Vol.2012-DBS-154 No.3
Vol.2012-IFAT-107 No.3
2012/8/1

複数のユーザアカウントを用いた ソーシャルブックマークスパムの検出

坂倉悠太[†1]　天笠俊之[†2†3]　北川博之[†2]

本稿では，閲覧者の増加や検索エンジン最適化のために多くのユーザアカウントを用いて対象 Web リソースをブックマークする“複数のユーザアカウントを用いた集中的ブックマーク（Intensive Bookmarking using Multiple Accounts, IBMA）”の検出手法を提案する．効率的に IBMA を検出するために，ブックマーク中の Web リソースもしくは Web サイトの共通性に基づいて，ユーザアカウントのクラスタリングをする．具体的には，類似した Web リソースもしくは Web サイトの集合をブックマークしているユーザアカウントをクラスタリングする．このために，二つのブックマーク集合に対する三つの類似度を提案する．実験により，提案手法が，実際のデータセットにおいて IBMA を正しく検出できることを示した．また，類似度ごとの提案手法の精度を評価し，特徴を述べる．

Detecting Social Bookmark Spams using Multiple User Accounts

YUTA SAKAKURA[†1]　TOSHIYUKI AMAGASA[†2†3]　HIROYUKI KITAGAWA[†2]

This paper proposes a scheme of detecting “Intensive Bookmarking using Multiple Accounts” (IBMA), where many social bookmark accounts are used to create bookmark entries linking to the target web resources with the aim of increasing site visitors or optimizing search result ranking. To efficiently detect IBMA, we propose to use clustering social bookmark user accounts according to the similarity with respect to the bookmarked web resources or web sites. Specifically, we cluster users who create bookmarks linking to similar set of web resources or web sites. For this, we propose three similarity measurements over two sets of bookmarks. We experimentally show that the proposed scheme successfully detects IBMA spammers in a real dataset. We also evaluate the accuracy of the proposed scheme with varying the similarity measurements, and characterize them.

1. はじめに

近年，ソーシャルブックマーク（Social Bookmark; SBM）が，ユーザのブックマーク情報を Web 上で保存や管理，検索，共有する手段として広く普及している．SBM では，ユーザが気に入った Web リソースへのリンクを含むブックマークエントリを作成することができ，URL を登録するだけでなく，検索や管理のためにタグやコメントを付与することができる．さらに，それらのブックマークエントリは同じ SBM のアカウント所有者の間で共有されるため，利用者は興味深い Web リソースを容易に見つけることができる．このため，SBM から得られる情報は有用であると考えられ，Web ページの推薦[1]や検索結果の再ランキング[2][3]などに用いられている．

一方，SBM は，ハイパーリンクやコメント，タグを含めたブックマークおよび関連情報を容易に作成，編集できるため，スパムの対象となっている．この特徴を用いて，スパマーは閲覧者数の増加や，検索エンジン最適化（Search Engine Optimization; SEO）[4]による検索エンジンにおける対象ページのランキングを操作しようとする．このようなスパマーによって作成されたブックマークエントリの自動検出が望まれる．

SBM におけるスパマー検出について，多くの関連研究がある．これらの研究は，スパムの対象によって，タグスパムとブックマークスパムの二種類に分類される．前者は，スパマーによってブックマークエントリに付与される不適切なタグを対象とし，実世界のイベントとタグの相関を利用してタグスパムの影響を低減させる方法[5]や，ユーザの行動やリソース，タグなどの特徴に基づいたタグスパムのクラス分類[6][7][8]がある．

一方，後者のスパムブックマークとブックマークスパマーが本稿の検出対象である．このスパムの場合，スパマーは対象 Web リソースへリンクしているブックマークエントリを作成する．多くの場合，対象 Web リソースは特定の商品の宣伝や不適切なコンテンツ（暴力的，性的なコンテンツなど）を含む．そのようなスパムを検出するために，対象 Web リソースの URL は非常に偏っているという特徴[5]を利用することができる．その他の研究では，非スパマーは，スパマーがブックマークしているような Web リソース（例えば不正コンテンツを含む，など）はほとんどブックマークしないため，スパマーがブックマークしている Web リソースは，そうでない Web リソースから孤立しているということが報告されている[8].

