加する．これは秘密分散処理で各ノードが自身より ID の高いすべてのノードに断片データを送信していることが原因である．

本提案手法は，秘密分散法を用いた分散データマイニングに参加するノードの数の増加によって，過剰な結託耐性を提供してしまうことに注目する．ユーザに必要な結託耐性数を事前に指定してもらい，その要求に合致する結託耐性を提供するように CRDM を変更することで，プライバシー保護した分散データマイニングの通信コストを削減する．手法として，はじめに節 3.4 に述べる通信相手決定アルゴリズムを用いて，事前にユーザの要求する結託耐性数を満たすように，秘密分散法で通信する相手を決定する．その後，修正後の CRDM を用いて分散データマイニングを行う．このようにすることで，通信コストの削減を図る．

3.4 通信相手決定アルゴリズム

提案手法では，ユーザに事前に結託耐性数 R $(1 \leq R \leq M-2)$ を指定してもらい，秘密分散法に発生する各ノー

Vol.2016-DPS-168 No.4
Vol.2016-SPT-21 No.4
Vol.2016-EIP-74 No.4
2016/11/17

ドの送信数と受信数の合計を R に制限することによって，ユーザのプライバシー保護に対する要求を満たしつつデータマイニングを行う．

具体的には，秘密分散処理を行う前に，各ノードが断片データをやりとりするノードを決定することで，結託耐性の調整を行う．ここでは，自身の断片データを渡す他ノードのリストを送信先リスト，他ノードから断片データを受け取る必要がある他ノードのリストを受信先リストと呼ぶ．本提案手法では，事前に**図 4** に示すアルゴリズムを用いて，送信先リストと受信先リストを生成する．なお，本アルゴリズムは各参加者ノード内で実行し，すべての参加者ノードの送信先リストと受信先リストを生成する．その後，生成された自身のノードの送信先リストと受信先リストを秘密分散法に用いる．

Input. ユーザの要求する結託耐性数 R
Output. 送信先リストと受信先リスト
Step.1 各参加者ノード N_i $(1 \leq i < M)$ のための送信先リストと受信先リストを初期化するために，i よりも ID が高いすべての参加者ノードを N_i の送信先リストに追加する．また，i よりも ID が低いすべての参加者ノードを N_i の受信先リストに追加する．
Step.2 ID の高い参加者ノード N_i のリストから順に，N_i の送信先ノードの数と受信先ノードの数の合計が R になるまで Step.2.1 と Step.2.2 を行う．
Step.2.1 N_i よりも ID の低い参加者ノードの中で，送信先ノードの数と受信先ノードの数の合計が最大である参加者ノードの ID をすべて求める．
Step.2.2 Step.2.1 で求めたノードの中で，ID が低いノード (N_j) を優先して選択し，N_i の受信先リストから N_j を削除する．また，N_j の送信先リストから N_i を削除する．この 2 つの処理は，N_j の送受信先のノードの数の合計が R であれば実行しない．

図 4 送信先リストと受信先リストの生成アルゴリズム

Step.1 では，管理者ノード以外の参加者ノードを通信相手として選択させる．ただし，ノード同士が自身の断片データを渡しあうとプライバシー保護性能が下がってしまう．これを防ぐために，ID の低いノードが ID の高いノードへ断片データを渡すルールを設ける．そのため，各参加者ノード N_i に対して，i より ID の高いノードを N_i の送信先リストに追加し，i より ID の低いノードを N_i の受信先リストに追加する．

Step.2 では，より多くの通信を必要とする参加者ノードを優先的に通信相手から除外することで，各参加者ノードの通信数を R へ調整する処理を行う．各ノードの通信数が R 以下になることを防ぐために，送受信先のノードの数の合計が R のノードに対しては処理を行わないルールを設ける．Step.2 の処理の流れを**図 5** と**図 6** に示す．ここでは，各ノードの通信数を 2 に制限することを目標とする．

図 5 に示す N_5 に対する処理では，他のすべてのノードの通信数が 4 なので，ID の低い N_1 と N_2 の 2 つを通信先から除外することで，自身の通信数を 2 に制限する．**図 6** に示す N_4 に対する処理では，はじめに通信数が多い N_3 優先して通信先から除外する．その後，さらに通信数を削減するために，ID の低い N_1 を通信先から除外する．この方針により，N_1 のみ通信が多くなるといった通信の偏りを防ぎ，余分な通信を減少させる．

