グラフ構造解析を用いたソーシャルブックマークにおけるスパマー検出

渡邊　桂太† 　高橋　翼† 　天笠　俊之†,†† 　北川　博之†,††

† 筑波大学大学院システム情報工学研究科　〒 305–8573 茨城県つくば市天王台 1–1–1
†† 筑波大学計算科学研究センター　〒 305–8573 茨城県つくば市天王台 1–1–1
E-mail: †{ein_vogel,tsubasa}@kde.cs.tsukuba.ac.jp, ††{amagasa,kitagawa}@cs.tsukuba.ac.jp

あらまし　近年，Web 上にブックマーク情報を登録，共有，管理するソーシャルブックマークサービスが，良質な情報源として注目を集めている．しかし，スパムサイトをブックマークしたり，特定サイト内の複数ページをブックマークしたりするような悪質なユーザの増加が問題となっている．そのため，ソーシャルブックマークの情報源としての有益性が損なわれつつある．本研究では，グラフ構造解析を用いたソーシャルブックマークにおけるスパマー検出手法を提案する．具体的には，ソーシャルブックマークを，ユーザとページをノードとしたブックマークの関係を表す 2 部グラフとみなす．このグラフに対してグラフ構造解析を用い，単一あるいは複数ユーザのスパム度合いを評価し，シングル ID，マルチ ID スパマーを検出する手法を提案する．実験により，シングル ID, マルチ ID スパマーいずれも高い精度で検出可能であることを確認した．

キーワード　ソーシャルブックマーク，グラフ構造解析，スパム検出

Analyzing Graph Structures for Spammer Detection in Social Bookmarking Systems

Keita WATANABE†, Tsubasa TAKAHASHI†, Toshiyuki AMAGASA†,††, and Hiroyuki KITAGAWA†,††

† Graduate School of Systems and Information Engineering, University of Tsukuba
Tennoudai 1–1–1, Tsukuba City, Ibaraki, 305–8573 Japan
†† Center for Computational Sciences, University of Tsukuba
Tennoudai 1–1–1, Tsukuba City, Ibaraki, 305–8573 Japan
E-mail: †{ein vogel,tsubasa}@kde.cs.tsukuba.ac.jp, ††{amagasa,kitagawa}@cs.tsukuba.ac.jp

Abstract　In recent years, social bookmark services, which enable us register, share, and manage bookmark information on the Web, are attracting increasing attentions as high-quality information sources. However, malicious users that are bookmarking spam sites or particular site's pages, are increasing in the social bookmark services. As a result, there are growing concerns about loss of beneficial social bookmark service as information sources. In this paper, we propose several methods to detect spammers using graph structure analysis. In particular, we model social bookmark as a bipartite graph in which a user or a web page is represented as a node, and a bookmark behavior is represented as an edge. Using graph structure analysis on this bipartite graph, our proposed methods estimate spam degree of users. We show the feasibility of the proposed methods by experimental evaluations.

Key words　Social Bookmark, Graph Structure Analysis, Spam Detection

1. ま え が き

近年, Blog や SNS などのサービスの浸透により, 誰もが容易に Web コンテンツを生成できるようになった．これに伴い，Web に情報を発信するユーザ人口は飛躍的な増加を遂げ，Web コンテンツの蓄積量もまた，爆発的な増加を遂げた．そのため，膨大な Web コンテンツから，有用な情報を識別，抽出する技術の重要性が，近年強く認識されている．

一方，ソーシャルブックマークサービス (Social Bookmark service: SBM) が近年注目を集めている．SBM は，ブラウザに搭載されているブックマーク機能をオンライン上に実現したもので，ブックマーク情報を Web で管理し，不特定多数のユーザと共有する機能を提供している．また，ブックマークしたページに対してタグやコメントでの注釈付けが行うことができ，独自の観点でブックマーク情報の管理が可能である．これまでに，SBM は Web2.0 を代表するサービスの一つとなるまでに成長し，近年の Web 研究において有益な情報源として注目されている．このような，SBM のブックマーク情報を，ユーザの嗜好，興味を反映した有益な情報源として利用し，新しい Web 検索や情報抽出などの研究が近年盛んである [3]〜[5].