しかし，上述のアプローチが有効ではない場合がある．一つ目は，対象 Web リソースが通常のコンテンツを含む，すなわち不適切なコンテンツを含まない場合である．スパマーは検索結果のランキングを操作するために，多くのハ

†1 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻
Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba
†2 筑波大学システム情報系情報工学域
Faculty of Engineering, Information and Systems, University of Tsukuba
†3 宇宙航空研究開発機構宇宙科学研究所宇宙科学情報解析研究系
Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency

ⓒ2012 Information Processing Society of Japan

イパーリンクを作成することがある．このような場合，対象 Web リソースは一般ユーザからも多くブックマークされる．その結果，上述のアプローチは有効ではなくなる．もう一つは, スパマーが多くのユーザアカウントを用いて，多くのハイパーリンクを作成する場合である．この場合，対象 Web リソースは多くのリンクを得るため，他の Web リソースと孤立させることが難しくなり，既存のアプローチは十分に機能しない．本稿では，この種のブックマークスパムを,“複数のユーザアカウントを用いた集中的ブックマーク（Intensive Bookmarking using Multiple Accounts; IBMA）”と呼ぶ.

本稿では，IBMA の検出問題に取り組む．我々の知るかぎりでは, この問題に取り組んだ研究はほとんどない. 我々のアプローチでは, SBM アカウントごとに作成されたブックマーク集合の類似度に基づいて SBM アカウントをクラスタリングする．例えば，もし二つの SBM アカウントがほとんど同じ Web リソースもしくは Web サイト集合をブックマークしていれば，それらの SBM アカウントは一つのクラスタにクラスタリングされる．これは，スパマーによって作成，管理されたアカウントであるならば，類似した Web リソース集合をブックマークしているという仮定に基づいている．提案手法では，いくつかの SBM アカウントが（ほとんど）同じ Web リソース集合をブックマークしている場合，それらの SBM アカウントを IBMA スパマーと判定する.

SBM アカウントのクラスタリングにおいて, 二つのブックマークの集合の類似性を測るために，いくつかの類似度を提案する．具体的には，ブックマークエントリに登録されている URL そのものだけでなく，URL から得られるサイト URL も用いて，様々な類似度を定義する.

我々は，SBM の実データセットにおいて，IBMA が検出されるかどうか実験を行った．実験結果は，提案手法が正しく IBMA を検出することを示した.

本稿の構成は以下の通りである. 2 節で SBM とスパマーについて説明する. 3 節で提案手法について説明し, 4 節で評価実験を行う. 5 節では関連研究を紹介し, 6 節で結論を述べる.

2. ソーシャルブックマークサービスにおけるスパマー

この節では，ソーシャルブックマーク（SBM）と SBM におけるスパマーについて説明する

2.1 ソーシャルブックマーク（SBM）

ソーシャルブックマークとは, Web 上で Web リソースに関するブックマークの保存や管理，検索，共有を行うことである．ソーシャルブックマークサービス（SBM service）は，ソーシャルブックマーキング機能を提供する Web サービスの一形態である．2003 年に，J. Schachter らが，最も有名な SBM サービスの一つである“delicious”[a]を立ち上げた．日本では，2005 年に，日本で最も有名な SBM サービスである“はてなブックマーク”[b]が開始された.

SBM サービスのユーザは，一つ以上のその SBM サービスの **SBM ユーザアカウント**を作成する．そのユーザアカウントの一つを使って，ユーザは Web リソースの URL を**ブックマーク**する. ユーザは URL を登録するだけでなく，任意の**コメント**や**タグ**を付与できる．ブックマークを検索したい時，ユーザはキーワードやタグを用いて，SBM 内のブックマークを検索できる．SBM の重要な特徴は，ユーザは自身のブックマークだけでなく，他のユーザアカウントのブックマークも検索することができる．そのため，SBM はより有用で，より興味深いものになる.