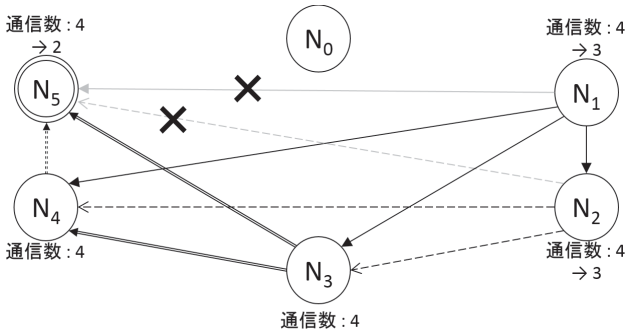

図 5 N_5 の送信先リストと受信先リストの生成の流れ

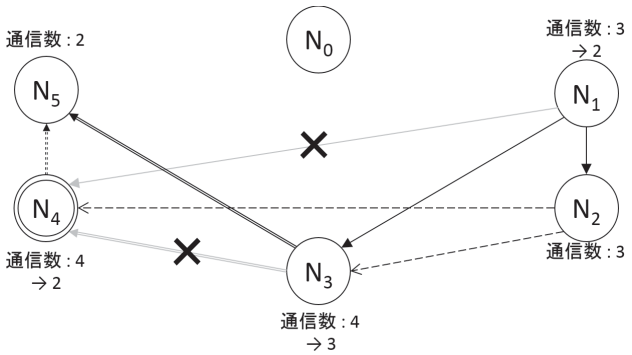

図 6 N_4 の送信先リストと受信先リストの生成の流れ

図 4 に示す送信先リストと受信先リストの生成アルゴリズムを用いて生成した送信リストと受信リストの例を，**表 1** に示す．この例では，ノード数を 6，ユーザの要求する結託耐性数 R を 2 とした場合のリストをそれぞれ示している．表より，各ノードの通信数が指定された結託耐性数 R の値である 2 になっていることがわかる．

表 1 生成される送信リストと受信リストの例

ノードID	送信先リスト	受信先リスト	通信数
1	2,3		2
2	4	1	2
3	5	1	2
4	5	2	2
5		3,4	2

上記のアルゴリズムを用いた後，生成した送信先リストと受信先リストを用いて，各ノードが計算したサポート値の集計を行う．集計処理アルゴリズムを**図 7** に示す．

集計処理アルゴリズムを用いた処理の流れの例を**図 8** と**図 9** に示す．この例では，ノード数は 6，ユーザの要求する結託耐性数を 2 とした．

図 8 では，分割した断片データ $V_{i,k,j}$ を他のノードへ送信する秘密分散処理の流れを示している．各ノードは，

Vol.2016-DPS-168 No.4
Vol.2016-SPT-21 No.4
Vol.2016-EIP-74 No.4
2016/11/17

Input. 各ノードの送信先リストと受信先リスト
Output. 秘密分散法によって，すべてのノードで共通する長さ k の頻出アイテムセットを抽出する．
Step.1 各ノードは，自身が計算したサポート値をランダムな値で送信先ノードの数 +1 の数の断片データに分割する．
Step.2 各ノードは，分割した断片データを送信先リストのノードにそれぞれ送信する．
Step.3 各ノードは，送られてきた断片データと自身のための断片データをすべて加算する．
Step.4 各ノードは，加算した値を管理者ノードへ送信する．
Step.5 管理者ノードは，Step.3 で計算された値を各ノードから受け取り，自身のサポート値と受け取った値の合計を計算する．その後，計算した値を min_sup と比較して頻出アイテムセットを生成する．

図 7 プライバシー保護した分散データマイニングアルゴリズム

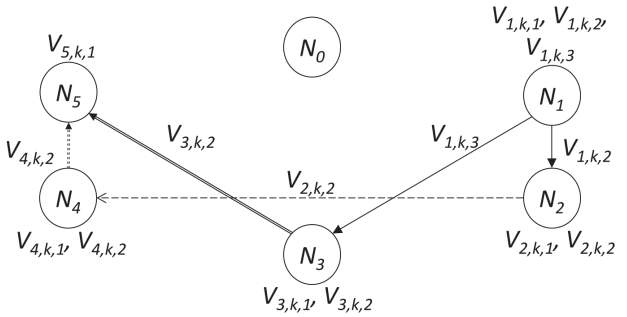

図 8 提案手法を用いた分散データマイニングの処理の流れ (秘密分散処理)

