DEIM Forum 2011 C5-4

複数パーティーにおける統合したグラフのプライバシ保護リンク解析

森井　正覚† 　佐久間　淳†† 　佐藤　一誠† 　中川　裕志†††

† 東京大学大学院情報理工学系研究科　〒 113–8656 東京都文京区本郷 7-3-1
†† 筑波大学システム情報工学研究科　〒 305–8577 茨城県つくば市天王台 1-1-1
††† 東京大学情報基盤センター　〒 113–8656 東京都文京区本郷 7-3-1
E-mail: †{morii,sato}@r.dl.itc.u-tokyo.ac.jp, ††jun@cs.tsukuba.ac.jp, †††n3@dl.itc.u-tokyo.ac.jp

あらまし　リンク解析はエンティティとそれらの関係を表すリンクによって表現されたグラフ構造から有用な情報を抽出する手法である．複数のパーティが秘密に保持するグラフを集めて統合することにより一つのグラフを構成し，エンティティのランク付けすることを考える．既存のリンク解析手法をそのまま用いると，各パーティが保持するデータのプライバシが守られない．本稿では，複数のパーティが秘密に保持するグラフを統合したモデルを考え，それらに対するプライバシを保護したリンク解析手法を提案する．我々の手法は，暗号学的ツールを組み合わせて作られており，加重平均をランダムシェアするプロトコルも提案する．

キーワード　プライバシ，データマイニング，リンク解析，ランダムシェア

Multi-Party Privacy-Preserving Link Analysis for Integrated Graphs

Shogaku MORIII†, Jun SAKUMA††, Issei SATO†, and Horoshi NAKAGAWA†††

† Graduate School of Information Science and Technology, The University of Tokyo
7–3–1 Hongo, Bunkyo, Tokyo, 113–8656 Japan
†† Graduate School of Systems and Information Engineering, University of Tsukuba
1–1–1 Tennodai, Tsukuba, Ibaraki, 305–8577 Japan
††† Information Technology Center, The University of Tokyo
7–3–1 Hongo, Bunkyo, Tokyo, 113–8656 Japan
E-mail: †{morii,sato}@r.dl.itc.u-tokyo.ac.jp, ††jun@cs.tsukuba.ac.jp, †††n3@dl.itc.u-tokyo.ac.jp

1. はじめに

情報技術の普及に伴い，Web 訪問履歴や購買履歴といった個人情報を扱うサービスが盛んに提供されている．そのような個人情報が漏えいした場合の社会的影響は深刻であり，サービス提供者には慎重な扱いが求められる．しかしながら，複数のサービス提供者の保持する詳細な個人情報を統合して用いるデータマイニングは，実社会における情報活用に大きく貢献すると期待される．そのような状況下で，データのプライバシを保護しながらデータマイニングを実現する研究 (プライバシ保護データマイニング) が盛んに行われている．多くのデータマイニング手法に対して，データのプライバシを保護する手法が提案されているが，本稿では，リンク解析におけるプライバシ保護について扱う．

リンク解析とは，エンティティとそれらの関係を表すリンクによって表現されたグラフ構造から有用な情報を抽出する手法である．既存のリンク解析は，解析者にエンティティのリンク構造全体が見えているということを前提としている．しかしながら，人間関係や企業取引などの実世界でのリンク情報は公であることは稀である．Sakuma らは，そのような秘密のリンク関係をもつエンティティのグラフにおけるプライバシモデルを定義し，それらのモデルに基づくプライバシを保護したリンク解析を提案した [3]．彼らの提案したリンク解析は自身に関係するリンク情報しか知り得ないモデルにおけるリンク解析である．それゆえ，彼らの研究ではグラフ上のエンティティ単体を一つのパーティと見なしている．我々の研究では，グラフ上の複数のエンティティ(例: 通信事業者の顧客またはシンクタンクによって調査された企業) に関するデータベースをパーティ (例: 通信事業者またはシンクタンク) が秘密に保持することを想定する．パーティがあるグラフに関する部分的なリンク情報を保持しており，そのリンク情報を他パーティに対しては知らせたくない状況を考える．我々の手法は，Sakuma らの手法と