2.2 SBM サービスにおけるスパマー

スパマーは，宣伝や SEO などの目的のために，Web リソースへの悪意のあるタグの付与や，悪意のあるブックマークエントリの作成をする．このような悪意ある行動は，それぞれ**スパムタギング**と**スパムブックマーキング**と呼ばれる.

スパムタギングは Web ページに対して不適切なタグを付与することである．一般にスパムタギングは大きく二つに分類される. 一つは, Web ページの価値を下げるために，そのページに悪意あるタグを付与すること，もう一つは，Web ページの価値を最適化するために，そのページに（不適当な）人気または一般的なタグを付与することである．本稿では，スパムタギングの検出は対象外とする.

スパムブックマークも二つに分類される．一つは，スパムコンテンツを含む URL のブックマーク，もう一つは，SEO 目的での特定の商品の宣伝のためのブックマークである．スパマーだけでなく一般ユーザも一つ以上の SBM アカウントを作成し，宣伝などの目的のためにユーザ自身の Web ページをブックマークすることがある. スパマーと非スパマーを区別する明確な根拠がないため，後者の検出の方が明らかに困難である．多くの場合，スパマーはブックマークエントリを作成するために，多くの SBM アカウントを用いる傾向がある．その目的は，より効率的に宣伝することと検索エンジンのスパムフィルターを回避することである．本稿では，この種のブックマークスパムを,“複数のユーザアカウントを用いた集中的ブックマーク（Intensive Bookmarking using Multiple Accounts; IBMA）”と呼び，その検出を試みる.

3. 提案手法

この節では, IBMA の検出手法を提案する. まず, 各 SBM アカウントがブックマークしている URL 集合の類似度に基づいてアカウントをクラスタリングする．その結果，類

a http://delicious.com/
b http://b.hatena.ne.jp/

似した Web リソース集合をブックマークするアカウントがクラスタを形成すると考えられる．クラスタのサイズが十分に大きく，十分に類似していれば，そのクラスタを IBMA を行なっているアカウントのクラスタとして検出する．

まずブックマーク情報を形式化し，データの前処理をする．その後，IBMA の検出手法について詳細な説明をする．

3.1 ブックマーク情報

SBM アカウントは，$u_i \in U\ (1 \le i \le |U|)$ と表現される．ここで，U は SBM アカウントの集合である．また，アカウントu_i のブックマークを$b_{ij} \in B_i\ (i \le j \le |B_i|)$ と表現する．ここで，B_i は u_i のブックマーク集合である．各ブックマーク b_{ij} において，対象 Web リソースの URL は $url(b_{ij})$ と表記する．

3.2 サイト URL の抽出

クラスタリングの前に，各ブックマークb_{ij} について，その URL $url(b_{ij})$ からサイト URL を抽出する．後で説明するクラスタリングは，URL だけでなくサイト URL についても行う．これは，望ましいクラスタリング結果を得るには，URL そのものでは詳細すぎることがあるからである．例えば，ブログについて考えると，IBMA スパマーは異なるブログ記事に対して多くのブックマークエントリを作成するかもしれない．この場合，各ブログ記事の URL よりもサイト URL を用いたほうが適切かもしれない．サイト URL を用いれば，各ブックマークエントリが異なる URL をであっても，サイト URL によってブログそのものを指すブックマークとして扱うことができる．

サイト URL の抽出は，[9]で述べられている簡単な手法を用いる．具体的には，URL をスラッシュ（“ / ”）で分割し，最初の 3 要素を抽出した．例えば，URL http://A/B/C/D/E/ からは，サイト URL http://A/B/C/ が抽出される．ブックマーク b_{ij} のサイト URL を$surl(b_{ij})$ と表現する．

3.3 SBM アカウントのクラスタリング

サイト URL を抽出したので，SBM アカウントのクラスタリングについて議論する．ここでの目的は，類似した Web リソース集合をブックマークしているアカウントのクラスタを形成することである．このために，二つの（サイト）URL 集合間の類似度を定義する必要がある．具体的には，1) URL とサイト URL のどちらを使うか，2) サイト URL の重複をどのように扱うか，に基づいて三つの類似度を提案する．