現在までに，SBM は多くのユーザを獲得してきたが，一方で宣伝目的に特定のページをブックマークしたり，スパムコンテンツを含むページをブックマークしたりするスパマーが出現し，問題となっている．このようなスパマーのスパム行為の横行により，良質な情報源としての SBM のブックマーク情報の品質低下が懸念されている．近年の SBM では，スパム行為の根源であるスパマーの検出，スパム行為の抑制は重要な課題とされている．

そこで本稿では，悪質なブックマーク行為を行うスパマーのブックマーク情報の統計的特徴に着目し，ユーザのスパム度合いを評価する *ibf* スコアと *idbf* スコアを提案する．*ibf* スコアは，関連研究 [6], [7] で確認された，ページの被ブックマーク数から得られるスパマーの特徴を用い，ユーザのスパム度合いをスコア化する．*idbf* スコアは，関連研究 [8] からスパマーのブックマーク行動の時間的特徴を見出し，この特徴を *ibf* スコアと組み合わせてユーザのスパム度合いをスコア化する．

しかし，複数のユーザ ID を用いてスパム行為を行うスパマーは，前述の *ibf* スコア，*idbf* スコアを用いてスパム度合いを評価することは困難である．そこで，複数ユーザのスパム度合いを評価する *Cibf* スコアを提案する．SBM をユーザとページをノード，ブックマークをエッジとした 2 部グラフとみなし，密なブックマーク構造を持つユーザ群を抽出したとき，*Cibf* スコアでは，ユーザ群毎に被ブックマーク数を評価し *ibf* スコアと同様の特徴から，ユーザ群のスパム度合いをスコア化する．このように，単一ユーザもしくは複数ユーザのスパム度合いをスコア化することにより，教師付き学習データを必要とせずにスパマーの検出が実現できる．

本論文の構成は以下の通りである．2 章では，SBM におけるスパム行為，スパマーの種類と彼らのブックマークの特徴を，関連研究を交えて説明する．3 章では，2 章で定義したシングル ID スパマー, マルチ ID スパマーを検出するための提案手法の詳細を述べる．4 章では，提案手法の検出性能の評価を行い，さらに関連研究の手法と本研究の手法の検出性能の比較を行う．最後に本研究のまとめと今後の課題を 5 章で述べる．

2. 前提となる知識および関連研究

2.1 SBM におけるスパム行為とスパマーの種類

SBM を始め多くのソーシャルメディアサービスの普及に伴い，SMO(Social Media Optimization) と呼ばれる SEO(Search Engine Optimization) に代わる Web 戦略が注目を集めている．SMO とは，特定のサイトやコンテンツの評価や認知度を上げることを目的に，ソーシャルメディアサービスを利用する Web 戦略である．本研究では，過剰な SMO 目的のブックマークやタグ付与もスパム行為として扱う．

本稿では，SBM におけるスパム行為をスパムブックマークとスパムタギングに分類する．スパムブックマークとは，スパムコンテンツを含むページをブックマークする行為，または SMO 目的に特定のサイトをブックマークする行為である．SMO 目的のスパムブックマークの例として，ブックマークを行うユーザ自身が運営・管理しているブログや EC サイトをブックマークする，セルフブックマークと呼ばれる行為が挙げられる．セルフブックマークは，一概にスパム行為として扱うには議論の余地があるが，本研究では，あるユーザのセルフブックマークの対象が，特定サイトのトップから末端のページまで範囲としている場合には，過剰な宣伝行為とみなしスパムブックマークとして扱う．このような，スパムブックマークを高い割合で行っているユーザを，本稿ではブックマークスパマーと呼ぶ．さらに，一つの ID を利用してスパムブックマークを行うブックマークスパマーをシングル ID スパマー，複数の ID を利用してスパムブックマークを行うブックマークスパマーをマルチ ID スパマーと呼ぶ．

スパムタギングとは，ブックマークしたページに対して，不適切なタグを付与する行為である．具体的には,「通常のページに対して，悪質なタグを付与する行為」と「SMO 目的に特定のページに対して，ユーザの閲覧を誘導するタグを付与する行為」が挙げられる．前者のスパムタギングは，タグ付与の対象のページに対して，悪い印象を与える効果がある．後者のスパムタギングは，タグ付与対象のページが，検索結果に出現しやすくすることを目的に行われている．このようなスパムタギングを高い割合で行うユーザを，本稿ではタグスパマーと呼ぶ．