以下の三つの点で異なっている．第一に，各パーティは複数のエンティティについてのリンク情報を保持していることである．第二に，各パーティは同じエンティティについて異なるリンク情報を保持しているかもしれないことである．最後に，パーティは単体のエンティティである必要はないということである．我々は，複数のパーティが秘密に保持するグラフを統合したグラフに対してのリンク解析を提案する．

本稿では，各パーティ $\mathrm{P}_k (k = 1, \ldots, l)$ は重みつき有向グラフを保持しており，計算能力が多項式オーダーに制限された計算機を保持していることを想定する．各パーティが保持するリンクとリンクの重み情報は他のパーティには公開したくない秘密情報とする．また，任意の 2 パーティ間の通信は，常に可能であるとする．我々は，そのような秘密情報を含むいくつかのグラフを統合したグラフに対する統合モデルを考える．

機密性やプライバシの問題のために，各エンティティ間のリンク情報が他パーティに対して公にできないことも少なくない．もし秘密のリンク情報を統合したネットワークを対象として，その秘匿性を損なうことなく安全にリンク解析を適用できれば，現実の多様なネットワークからの情報抽出を可能にする．本稿では，その統合したグラフに対する安全なリンク解析アルゴリズムを提案する．

2. リンク解析

ノード集合 $V = \{1, \ldots, n\}$，リンク集合 $E = \{e_{ij}\}$，および重み行列 $W = (w_{ij})$ からなる非負の重み付き有向グラフ $G = (V, E, W)$ を考える．エンティティはノードとして抽象化される．ノード ij 間にリンクが存在しなければ，$w_{ij} = 0$ とする．ノード i の度数は $d_i = \sum_{j \in V} w_{ij}$ と定義される．$D = \mathrm{diag}(d_1, ..., d_n)$ を度数行列と呼ぶ．ノード ij 間にリンクが存在したら (i, j) 成分が 1，そうでなければ 0 であるような行列を隣接行列 $A = (a_{ij})$ とする．

Spectral Ranking： リンク解析は，与えられたグラフのリンク構造の特徴を考慮して各ノードに何らかのスコアを与えるアルゴリズムである．グラフ上のマルコフランダムウォークにおける定常分布密度によりノードのスコアを計算する方法を spectral ranking と呼ぶ．ノード i からノード j に，確率 p_{ij} で遷移するマルコフ連鎖を考える．ただし状態遷移確率行列 $P = (p_{ij})$ を $P = D^{-1}W$ として定義する．定常分布 $\boldsymbol{x} = (x_1, \ldots, x_n)^T$ は遷移後もその分布を変えないことから以下を満たす:

$$\boldsymbol{x}^T = \boldsymbol{x}^T P. \tag{1}$$

ただし $\sum_i x_i = 1$ である．この定常分布は，P^T の最大の固有値 $(= 1)$ と対になる固有ベクトル (主固有ベクトル) に対応することが知られている．主固有ベクトルの計算にはべき乗法がしばしば用いられる．初期値として $\sum_i x_i^{(0)} = 1$ なるベクトルを与え，以下の更新式を繰り返す:

$$(\boldsymbol{x}^{(t)})^T \leftarrow (\overline{\boldsymbol{x}}^{(t-1)})^T P, \quad \overline{\boldsymbol{x}}^{(t)} \leftarrow \frac{\boldsymbol{x}^{(t)}}{\| \boldsymbol{x}^{(t)} \|}. \tag{2}$$

3. 問題の定式化

本章では，各パーティが複数のエンティティのリンク情報を秘密に保持している状況で，それらのグラフを統合したグラフにおけるプライバシモデルを定義する．そして，そのモデルに基づいたリンク解析を定義する．各パーティ P_k $(k = 1, \ldots, l)$ が重みつき有向グラフ G^k $(k = 1, \ldots, l)$ をそれぞれ秘密に保持しているとする状況を考える．