SBM アカウント u_i における，u_i のブックマーク B_i の URL 集合 $P(u_i)$ を次のように表現する．

$$P(u_i) = \{url(b) \mid b \in B_i\}$$

同様に，サイト URL 集合 $S(u_i)$ を次のように表現する．

$$S(u_i) = \{surl(b) \mid b \in B_i\}$$

二つの SBM アカウント u_i と u_j 間の三つの類似度を定義する．

(1) URL に基づく類似度

これは，u_i と u_j がブックマークしている全 URL 数に対する，共通してブックマークしている URL 数の割合を示している．

$$sim_{page}(u_i, u_j) = \frac{|P(u_i) \cap P(u_j)|}{|P(u_i) \cup P(u_j)|}$$

(2) サイト URL に基づく類似度（重複なし）

これは，u_i と u_j がブックマークしている全サイト URL 数に対する，共通してブックマークしているサイト URL 数の割合を示している．

$$sim_{site}(u_i, u_j) = \frac{|S(u_i) \cap S(u_j)|}{|S(u_i) \cup S(u_j)|}$$

(3) サイト URL に基づく類似度（重複あり）

上述の(2)と異なり，これは各サイト URL の重み，すなわちどれだけの URL が同じサイト URL に集約されるか，を考慮している．

$$sim_{sitedup}(u_i, u_j) = \frac{\sum_{t \in S(u_i) \cap S(u_j)} w(t, u_i, u_j)}{|P(u_i) \cup P(u_j)|}$$

ここで，$w(t, u_i, u_j)$ は u_i と u_j がブックマークしているサイト URL t の重みであり，次のように定義される．

$$w(t, u_i, u_j) = |\{p \mid p \in P(u_i) \cup P(u_j) \wedge surl(p) = t\}|$$

図 1 は，上述の類似度を用いたクラスタリングの例を示している．このグラフは，ノードである SBM アカウント（u_i）と Web リソース（p_i）と，エッジであるブックマークから構成される．また，Web リソース p_1 から p_5 のサイト URL は s_1，Web リソース p_6 から p_8 のサイト URL は s_2，Web リソース p_9 のサイト URL は s_3 である．

$sim_{page}(u_i, u_j)$ を用いることで，共通の Web リソースをブックマークしているような SBM アカウントをクラスタリングする．この例では，SBM アカウント u_1 と u_2 は同一の Web リソース集合をブックマークしているので，これらの SBM アカウントの類似度（$sim_{page}(u_1, u_2)$）は 1.0 である．

$sim_{site}(u_i, u_j)$ を用いることで，ブックマークしている URL 集合が異なっていても，二つの SBM アカウントをクラスタリングすることができる．URL 集合の代わりに，サイト URL 集合が類似しているかどうかを調べる．この例では，SBM アカウント u_3 と u_4 は，ページ URL の観点では類似していない（$sim_{page}(u_3, u_4) = 0.2\ (= 1/5)$）．しかし，$u_3$ と u_4 はサイト URL s_1 だけをブックマークしているので，サイト URL の観点では類似している（$sim_{site}(u_3, u_4) = 1.0$）．

$sim_{site}(u_i, u_j)$ を用いることでサイト URL を利用できるが，あるブックマークにおけるサイト URL の頻度を考慮することはできない．例えば，ある SBM アカウントは，サイト URL s_1 と s_2 の両方をブックマークしているが，ほと

ⓒ2012 Information Processing Society of Japan

んど s_1 をブックマークしているかもしれない．このような傾向を扱うために，$sim_{sitedup}(u_i, u_j)$ を用いる．

図 2 を用いて，$sim_{sitedup}(u_i, u_j)$ の計算方法を示す．図 2 は，図 1 における SBM アカウント u_5 と u_6 と，これらのアカウントのブックマーク，ブックマークしている Web リソース（p_5 から p_9）を抽出したグラフである．