2.2 関 連 研 究

SBM のようなコンテンツ共有サービスのスパムの問題を扱った研究として，Heymann と Koutrika の研究が挙げられる [1], [2]. Heymann ら [1] は, コンテンツ共有サービスにおけるスパム問題の概要を述べた．これらのスパムの対策方法として, スパムを検出する (detection), コンテンツのランク付けによってスパムの効果を減衰させる (demotion), タグ付与回数に上限を設けるなど, ユーザの行動を一部制限することによって, スパム行為を予防 (prevention) する, という三つのアプローチを述べた．Koutrika ら [2] は, コンテンツにタグの付与ができるタギングシステムでのスパムの問題を扱っている．「仮想的なタギングシステム下で有効なスパム対策手法は, 現実世界のタギングシステム下でも有効である」という仮定から, ユーザのタグ付与行動をモデル化し, 仮想タギングシステムを設計し, また, この仮想タギングシステム内で, 有効なスパムの対策手法を提案した．

本研究と同じ SBM のスパマー検出を目的とする研究として, 宗方ら [6], [7] と数原ら [8] の研究が挙げられる．

2.2.1 宗方らのアプローチ

宗方らは，スパマーのブックマークの特徴として,「ブックマークページの多くは，被ブックマーク数が 1 ユーザのみであることが多い」,「ブックマークページをより多くの検索語で検索結果に出現させるために，たくさんのタグを使用する傾向がある」ことを示した．彼らは，日本最大の SBM サービスであるはてなブックマーク(注1)のデータを収集し，ユーザ毎にスパマーと被スパマーのラベル付けをした．ラベル付けしたデータから教師付き学習データを作成し，機械学習によるスパマーの分類器 (決定木) を生成する検出手法を提案した．

2.2.2 数原らのアプローチ

数原らは，スパマーのタグ付与の特徴として,「短期間にタグの付与数が多い」,「スパマーの使用タグ群から算出されるエントロピーが高い」傾向があることを示した．彼らは，bibsonomy(注2)という SBM サービス が提供している，スパマーと非スパマーのラベル付けがされている RSDC'08 データセットを用いている．RSDC '08 データセットは，ECML PKDD 2008 の Discovery Challenge (RSDC '08)(注3)において，ソーシャルスパマー発見タスク (Spam Detection in Social Bookmarking Systems) で提供された機械学習によるスパマー，スパムコンテンツの分類器生成のための教師付き学習データとテストデータである．彼らは，このデータを用いて，前述の特徴を学習のパラメータとした機械学習によってスパマー分類器を生成する検出手法を提案した．

2.3 本研究と関連研究の比較

2.2.1 節と 2.2.2 節で述べた関連研究では，検出対象のスパマーが異なる．2.2.1 節では，ユーザのブックマークしたページの被ブックマーク数と，ブックマークページ数と使用タグの種類の相関を学習のパラメータとしているため，ブックマークスパマーと，一部のタグスパマーの検出に対応している．しかし，マルチ ID スパマーは被ブックマーク数の操作が可能ため，2.2.1 節で述べられた特徴を用いた手法では，マルチ ID スパマーの検出に十分な対応なされていないと考えられる．2.2.2 節では，ユーザ毎にタグ付与行動の不自然さを評価しているため，タグスパマーの検出に特化している．しかし，タグの付与を一切行わないようなブックマークスパマーの検出には十分な対応がなされていないと考えられる．

また，いずれも教師付き学習データを必要とする手法をとっているが，通常，教師付き学習データの入手は困難であることが多いため，教師付き学習データを必要としないスパマー検出手法が必要であると考えられる．さらに，いずれの手法もマルチ ID スパマーの検出への対応が不十分なので，マルチ ID スパマーの検出に対応した手法も重要となる．本研究では，宗方らが報告したシングル ID スパマーの特徴から，マルチ ID スパマーの特徴の類推し，シングル ID スパマーのみならず，マルチ ID スパマーも検出対象とし，かつ教師付き学習データを用いずにスパマーを検出する手法を提案する．