定義 1. *(Additive integrated private-weighted graph ; Additive IPWG)* パーティ P_k $(k = 1, \ldots, l)$ はグラフ $G^k = (V, E^k, W^k)$ だけを知っているとする．ただし，$|V| = n$ とする．有限集合を S とし，リンク統合関数と重み統合関数をそれぞれ $g(E^1, \ldots, E^l) = \cup_{k=1}^{l} E^k$ と $h((E^1, W^1), \ldots, (E^l, W^l)) = \sum_{k=1}^{l} W^k$ とする．それらの統合関数を用いて，*additive integrated private-weighted graph* G を以下で定義する．

$$G = (V, g(E^1, \ldots, E^l), h(W^1, \ldots, W^l)) \tag{3}$$

続いて，IPWG 上でのリンク解析を定義する．$f : \mathbb{R}^{n \times n} \to \mathbb{R}^n$ をリンク解析のためのスコアリング関数とする．f は重み行列 $W \in \mathbb{R}^{n \times n}$ を入力として，スコアベクトル $\boldsymbol{x} \in \mathbb{R}^n$ を出力する．このとき secure integrated link analysis は以下のように定義される．

定義 2. *(Secure integrated link analysis)* $G = (V, E, W)$ を *IPWG* とする．*Secure integrated link analysis* の実行後，$f(W) \to \boldsymbol{x}$ は正しく評価され，各パーティは $\boldsymbol{x}$ を知るが，それ以外の知識は得ない．

4. 暗号学的ツール

本章では，secure integrated spectral ranking (SISR) を実現するために必要ないくつかの暗号学的ツールを導入し，weighted average random share protocol を提案する．

準同型性公開鍵暗号： 公開鍵暗号系において，暗号化は公にされた公開鍵 pk を，解読には受信者のみが保持する公開鍵に対応した秘密鍵 sk を用いる．平文 m について，$c = \mathrm{Enc}_{\mathsf{pk}}(m; \rho)$ は m の確率暗号による暗号化を，$m = \mathrm{Dec}_{\mathsf{sk}}(c)$ はその解読をあらわす．ρ が $\mathbb{Z}_N (= \{0, 1, \ldots, N-1\})$ 上で一様ランダムに選ばれたならば，暗文 c も同様に $\mathbb{Z}_N$ で一様ランダムに分布する．加法的準同型公開鍵暗号は，秘密鍵の知識なしに，暗文同士の加算

$$\mathrm{Enc}_{\mathsf{pk}}(m_1 + m_2; \rho) = \mathrm{Enc}_{\mathsf{pk}}(m_1; \rho_1) * \mathrm{Enc}_{\mathsf{pk}}(m_2; \rho_2) \tag{4}$$

が可能である．ここで，ρ_1 か ρ_2 の少なくともどちらか一つが $\mathbb{Z}_N$ 上で一様ランダムならば，同様 ρ は一様ランダムである．この性質に基づき，定数 κ と暗文 $\mathrm{Enc}_{\mathsf{pk}}(m; \rho)$ の乗算が，$*$ の繰り返しにより以下のように実現される．

$$\mathrm{Enc}_{\mathsf{pk}}(\kappa m; \rho) = \prod_{i=1}^{\kappa} \mathrm{Enc}_{\mathsf{pk}}(m; \rho_i) = \mathrm{Enc}_{\mathsf{pk}}(m)^{\kappa}. \tag{5}$$

ρ は $\rho_1, \ldots, \rho_\kappa$ の少なくともどちらか一つが $\mathbb{Z}_N$ 上で一様ランダムならば，同様に一様ランダムである．以降は，簡単のために乱数 ρ は表示しない．

(u, y)-閾値暗号系では，u パーティが共通の公開鍵 pk を保持し，各パーティはそれぞれ異なる秘密鍵 $\mathsf{sk}^i, \ldots, \mathsf{sk}^j$ を保持している．各パーティは共通の公開鍵により任意のメッセージを暗号化可能である．一方，解読には，少なくとも y 以上のパーティのグループが協力し，公開鍵とそれぞれのノードが持つ decryption shares $\mathrm{Dec}_{\mathsf{sk}^i}(c), \ldots, \mathrm{Dec}_{\mathsf{sk}^j}(c)$ を引数にとる recovery アルゴリズムを実行する必要がある．本稿で示すプロトコルは，加法的準同型性を持つ閾値暗号系を用いる．