$sim_{sitedup}(u_5, u_6)$ を計算するために，u_5 と u_6 が共通してブックマークしているサイト URL（この例では s_2）の重みを計算する．実際に計算すると，3 個の異なる URL（p_6 から p_8）が集約されるので，サイト URL s_2 の重みは 3 である．その結果，$sim_{sitedup}(u_5, u_6) = 0.6\ (= 3/5)$ である．他の類似度を用いると，$sim_{page}(u_5, u_6) = 0.2\ (= 1/5)$ と $sim_{site}(u_5, u_6) = 0.33 \ldots (= 1/3)$ であり，類似していないことが分かる．

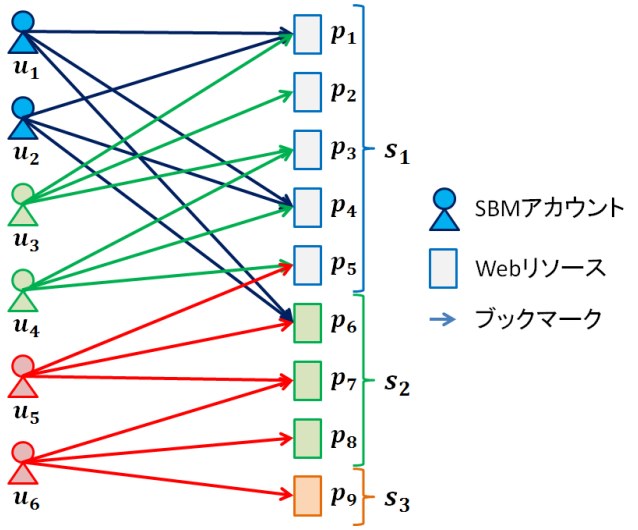


図 1　SBM アカウントのクラスタリング
Fig. 1　Clustering SBM user accounts

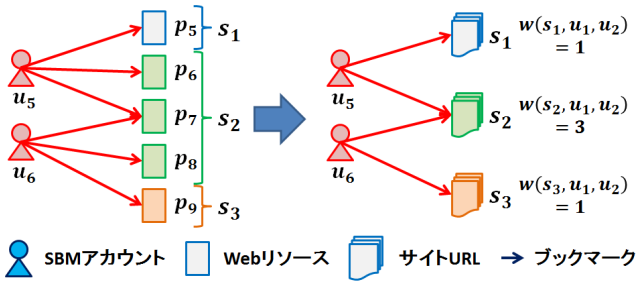


図 2　$sim_{sitedup}(u_i, u_j)$ の計算
Fig. 2　Calculating $sim_{sitedup}(u_i, u_j)$

3.4 IBMA の検出

上述の類似度を用いて，以下の手順で IBMA を検出する．

(1) データの前処理

与えられたブックマークのデータを，今後の処理のために変換する．サイト URL の抽出もここで行う．

(2) SBM アカウントのクラスタリング

上述の類似度の一つを用いてクラスタリングを行い，SBM アカウントのクラスタを得る．クラスタリングアルゴリズムには何を使っても良いが，本稿では，閾値を設定することでクラスタ数を自由に調整できる凝集型階層的クラスタリングを用いた．

(3) IBMA の検出

十分にサイズが大きく，結合力（クラスタ内の類似度）が大きいクラスタを IBMA スパマーのクラスタとみなす．

4. 評価実験

提案手法の精度を示すための実験を行った．まずデータセットの詳細について，次にデータの前処理について説明する．その後，実験結果を示す．

4.1 データセット

本実験では，Livedoor クリップ[10]が提供しているデータセットを用いた．このデータセットには，各ブックマークエントリにおけるユーザアカウント ID，URL，タグ，タイムスタンプが含まれている．ただし，このデータセットには，三つ以上のユーザアカウントからブックマークされている URL だけが含まれている．データセットの詳細を表 1 に示す．