各研究のスパマーの検出範囲を表 1 に示す．

表 1　各研究のスパマー検出範囲

	ブックマークスパマー		
	シングル ID スパマー	マルチ ID スパマー	タグスパマー
宗方 (2009)	○	×	×
数原 (2009)	×	×	○
本研究	○	○	×

3. 提案手法

本研究では，ブックマークスパマーのうちシングル ID スパマーを検出する手法とマルチ ID スパマーの検出する手法を提案する．前者では，先行研究 [10] で提案した *ibf* を基に，ユーザのスパム度合いをスコア化する．後者では，SBM データをユーザとページをノード，ブックマークをエッジとした 2 部グラフと考え，ブックマークのエッジが密になっている構造を抽出し，そこに含まれるユーザ群のスパム度合いをスコア化する．

3.1 シングル ID スパマーの検出手法

2.2.1 節で述べたように，シングル ID スパマーがブックマークするページの特徴として,「被ブックマーク数が 1 であることが多い」ことが指摘されている．そこで本研究では，シングル ID スパマーの特徴を「被ブックマーク数が極めて少ないページを高い割合でブックマークしている」と考え，ユーザのスパム度合いを評価する．このとき，スコアの高いユーザをシングル ID スパマーとして検出する．

3.1.1 Inverse Bookmark Frequency (ibf)

先行研究 [10] では,「被ブックマーク数が極めて少ないページ」を評価する指標として，*ibf (Inverse Bookmark Frequency)* を提案した．*ibf* は，tf-idf 法における大域的重み *idf (Inverse Document Frequency)* の考えに基づくページの評価指標である．*idf* の基本的な考え方は，多くの文書に出現する語は一般的に利用される語と考え低い重みを与え，特定の文書に偏って出現する語は特徴的な語と考えて高い重みを与えるというものである．*ibf* は，ブックマーク対象のページの特徴を表すもので，被ブックマーク数の多いページには低い値が，被ブックマーク数の少ないページには高い値が与えられる．

ページ p をブックマークしているユーザの集合を，$users(p)$ としたときの *ibf* は，以下のように算出する．

$$ibf(p) = \frac{1}{\log_q(|users(p)| + q - 1)} \quad (1)$$

ただし，q は，log 演算の底を表すパラメータである．*ibf* は，q の値が高いと，収束値が高くなり，低いと収束値は低くなるといった特徴があるため,「被ブックマーク数が極めて少ないページ」の特徴を表現するには，q は低い値に設定する必要がある．

3.1.2 ibf スコア (ibf score)

先述で述べたシングル ID スパマーの特徴を「ibf の高いページを高い割合でブックマークしている」とみなし，ユーザのスパム度合いを評価する *ibf* スコア *(ibf_score)* を定義する．*ibf* ス

(注1)：http://b.hatena.ne.jp/
(注2)：http://www.bibsonomy.org/
(注3)：http://www.kde.cs.uni-kassel.de/ws/rsdc08/

コアは，ユーザがブックマークしているページ全ての *ibf* を算出し，平均を取ったものである．*pages(u)* を，ユーザ u がブックマークしているページの集合としたとき，ユーザ u の *ibf* スコアは，以下のように算出される．

$$ibf\_score(u) = \frac{1}{pages(u)} \sum_{p \in pages(u)} ibf(p) \quad (2)$$

ibf の高いページを高い割合でブックマークしているユーザは，*ibf* スコアが高くなるため，シングル ID スパマーである確率が高い．一方，*ibf* の低いページを高い割合でブックマークしているユーザは，*ibf* スコアは低くなり，シングル ID スパマーである確率は低いこと表す．

3.1.3 inverse date bookmark frequency スコア (idbf スコア)