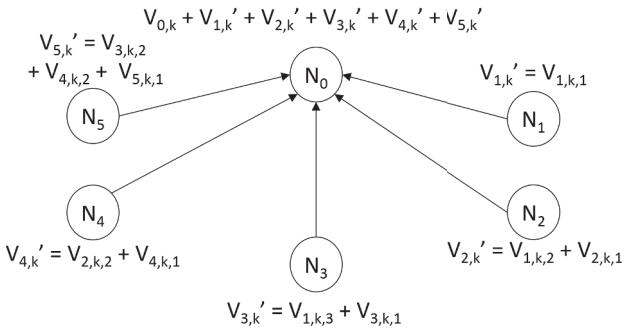

図 9 提案手法を用いた分散データマイニングの処理の流れ (集計処理)

自身が計算したサポート値 $V_{i,k}$ をランダムな値で送信先ノードの数 +1 の数の断片データに分割する．分割した断片データを送信先リストに示されたノードへ送信する．

図 9 では，参加者ノード N_i は受信先リストに示されたノードから受信した断片データと自身のための断片データ $V_{i,k,1}$ をすべて加算し，管理者ノードへ送信する流れを示している．管理者ノードはすべてのノードから断片データの合計値を受け取り，自身のサポート値と加算することですべてのノードのサポート値を集計することが出来る．

以上の処理により，各ノードのサポート値を秘匿しながら分散データマイニングの集計処理を行うことができる．また，各ノードの通信数を R に制限することで，少ない通信数で集計処理を行うことができる．

4. 性能評価

本節では，提案手法の結託耐性の評価を行う．また，分散データマイニングに必要な通信コストについて既存研究である CRDM と比較を行う．

4.1 結託耐性について

関連研究である CRDM の結託耐性数は，$M-2$ となる．この値は，攻撃ノードは攻撃対象のノード以外のすべてのノードと結託する必要があることを示す．しかし，ノード数が増加することによって過剰な結託耐性を提供してしまう．例として，100 ものノードがデータマイニングに参加した場合，実際には結託耐性数は 10 で十分であるにも関わらず，結託耐性数が 98 の環境で処理を行う．

提案手法のアルゴリズムを用いた場合，各ノードの結託耐性数は図 4 に述べたアルゴリズムによって，分割したサポート値の送信と受信の数を R になるように調整されるため，結託耐性数はユーザの要求する結託耐性数 R となる．

4.2 通信コストの比較

本研究では，通信コストを「秘密分散処理で発生する参加者ノード間の通信の数」とする．各ノードは，送信先リストと受信先リストに記されたノードの数だけ通信を行う．本提案手法の送信先リストと受信先リストの生成アルゴリズムによって，各ノードで発生する通信数は R に制限している．そのため，ノード数を M，ユーザの要求する結託耐性数を R とすると，各ノードが通信する数である R と参加者ノードの数 $M-1$ をかけ，送受信の重複を除外するために 2 で割った値が通信コストとなる．

本提案手法によって発生する通信コストの理論値は，$((M-1) \times \frac{R}{2})$ となる．この式を確認するために，ノード数とユーザの要求する結託耐性数をそれぞれ変化させ，理論値と実際の通信コストの比較を行った．比較した結果を図 10 に示す．この図では，ユーザが要求する結託耐性数 R を $\frac{M-1}{2}$ と $\frac{2}{3}(M-1)$ としている．

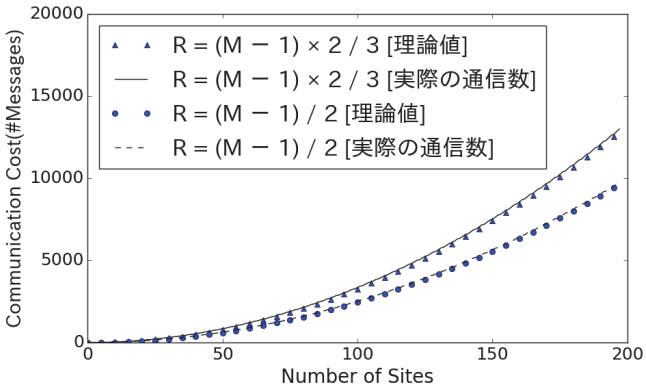

図 10 通信コストの理論値と実際の値の比較

Vol.2016-DPS-168 No.4
Vol.2016-SPT-21 No.4
Vol.2016-EIP-74 No.4
2016/11/17

図 10 より，理論値と実際の通信コストはほぼ一致しており，本提案手法で発生する通信コストは $((M-1) \times \frac{R}{2})$ で導くことができることがわかる．

上記の理論値を用いた通信コストにおける提案手法と CRDM の比較結果を図 11 に示す．提案手法では，ユーザが要求する結託耐性数 R を $\frac{M-1}{2}$ と $\frac{2}{3}(M-1)$ とした場合の結果を比較に用いた．