表 1　データセットの詳細
Table 1　The detail of dataset

ユーザアカウント数	55,278
URL 数	410,002
ブックマーク数	3,005,129

4.2 前処理

3.2 節で説明した手法で，データセットからサイト URL を抽出した．410,002 個の URL から，260,502 個のサイト URL を抽出した．

さらに，検出処理における計算を簡単にするために，以下の手法に基づいてアカウントの絞り込みを行う．[5]で報告されているように，スパマーにブックマークされている URL の違いは少ない傾向にある．言い換えると，スパマーにブックマークされている URL は，小さなサイト URL 集合に偏っている傾向がある．このため，必ずしも全てのブックマークを調査しなくても良い．

そのようなブックマークを抽出するために，前の処理で抽出されたサイト URL を利用する．具体的には，各アカウント u_i において，ブックマークしている URL に対する，サイト URL の割合 $R(u_i)$ を計算する．$R(u_i)$ は以下のように定義される．

$$R(u_i) = \frac{|S(u_i)|}{|P(u_i)|}$$

ここで，u_i はユーザアカウント，$P(u_i)$ は u_i がブックマー

クしている URL 集合，$S(u_i)$ は u_i がブックマークしているサイト URL 集合である．$R(u_i)$ の値が定義した閾値以下であるユーザアカウントを抽出した．

以下の実験では，閾値を0.3 に設定し，369 アカウントを抽出した．

4.3 実験方法

以下の手順で実験を行う．4.2 節で抽出したアカウントをクラスタリングする．クラスタリングアルゴリズムとして凝集型階層的クラスタリング，距離関数としてウォード方を用いる．3.3 節で説明した三つの類似度を用いてクラスタリングをする．

クラスタリング結果が得られたら，閾値を0.5 に設定してクラスタを抽出する．クラスタリングには R[11]を用いた．サイズが4 以上のクラスタを，IBMA スパマーのクラスタであるかどうかを判別する．

精度という観点から，提案手法の性能を評価する．SBM アカウントの集合が IBMA を行なっているかどうかの判断は目視によって行う．手順は以下の通りである．

1. クラスタに含まれるユーザアカウントがブックマークしている URL を無作為に50 個抽出する．URL の総数が50 個未満である場合は，すべての URL を抽出する．
2. 抽出した各 URL を訪問し，そのブックマークがサイトの宣伝や商品販売のために作られたものであるかどうかを判別する．
3. ブックマークされている URL にアクセスできない場合は，インターネットアーカイブサービス Wayback Machine[12] を用いる．
4. 存在しない URL が 90% を超える場合，そのクラスタを判別不可能とする．

4.4 実験結果

以下では，実験結果の詳細を説明する．

4.4.1 URL に基づく類似度

提案手法を用いることで生成されたクラスタ数は 214，サイズが 4 以上のクラスタ数は 17 であった．17 クラスタから 106 アカウントが抽出された．

結果を表 2 に示す．クラスタは SBM アカウントのペアごとの類似度の平均（平均類似度）の降順に並んでいる．

表 2 から分かるように，誤検出されたクラスタは存在しなかった．より詳しく調査したところ，次のような IBMA スパムを発見した．

- 単一サイトのページのみをブックマークしている．
- 商品販売目的のページをブックマークしている．
- 特定のブログのページ（記事）を多くブックマークしている．
- コンテンツが異なっているが，デザインが類似しているページを多くブックマークしている．

IBMA スパムの例を図 3 [c]と図 4 [d]に示す．両方の Web ページとも異なる SBM アカウントにブックマークされている．図 3 と図 4 から分かるように，ページのタイトルやコンテンツは異なっているが，デザインが非常に類似している．そのため，これらのブックマークは，特定の商品（例ではマスク）の販売のために作られたものであると判断した．

表 2 結果：URL に基づく類似度 $sim_{page}(u_i, u_j)$
Table 2 Result: similarity based on raw URL $sim_{page}(u_i, u_j)$

クラスタ	アカウント数	平均類似度	判別結果
1	5	1.000	判別不可能
2	9	0.995	○
3	18	0.992	○
4	5	0.985	判別不可能
5	4	0.984	○
6	13	0.977	○
7	4	0.977	○
8	5	0.970	判別不可能
9	7	0.940	○
10	4	0.933	○
11	5	0.888	判別不可能
12	5	0.868	○
13	4	0.853	○
14	4	0.790	○
15	4	0.707	判別不可能
16	5	0.606	判別不可能
17	5	0.599	判別不可能