スパマーの特徴として，さらに「特定の期間にまとめてスパム行為を行う」という点が指摘されている [8]．そこでシングル ID スパマーの特徴を「被ブックマーク数が極めて少ないページを特定の日に多くブックマークしている」とみなし，この特徴を用いてユーザのスパム度合いを評価する指標として *idbf(inverse date bookmark frequency)* スコアを定義する．*idbf* スコアは，ユーザのブックマーク数の多い日にブックマークされたページに高い重みを，少ない日には低い重みをページに与え，*ibf* スコアと同様の手順で算出する．*dates(u)* をユーザ u がブックマークを行った日付の集合，*page(u,dt)* を，ユーザ u が dt の日にブックマークしたページの集合としたとき，*idbf* スコアは以下のように算出される．

$$date\_weight(u, dt) = \frac{|pages(u, dt)|}{|dates(u)|} \quad (3)$$

$$normalizer(u) = \sum_{dt \in dates(u)} date\_weight(u, dt) * |pages(u, dt)| \quad (4)$$

$$idbf\_score(u) = \frac{\sum_{dt \in dates(u)} date\_weight(u, dt) \sum_{p \in pages(u,dt)} ibf(p)}{normalizer(u)} \quad (5)$$

date_weight は，ユーザ u が日付 dt にブックマークしたページに付与される重みである．また，*normalizer* は，*idbf* スコアの値域を 0〜1 に正規化するために用いられる．

4.1 節では，*ibf* スコア，*idbf* スコアに閾値 σ を設け，σ 以上の *ibf* スコア，*idbf* スコアを持つユーザを，シングル ID スパマーとして検出し，検出性能を評価する実験を行った．

3.2 マルチ ID スパマーの検出手法

Web スパムにおいて，リンク登録サイトを利用したハブ・オーソリティ型スパムと呼ばれるスパムが確認されている [9]．リンク登録サイトとターゲットサイトをノード，リンクをエッジとした 2 部グラフで表したとき，ハブ・オーソリティ型スパムは，一方向にリンク（エッジ）を密に張ることが確認されている．この密なリンク構造は，2 部クリークもしくは擬似 2 部クリークで捉えることができる．逆に，(擬似) 2 部クリークの検出することで，ハブ・オーソリティ型スパムの検出が可能である (図 1)．そこで本研究では，マルチ ID スパマーの特徴を「ブックマークが類似しているユーザ群のブックマークが，特定のページに偏っている」と考え，類似ユーザ群を抽出し，類似ユーザ群のスパム度合いを評価する．このとき，スコアの高い類似ユーザ群をマルチ ID スパマーとして検出する．

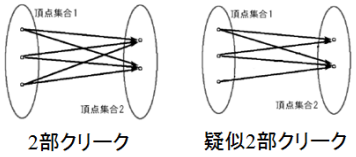

図 1　2 部クリークと擬似 2 部クリークの例

3.2.1 類似ユーザ群の抽出

今回，ユーザと各ユーザがブックマークしているページから構成される隣接行列を利用し，クラスタリングによって類似ユーザ群を抽出した．隣接行列を *ibf* で重み付けしたデータを用い，ユーザをクラスタ分けする (図 2)．*ibf* 値による隣接行列の重み付けは，被ブックマーク数の少ないページを共通にブックマークしているユーザ同士が，同じクラスタにマージされるよう補正する役割を持つ．

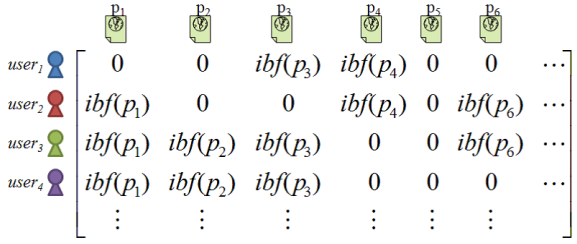

図 2　ユーザベクトル

3.2.2 Cluster ibf (Cibf)

マルチ ID スパマーのスパムブックマーク対象ページの特徴として，「特定の類似ユーザ群から，極めて高い割合でブックマークを受けているページ」であることが考えられる．そこで，抽出した類似ユーザ群毎に，この特徴を評価する指標として，*Cibf* (*Cluster ibf*) を定義する．類似ユーザ群 Q のユーザ集合を Q_{users} とし，Q_{users} に含まれるユーザがブックマークしているページ集合を Q_{pages}，Q_{pages} に含まれるページ p が与えられたとき，*Cibf* は以下のように算出する．

$$Cibf(Q, p \in Q_{pages}) = \frac{1}{log_q(|users(p)| + q - |Q_{users}(p)|)} \quad (6)$$

$$(Q_{users}(p) = Q_{users} \cap users(p))$$

q は log 演算の底を表すパラメータである．*Cibf* は，*ibf* をマルチ ID スパマーのブックマークページの特徴を評価するために拡張した指標のため，3.1.1 節と同様の理由で，q は低い値で設定する必要がある．