ランダムシェア： 行列 $X \in \mathbb{Z}_N^{n_1 \times n_2}$ の各要素 $x_{ij} \in \mathbb{Z}_N$ (for all i, j) が，$\sum_{k=1}^{l} r_{ij}^k \bmod N = x_{ij}$, を満たすように $\mathbb{Z}_N$ から一様ランダムに選択された $r_{ij}^1, \ldots, r_{ij}^l$ に分割されているとする．パーティ $\mathrm{P}_1, \ldots, \mathrm{P}_l$ が X を知らずに $r_{ij}^1, \ldots, r_{ij}^l$ をそれぞれ保持しているとき，これを X のランダムシェアによる秘密共有，と呼ぶ．この論文中では，すべての演算が有限体 $\mathbb{Z}_N$ 上で行われるように，N は十分大きい自然数とする．

Secure Scalar Product Protocol： パーティ P_1 と P_2 がそれぞれベクトル $\boldsymbol{x}^1 = (x_1^1, \ldots, x_{n_1}^1)^T$ と $\boldsymbol{x}^2 = (x_1^2, \ldots, x_{n_1}^2)^T$ を保持しているとする．それらの内積のランダムシェアを計算するために，Goethals らによって提案された secure scalar product protocol [2] を用いる．以下では，secure scalar product protocol を $\mathcal{P}_{\mathrm{SSP}}$ と略記し，アルゴリズムの説明に用いる．

Weighted Average Random Share Protocol： パーティ P_k $(k = 1, \ldots, l)$ が自然数 x^k, a^k のペア (x^k, a^k) を保持しているとする．各パーティは互いに協力し，$\sum_{k=1}^{l} x^k / \sum_{k=1}^{l} a^k$ をランダムシェアによって秘密共有を試みる．つまり，以下の機能を持つ安全なプロトコルが必要である．

$$((x^1, a^1), (x^2, a^2), \ldots, (x^l, a^l)) \mapsto (r^1, r^2, \ldots, r^l), \qquad (6)$$

ここで，r^k $(k = 1, \ldots, l)$ は $\sum_{k=1}^{l} r^k = \sum_{k=1}^{l} x^k / \sum_{k=1}^{l} a^k$ を満たすような一様ランダムな数である．式 (6) は，パーティ P_k がプロトコルに入力として (x^k, a^k) を与え，出力として r^k を受け取ることを意味する．

以下では，weighted average random share protocol を $\mathcal{P}_{\mathrm{WARS}}$ とおく．閾値準同型性公開鍵暗号系の鍵集合を $\mathcal{K} = \{\mathsf{pk}, \mathsf{sk}^i, \ldots, \mathsf{sk}^j\}$ とする．秘密鍵は，$\mathsf{sk}^i, \ldots, \mathsf{sk}^j$ と m 個に分割され，sk^i はパーティ P_i のみが保持する．公開鍵 pk は全ノードが共有する．また，$\Delta \geqq \sum_{k=1}^{l} a^k$ であり，Δ はすべてのパーティが知っているとする．$\mathcal{P}_{\mathrm{WARS}}$ のアルゴリズムを手続き 1 に示す．

全てのパーティが semi-honest に振舞うことを想定する．つまり，各パーティは定められたプロトコルを逸脱しないが，実行途中で受け取った全ての情報から他パーティの情報を推測しようとする．このとき，以下の定理が示される．

定理 1. 全てのパーティが *semi-honest* に振舞い，かつ u パーティ以上の結託がないならば，手続き *1* に示されたプロトコルは正しくかつ安全に加重平均のランダムシェアを計算する．

手続き 1 Weighted Average Random Share Protocol

Require: 公的な入力: $\Delta \geqq \sum_{k=1}^{l} a^k$ なる Δ.
Require: パーティ P_k $(k = 1, \ldots, l)$ の秘密の入力: x^k, a^k.
Require: 鍵の設定: すべてのパーティが共同で鍵集合 $\mathcal{K} = \{\mathsf{pk}, \mathsf{sk}^i, \ldots, \mathsf{sk}^j\}$ を生成し，pk はすべてのパーティが所持し，sk^i はパーティ P_i だけが所持するように配布する．