4.4.2 サイト URL に基づく類似度（重複なし）

生成されたクラスタ数は 207，サイズが 4 以上のクラスタ数は 18 であった．18 クラスタから 121 アカウントが抽出された．結果を表 3 に示す．

この場合，2 クラスタが IBMA スパマーのクラスタとして誤検出された．実際，誤検出されたクラスタの平均類似度は 1.0 であった．これらのクラスタをより詳しく調査したところ，これらのクラスタに含まれる SBM アカウントは，ニュースサイトやウェブマガジンの記事を多くブックマークしていることが分かった．これらの SBM アカウントが，偶然同一のサイトやウェブマガジンの集合をブックマークしていたと考えることは自然である．後で考察するが，ただサイト URL を用いるだけでは，効率的に IBMA

c http://www.winsta09053.com/influenza/influenza_1/0/1/%E3%82%A6%E3%82%A3%E3%83%AB%E3%82%B9%20%E3%83%9E%E3%82%B9%E3%82%AF.html
d http://www.winsta09053.com/influenza/influenza_1/0/1/%E6%8A%97%E4%BD%93%E3%83%9E%E3%82%B9%E3%82%AF.html



図 3　IBMA スパムの例（1）

Fig. 3　IBMA spam example (1)



図 4　IBMA スパムの例（2）

Fig. 4　IBMA spam example (2)

スパマーを検出できない．

4.4.3 サイト URL に基づく類似度（重複あり）

生成されたクラスタ数は 138，サイズが 4 以上のクラスタ数は 25 であった．18 クラスタから 181 アカウントが抽出された．結果を表 4 に示す．

この場合，5 クラスタが IBMA スパマーのクラスタとして誤検出されたが，これらのクラスタは比較的高い平均類似度であった．誤検出された SBM アカウントの特徴は，

表 3　結果：サイト URL に基づく類似度（重複なし）

$sim_{site}(u_i, u_j)$

Table 3　Result: similarity based on site URL w/o duplicate

$sim_{site}(u_i, u_j)$

クラスタ	アカウント数	平均類似度	判別結果
1	12	1.000	○
2	7	1.000	○
3	6	1.000	×
4	5	1.000	×
5	5	1.000	○
6	5	1.000	判別不可能
7	5	1.000	判別不可能
8	4	1.000	判別不可能
9	4	1.000	○
10	4	0.996	○
11	4	0.992	○
12	13	0.992	○
13	9	0.963	○
14	19	0.959	○
15	5	0.933	判別不可能
16	5	0.872	○
17	5	0.830	○
18	4	0.777	判別不可能

4.4.2 節で述べた特徴と同様であった．

4.5 考察

上述の結果をまとめるために，上位 k クラスタの適合率を算出した．その結果を図 5 に示す．図 5 のグラフにおいて，“page”（青），“site”（赤），“sitedup”（緑）はそれぞれ $sim_{page}(u_i, u_j)$，$sim_{site}(u_i, u_j)$，$sim_{sitedup}(u_i, u_j)$ の適合率を表している．グラフから分かるように，適合率において，“page”と“sitedup”は，“site”より高い性能を示している．$k = 13$ とした場合，これらの適合率は1.0 になる．すなわち，k の値を注意深く決定することで，高い性能を達成することができることを示している．一方，“site”は，低い性能を示している．これは，誤検出されたクラスタの平均類似度が高かったため，それらのクラスタが上位に現れてしまったことが原因である．

異なる観点から考察するために，各検出結果で検出された SBM アカウントがどのように重複しているかを調査した．すなわち，一部のアカウントは二つ以上の場合に検出されるかもしれないが，あるアカウントは一つの場合にしか検出されないかもしれない．

その結果をまとめたものを表 5 に示す．誤検出された SBM アカウントを除いて，107 アカウントが IBMA スパムアカウントとして検出された．検出された SBM アカウントのうち，65 アカウントが共通して検出された．一方，2 ア