3.2.3 Cibf スコア

3.1.2 節と同様に，マルチ ID スパマーの特徴を「*Cibf* の高いページを高い割合でブックマークしている」と考える．そこで，類似ユーザ群のスパム度合いを評価する指標として，*Cibf* スコア (*Cibf_score*) を以下のように提案する．

$$Cibf\_score(Q) = \frac{1}{|Q_{users}|} \sum_{u \in Q_{users}} \frac{1}{|pages(u)|} \sum_{p \in pages(u)} Cibf(Q,p) \quad (7)$$

Cibf スコアは，*Cibf* 値の高いページを多くブックマークしているユーザ群には高いスコアを，*Cibf* 値の低いページを多くブックマークしているユーザ群は低いスコアが与えられるといった特徴がある．

4. 評価実験

本章では，提案するシングル ID スパマーとマルチ ID スパマーの検出手法の性能の評価を行い，その有効性を検証する．

4.1 シングル ID スパマーの検出実験

4.1.1 実験データの概要

本実験では，我々の研究室で収集しているはてなブックマークの SBM データを用いて *ibf* スコア，*idbf* スコアの性能を評価する．実験データの詳細は以下の通りである（表 2）．

表 2　シングル ID スパマーの検出に用いる実験データ

ユーザ数	3,735
ページ数	1,265,208
ブックマーク (エッジ) 数	3,991,955

4.1.2 検出性能の評価 (ibf スコア)

ibf スコアを用いたシングル ID スパマーの検出実験は以下の手順で行う．

（1） 実験データのユーザ全ての *ibf* スコアを算出．

（2） *ibf* スコアの閾値 σ を設定し，$ibf\_score(u) \geqq \sigma$ となるユーザ u をシングル ID スパマーとして検出．

（3） 検出ユーザがスパマーかどうかを目視で判別．

実験結果を表 3 に示す．

表 3　シングル ID スパマーの検出結果 (*ibf* スコア)

ibf スコアの閾値 σ	検出ユーザ数	検出スパマー数	適合率
0.70	51	34	0.667
0.75	45	33	0.733
0.80	36	30	0.833
0.85	26	24	0.923
0.90	19	19	1.00

閾値 σ の設定を高くすることにより，検出スパマー数の減少が見られたが，一方で適合率の増加が確認できた．

4.1.3 検出性能の評価 (idbf スコア)

idbf スコアを用いたシングル ID スパマーの検出実験は以下の手順で行う．

（1） 実験データのユーザ全ての *idbf* スコアを算出．

（2） *ibf* スコアの閾値 σ を設定し，$idbf\_score(u) \geqq \sigma$ となるユーザ u をシングル ID スパマーとして検出．

（3） 検出ユーザがスパマーかどうかを目視で判別．

実験結果を表 4 に示す．

表 4　シングル ID スパマーの検出結果 (*idbf* スコア)

idbf スコアの閾値 σ	検出ユーザ数	検出スパマー数	適合率
0.70	77	39	0.506
0.75	56	36	0.643
0.80	44	34	0.773
0.85	30	29	0.967
0.90	23	23	1.00

4.1.2 節と同様，閾値 σ の設定を高くすることにより，検出スパマー数の減少が見られたが，一方で適合率の増加が確認できた．

ibf スコアと *idbf* スコアの検出スパマー数を適合率別に比較したグラフを図 3 に示す．図 3 より，*idbf* スコアは *ibf* スコアより検出スパマー数が多いことが確認できた．