1. パーティ P_1 は，暗文のベクトル $\boldsymbol{c} = \left(\mathrm{Enc}_{\mathsf{pk}}(1), \mathrm{Enc}_{\mathsf{pk}}\left(\frac{\Delta!}{1}\right), \ldots, \mathrm{Enc}_{\mathsf{pk}}\left(\frac{\Delta!}{\Delta}\right)\right)$ を計算する．$\boldsymbol{c}[i]$ を $\boldsymbol{c}$ の i 番目の要素とする．
2. パーティ P_k $(k = 1, \ldots, l)$ は以下を行う．
   (a) パーティ P_k は $\boldsymbol{c} \leftarrow (\boldsymbol{c}[a^k + 1], \boldsymbol{c}[a^k + 2], \ldots, \boldsymbol{c}[\Delta + 1], \boldsymbol{c}[1], \boldsymbol{c}[2], \ldots, \boldsymbol{c}[a^k])$ を計算する．
   (b) パーティ P_k は $\boldsymbol{c} \leftarrow (\boldsymbol{c}[1] \cdot \mathrm{Enc}_{\mathsf{pk}}(0), \boldsymbol{c}[2] \cdot \mathrm{Enc}_{\mathsf{pk}}(0), \ldots, \boldsymbol{c}[\Delta + 1] \cdot \mathrm{Enc}_{\mathsf{pk}}(0))$ を計算する．
   (c) パーティ P_k は $\boldsymbol{c}$ をパーティ P_{k+1} へ送る．
3. パーティ P_l は $\boldsymbol{c}[1]$ を全てのパーティに知らせる．
4. パーティ P_k $(k = 1, \ldots, l)$ は以下を行う．
   (a) パーティ P_k は乱数 $s^k \in_r \mathbb{Z}_N$ を生成する．
   (b) パーティ P_k は $c_k = x^k - s^k$ をパーティ $\mathrm{P}_{k+1 \bmod l}$ に送る．パーティ $\mathrm{P}_{k+1 \bmod l}$ は $c'_k = \boldsymbol{c}[1]^{c_k + s^{k+1 \bmod l}}$ をパーティ P_k に送る．
   (c) パーティ P_k は recovery アルゴリズムを実行し，$r^k = \mathrm{Dec}_{\mathsf{sk}}(c'_k)/\Delta!$ を得る．

5. Secure Integrated Spectral Ranking

本章では，前章に示した定義に基づく IPWG に対する SISR を提案する．SISR の全体的な手順を手続き 2 に示す．以下の節では，SISR の各ステップについての説明を行う．

5.1 確率遷移行列のランダムシェア

問題の定式化で述べたように，リンクとリンク間の重みは公にすべき情報ではない．Spectral ranking には IPWG から計算される確率遷移行列 P が必要であるが，各パーティは P について何も情報を得ない状況で，SISR を実現したい．そのために，P をランダムシェアで秘密共有する．ランダムシェアを計算するために，$\mathcal{P}_{\mathrm{WARS}}$ を用いる．具体的には，パーティ P_k の入力を $(w_{ij}^k, \sum_{j=1}^{n} w_{ij}^k)$ として，$\mathcal{P}_{\mathrm{WARS}}$ を実行し，その出力を P^k の (i, j) 成分とすればよい．もし，$\mathcal{P}_{\mathrm{WARS}}$ の出力が整数でなければ，全パーティが十分大きい自然数をかけることによって整数に拡大すればよい．

5.2 初期設定

手続き 2 のステップ 1 で各パーティは確率遷移行列 $P = D^{-1}W$ をランダムシェアにより秘密共有する．典型的な暗号系は整数のみを引数に取るため，p_{ij}^k は十分に大きい定数 L に基づいて，$b_{ij}^k (= Lp_{ij}^k) \in \mathbb{Z}_N$ となるように拡大される．同様の理由により，初期定常分布 $\boldsymbol{q}$ も十分に大きい定数 K を用いて，$\sum_{k=1}^{l} \sum_{i=1}^{n} q_i^k = K$ となるように初期化する．

5.3 べき乗法

IPWG である $G = (V, E, W)$ に対して，$B = LP$ を l パーティでランダムシェアによる秘密共有をする．ここで，$B = \sum_{k=1}^{l} B^k$ であり，パーティ P_k は，B^k のみを知っている．また，パーティ P_k のみが知っているベクトル $\boldsymbol{q}^{(t),k}$ に対して，$\boldsymbol{q}^{(t)} = \sum_{k=1}^{l} \boldsymbol{q}^{(t),k}$ とおく．正規化ステップを省略した

手続き 2 Secure Integrated Spectral Ranking

Require: 公的な入力: $K \in \mathbb{Z}_N, L \in \mathbb{Z}_N$ $s.t.$ $Lp_{ij}^k \in \mathbb{Z}_N$ for all i, j, k, IPWG の種類.