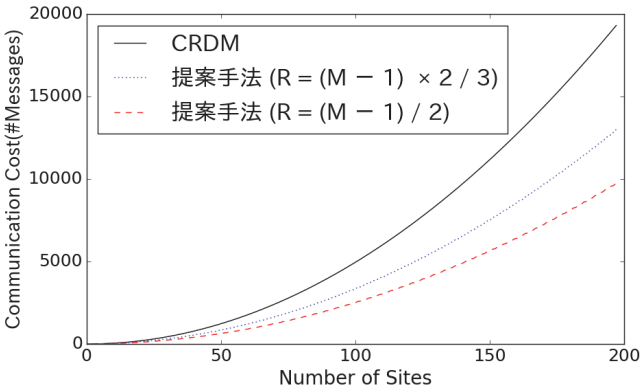

図 11　通信コストの比較

4.3　処理時間の比較

ここでは，処理時間における提案手法と CRDM を比較した結果を示す．データマイニングに用いた実験環境を表 2 に示す．また，分散データマイニングに用いたパラメータを表 3 に示す．表 2 と表 3 に示した環境で処理時間を比較した結果を表 4 に示す．

表 2　実験環境

開発言語	Java
CPU	Intel Core i5-440 3.10GHz
OS	Ubuntu Server
メモリ	8GB
Docker Version	1.71

表 3　分散データマイニングに用いたパラメータ

ノード数	10
各サイトのトランザクション数	10k
アイテムの種類	1000
min_sup	5%
ユーザの要求する結託耐性数 (提案手法のみ)	R = 3, 5

表 4　処理時間の比較

	CRDM	提案手法 (R = 5)	提案手法 (R = 3)
処理時間 *[s]*	*182.5*	*182.4*	*183.0*

4.4　考察

本提案手法を用いることによって，各ノードは指定した結託耐性数 R を満たした環境で分散データマイニングを行うことができる．

さらに図 11 より，本提案手法を用いることによって，R で示した割合まで通信コストを削減できたことがわかる．また，ノード数が増加するにつれて，CRDM よりも通信コストを大きく削減できる．これにより，ネットワーク帯域の圧迫による通信時間の増加を抑えることが期待できる．

しかし，実験環境においては，CRDM と提案手法の処理時間の大きな変化は見られなかった．これは，単一の PC で利用環境を再現したために，ネットワーク帯域が十分に広く，帯域の圧迫による処理時間の増加が発生しなかったことが考えられる．今後は，帯域の狭い環境で処理時間の変化を計測し，再評価を行うことで本研究の優位性を確認する．

5.　まとめ

本論文は，医療機関のような複数地域にまたがって点在している医院のデータベースを対象とした分散型相関ルールマイニングのプライバシー保護と通信コストの削減手法を提案した．利用者の要求する結託耐性数に合わせてプライバシー保護処理に必要な通信を制御することで，通信コストの削減を図った．性能評価では，要求する結託耐性数を満たす環境でデータマイニングができることを示した．さらに，通信コストについて関連研究である CRDM と比較を行い，少ない通信数でデータマイニングが行えることを示した．これにより，ネットワーク帯域の圧迫による通信時間の増加を抑えることが期待できる．今後は，帯域の狭い環境で処理時間の変化を計測し，再評価を行うことで処理時間における本研究の優位性を確認する．

参考文献

[1] S. Urabe, J. Wang, E. Kodama, T. Takata : A High Collusion-Resistant Approach to Distributed Privacy-preserving Data Mining, IPSJ Transactions on Databases, Vol. 48, No. SIG 11, pp.104–117, 2007.

[2] D.W. Chung, J. Han, V.T. Ng, A.W. Fu, and Y. Fu : Fast Distributed Algorithm for Mining Association Rules, Proc. International Conference on Parallel and Distributed Information Systems, pp.31–42, 1996.

[3] R. Agrawal, R. Srikant : Fast Algorithms for Mining Association Rules in Large Databases, Proc. 20th International Conference on Very Large Data Bases, pp.487–499, 1994.

[4] M. Kantarcioglu and C. Clifton : Privacy-preserving Distributed Mining of Association Rules on Horizontally Partitioned Data, IEEE Transaction on Knowledge and Data Engineering, Vol.16, No.9, pp.1026–1037, 2004.

[5] Goldreich, O: Secure multi-party computation (working draft), Available from <http://www.wisdom.weizmann.ac.il/~oded/pp.html>(Sept. 1998).