4.1.4 関連手法との性能比較

今回，宗方らの手法とシングル ID スパマー検出性能の比較実験を行った．宗方らが行った実験のデータと結果を表 5,6 に示す．

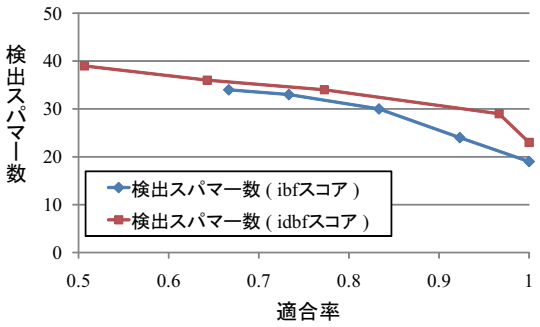

図 3　*ibf* スコアと *idbf* スコア検出性能の比較

表 5　宗方らの実験データ

	教師付き学習データ	テストデータ
ユーザ総数	1000	66
スパマー数	87	17

表 6　宗方らのスパマー検出結果

適合率	0.607
再現率	1

今回行う比較実験は，実験データからランダムに 100 人のユーザを選択し，4.1.2，4.1.3 節と同様の実験手順で，*ibf* スコアと *idbf* スコアの検出性能を評価する．以上の試行を 5 回行い，平均をとった結果を宗方らの手法の検出性能 (表 6) と比較する．比較結果を図 4, 5 に示す．図 4, 5 より，*ibf* スコアは閾

値 σ が 0.6 から 0.8 のとき，*idbf* スコアは閾値 σ が 0.7 から 0.8 のときに，宗方らの手法の F 値を上回る性能が得られることが確認できた．

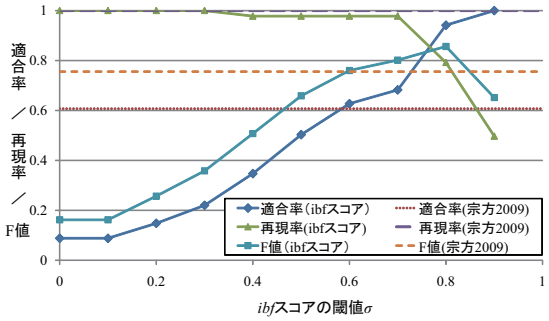

図 4　シングル ID スパマーの検出性能の比較 (*ibf* スコアと宗方らの手法)

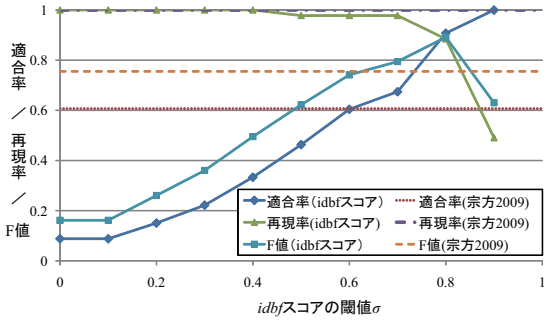

図 5　シングル ID スパマーの検出性能の比較 (*idbf* スコアと宗方らの手法)

4.2　マルチ ID スパマーの検出実験

4.2.1　実験データの概要

4.1 節で示した SBM データは，マルチ ID スパマーを解析する上で十分な規模でなかったため，本実験では，livedoor クリップ(注4)が提供している研究用データセットを用いて，マルチ ID スパマーの検出性能を評価した．実験データは，2009 年 12 月に公開されたものを用い，詳細を表 7 に示す．

表 7　マルチ ID スパマー検出の実験データ

ユーザ数	45,031
ページ数 (被ブックマーク数 $\geqq$ 3 ユーザ)	330,498
ブックマーク (エッジ) 数	2, 467,438

注意点として，本データセットには被ブックマーク数が 3 ユーザ未満のページは含まない仕様となっている．このため，シングル ID スパマーはほとんど含まれていないと考えられる．

(注4)：http://clip.livedoor.com/

4.2.2　検出性能の評価

Cibf スコアを用いたマルチ ID スパマーの検出性能の評価は，以下の手順で行う．また，類似ユーザ群を抽出するためのクラスタリング手法として，bisecting k-means を用いた．

（1） 生成するクラスタ数 k を設定し，k 個の類似ユーザ群を抽出．

（2） 1. で抽出した類似ユーザ群の *Cibf* スコアを算出．

（3） *Cibf* スコアの閾値 σ を設定し，$Cibf\_score(Q) \geqq \sigma$ となる類似ユーザ群 Q をマルチ ID スパマーとして検出．