Require: パーティ P_k $(k = 1, \ldots, l)$ の秘密の入力: W^k, A^k.

Require: 鍵の設定: すべてのパーティが共同で鍵集合 $\mathcal{K} = \{\mathsf{pk}, \mathsf{sk}^i, \ldots, \mathsf{sk}^j\}$ を生成し，pk はすべてのパーティが所持し，sk^i はパーティ P_i だけが所持するように配布する.

1. (B のランダムシェア) 各パーティは $\mathcal{P}_{\mathrm{WARS}}$ を用いて $B = LP$ のランダムシェアを得る．パーティ P_k はランダムシェア B^k を保持している．ここで，$B = \sum_{k=1}^{l} B^k$ である.

2. (初期化) パーティ P_k はすべての i について，以下のように設定する.
$q_i^{(0),k} \leftarrow K_i^k$ $s.t.$ $\sum_{k=1}^{l}\sum_{i=1}^{n} K_i^k = K$, $t \leftarrow 1$

3. (べき乗法) 各パーティ P_k は以下の計算を収束するまで繰り返す.
(a) パーティ P_k は $B^k \boldsymbol{q}^{(t-1),k}$ を計算する.
(b) すべての i と $k' \neq k$ なるすべての k' について，パーティ P_k は $\mathcal{P}_{\mathrm{SSP}}$ を用いてランダムシェア $r_{i,k'}^{(t-1),k}, s_{i,k}^{(t-1),k'}$ を得る．ここで，$r_{i,k'}^{(t-1),k} + s_{i,k}^{(t-1),k'}$ は，B^k の i 番目の行と $\boldsymbol{q}^{(t-1),k'}$ の内積である.
(c) すべての i について，パーティ P_k は $\boldsymbol{q}^{(t),k} \leftarrow B^k \boldsymbol{q}^{(t-1),k} + \sum_{k' \neq k} \left(\boldsymbol{r}_{k'}^{(t-1),k} + \boldsymbol{s}_{k'}^{(t-1),k} \right)$ を計算する.
(d) パーティ P_i と ランダムに選ばれたパーティ P_j $(j \in_r \{1, \ldots, l\} \backslash \{i\})$ は収束を判定するプロトコルを実行する．収束していなければ，"未収束" と全てのパーティに知らせる．もしそのようなメッセージがなければ，ステップ 4 へ進み，そうでなければ，ステップ 3(a) へ戻る.

4. (復号) パーティ P_k は，復号を実行し $\boldsymbol{q}^{(t),k}$ を得，それを全てのパーティに知らせる．それゆえ，出力は $\boldsymbol{\pi}^{(t)} = \sum_{k=1}^{l} \boldsymbol{q}^{(t),k} / KL^{t-1} = \boldsymbol{q}^{(t)} / KL^{t-1}$ となる.

べき乗法の更新式は以下のようになる.

$$
\begin{aligned}
B^T \boldsymbol{q}^{(t)} &= \sum_{k=1}^{l} \left((B^k)^T (\boldsymbol{q}^{(t),1} + \ldots + \boldsymbol{q}^{(t),l}) \right) \\
&= (B^1)^T \boldsymbol{q}^{(t),1} + \ldots + (B^1)^T \boldsymbol{q}^{(t),l} + \ldots \\
&\quad + (B^l)^T \boldsymbol{q}^{(t),1} + \ldots + (B^l)^T \boldsymbol{q}^{(t),l}. \qquad (7)
\end{aligned}
$$