表 8 に，$k = 1000$, $\sigma = 1$ と設定したときの検出結果を示す．類似ユーザ群を，「ブックマークが完全一致している類似ユーザ群」と「ブックマークが部分一致している類似ユーザ群」に仕分けたところ，「ブックマークが完全一致している類似ユーザ群」に注目した際に，最も高い検出精度が得られた．

表 8　マルチ ID スパマーの検出結果 ($k = 1000, \sigma = 1.00$)

	類似ユーザ群	ブックマークが完全一致している類似ユーザ群	ブックマークが部分一致している類似ユーザ群
総検出数	143	103	40
スパム検出数	134	101	33
適合率	0.937	0.980	0.825

5.　まとめと今後の予定

本研究では，SBM におけるシングル ID スパマー，マルチ ID スパマーを対象に，ユーザのスパム度合いを評価する指標を提案した．シングル ID スパマーの評価手法では，ユーザのスパム度合いを *ibf* スコア，*idbf* スコアで評価し，関連手法との比較を行った．マルチ ID スパマーの評価手法では，ブックマークが類似しているユーザ群を抽出し，ユーザ群のスパム度合いを *Cibf* スコアで評価，指標の妥当性を検証した．

今後は，マルチ ID スパマーの検出実験を様々なパラメータで行い，より詳細に検出性能を評価する．さらに，クラスタリングによる類似ユーザ群の抽出ではなく，クリーク探索アルゴリズムを用いて（擬似）2 部クリークを抽出し，抽出した（擬似）2 部クリークに含まれるユーザ群で *Cibf* スコアの性能評価を行う．また，最適な検出性能が得られるパラメータを自動設定する手法について検討する．

謝　　辞

本研究の一部は科学研究費補助金特定領域研究（# 21013004）による．

文　　献

[1] Paul Heymann, Georgia Koutrika, and Hector Garcia-Molina. Fighting spam on social web sites: A survey of approaches and future challenges. *IEEE Internet Computing*, 11(6):36–45, 2007.

[2] Georgia Koutrika, Frans Adjie Effendi, Zoltán Gyöngyi, Paul Heymann, and Hector Garcia-Molina. Combating spam in tagging systems. In *Proceedings of the 3rd international workshop on Adversarial information rtrieval on the*

web, pages 57–64, New York, NY, USA, 2007. ACM.

[3] 山家雄介, 中村聡史, アダム ヤトフト, 田中克己. ソーシャルブックマークの周期性発見に基づく時期連動型検索ランキング手法. Web とデータベースに関するフォーラム (*WebDB Forum2008*), 2008.

[4] Tsubasa Takahashi and Hiroyuki Kitagawa, "A Ranking Method for Web Search Using Social Bookmarks", Proc. International Conference on Database Systems for Advanced Applications (*DASFAA 2009*), pp. 585-589, Brisben, Australia, April 21 - 23, 2009.

[5] 百田信, 伊東栄典. ソーシャルブックマークに基づく情報発見. データ工学ワークショップ (*DEWS2008*), 2008.

[6] 宗片健太朗, 福原知宏, 山田剛一, 絹川博之, 中川裕志. ソーシャルブックマークにおけるスパマー検出. 情報処理学会第 71 回全国大会, 2009.

[7] 宗片健太朗, 福原知宏, 山田剛一, 絹川博之, 中川裕志. ソーシャルブックマークにおけるスパム検出のための特徴とその評価. 第 8 回情報科学技術フォーラム (*FIT 2009*), 2009.

[8] 数原良彦, 植松幸生, 井上孝史, 片岡良治. ソーシャルブックマークにおけるタグ付与行動に基づくスパマー検出. DBSJ Journal Vol.7 No.4, 2009.

[9] 小野拓史，豊田正史，喜連川優， リンク解析を用いたウェブ上のスパム発見手法に関する一考察， 電子情報通信学会第 17 回データ工学ワークショップ (*DEWS 2006*), 2006.

[10] 渡邊桂太，高橋翼，北川博之，ソーシャルブックマークにおけるユーザ間の類似度を考慮したスパマー検出， 第 1 回データ工学と情報マネジメントに関するフォーラム (*DEIM 2009*), 2009.