各パーティ P_k は，$(B^k)^T \boldsymbol{q}^{(t),k}$ を独自に計算できるので，式 (7) の計算は，$(B^k)^T \boldsymbol{q}^{(t),k'} (\forall k \neq k')$ の安全な計算ができればよいことになる．これは，$\mathcal{P}_{\mathrm{SSP}}$ を用いることにより計算が可能である．各パーティは，$(B^k)^T$ の行と $\boldsymbol{q}^{(t),k'}$ の内積をランダムシェアによって秘密共有する．$\mathcal{P}_{\mathrm{SSP}}$ を用いた後，パーティ P_k と $\mathrm{P}_{k'}$ がランダムシェア $\boldsymbol{r}_{k'}^{(t),k}$ と $\boldsymbol{s}_{k}^{(t),k'}$ をそれぞれ保持しているとする．パーティ P_k は以下のようにして $\boldsymbol{q}^{(t),k}$ を更新できる.

$$
\boldsymbol{q}^{(t+1),k} \leftarrow (B^k)^T \boldsymbol{q}^{(t),k} + \sum_{k' \neq k} \left(\boldsymbol{r}_{k'}^{(t),k} + \boldsymbol{s}_{k'}^{(t),k} \right), \qquad (8)
$$

ここで，$\sum_{k=1}^{l} \boldsymbol{q}^{(t+1),k}$ は，$B^T \boldsymbol{q}^{(t)}$ に等しい．このように，全パーティは式を秘密に更新できる.

プロトコルの安全性については，以下の定理が成り立つ.

定理 2. 全てのパーティが *semi-honest* に振舞い，かつ u パーティ以上の結託がないならば，*SISR* は定義 *2* の意味で安全に *IPWG* の定常分布を計算する.

表 1　確率遷移行列のランダムシェアの計算時間 (秒)

	$n = 10$	$n = 33$	$n = 100$
$l = 2$	8.3	87.6	737.9
$l = 3$	12.4	129.6	1,193.0
$l = 4$	16.2	173.2	1,589.0

6. 実　　験

プロトコルの有効性を検証するための実験を行った．プログラムは Java 1.6.0 で実装し，暗号系としては [1] で 1024-bit の鍵を用いた．実験は，Xeon 5160 3.0GHz 2 core x 2(CPU), 32GB(RAM) の Linux のもとで行った.

提案プロトコルである SISR は，プライバシが考慮されない場合の spectral ranking と同様の結果を返すことが保障されているため，プロトコルがどのような結果を得るか，ではなく，プロトコルの計算効率性およびプロトコル実行の結果として開示される情報を検証の対象とした．SISR で用いられるサブプロトコルとして，我々が提案したプロトコルである $\mathcal{P}_{\mathrm{WARS}}$ を実装し，ランダムに生成した人工データを用いて実験した.

$\mathcal{P}_{\mathrm{WARS}}$ を用いた，確率遷移行列のランダムシェアの時間計算量は，パーティ数 l と エンティティ数 n に依存している．各パーティは同じ計算を実行しないため，l パーティの計算時間の総和を計算時間とみなしている．各パーティ数 $l = 2, 3, 4$ に対する，エンティティ数 $n = 10, 33, 100$ と計算時間の関係を表 1 に示す．以上の全ての場合で，$\Delta = n$ である.

7. 終 わ り に

本稿では，integrated private-weighted graph と呼ばれる情報分割モデルを導入し，それに対応した secure integrated link analysis を提案した．我々は，secure integrated spectral ranking の問題が，加重平均のランダムシェアの計算に変形できることを示し，準同型性公開鍵暗号を用いて，weighted average random share protocol を安全に実現した．このプロトコルは，integrated private-weighted graph の確率遷移行列のランダムシェアの計算を可能にする．べき乗法の更新を安全に実行することにより，統合したグラフの spectral ranking を達成できる．今後の課題として，ノードクラスタリング，リンク予測，頻出構造の発見など，グラフマイニングにおける様々な問題を integrated private-weighted graph 上で解決することがあげられる.

文　　献

[1] I. Damgård and M. Jurik. A generalisation, a simplification and some applications of paillier's probabilistic public-key system. In *Public Key Cryptography*, pages 119–136. Springer, 2001.

[2] B. Goethals, S. Laur, H. Lipmaa, and T. Mielikäinen. On private scalar product computation for privacy-preserving data mining. *Information Security and Cryptology-ICISC 2004*, pages 104–120, 2005.

[3] J. Sakuma and S. Kobayashi. Link analysis for private weighted graphs. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 235–242. ACM, 2009.