ⓒ2012 Information Processing Society of Japan

Vol.2012-DBS-154 No.3
Vol.2012-IFAT-107 No.3
2012/8/1

表 4　結果：サイト URL に基づく類似度（重複あり）
$sim_{sitedup}(u_i, u_j)$

Table 4　Result: similarity based on site URL w/duplicates
$sim_{sitedup}(u_i, u_j)$

クラスタ	アカウント数	平均類似度	判別結果
1	12	1.000	○
2	7	1.000	○
3	5	1.000	判別不可能
4	4	1.000	判別不可能
5	4	0.999	○
6	4	0.998	○
7	13	0.998	○
8	5	0.998	判別不可能
9	9	0.997	○
10	18	0.992	○
11	4	0.983	○
12	8	0.978	判別不可能
13	7	0.970	○
14	8	0.967	×
15	4	0.948	○
16	6	0.946	○
17	6	0.943	×
18	9	0.907	×
19	8	0.895	○
20	5	0.893	○
21	4	0.852	○
22	5	0.849	×
23	4	0.847	×
24	18	0.834	判別不可能
25	4	0.728	判別不可能

カウントが$sim_{page}(u_i, u_j)$だけで検出され，12 アカウントが $sim_{sitedup}(u_i, u_j)$ だけで検出された．$sim_{site}(u_i, u_j)$ だけで検出された SBM アカウントは存在しなかった．

$sim_{page}(u_i, u_j)$ だけで検出されたスパムアカウントは，URL は類似していたが，サイト URL は類似していなかった．階層が少ないため，サイト URL をうまく抽出できない URL が存在する．その結果，これらのスパムアカウントは，$sim_{site}(u_i, u_j)$ と $sim_{sitedup}(u_i, u_j)$ では検出することができなかった．

また，$sim_{sitedup}(u_i, u_j)$ だけで検出されたスパムアカウントは，典型的に，長さの異なる URL を持つ Web リソースをブックマークしていた．例えば，URL の末尾の“/index.html”が省略されていた．また，これらのスパムアカウントは，他のサイト URL より特定のサイト URL をより多くブックマークしていた．このような特徴をとらえるためには，$sim_{site}(u_i, u_j)$ ではサイト URL の傾向の違いを考慮することができないため，$sim_{sitedup}(u_i, u_j)$ が有用である．

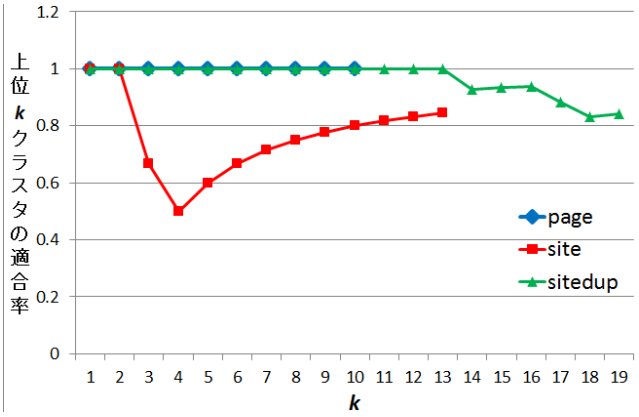


図 5　上位 k クラスタの適合率

Fig. 5　Precision at top k ckusters

表 5　検出された総 IBMA をスパムアカウント数

Table 5　Total number of IBMA spam accounts detected

	sim_{page}	sim_{site}	$sim_{sitedup}$
検出スパムアカウント数	72	87	105
当該類似度でのみ検出されたスパムアカウント数	2	0	12

5. 関連研究

ソーシャルメディアサービスの普及に伴い，これらのサービスにおけるスパム検出は重要になっている．多くの研究者が異なる観点からこの問題に取り組んでいる．

Wetzker ら[5]は，SBM の集中的な統計分析を行い，SBM のブックマークとタグにおける特定のパターンを報告した．また，多くのスパムブックマークが自動的に投稿されており，ドメインの違いが限られていると報告している．

Heymann ら[13]は，ソーシャルメディアサービスにおけるスパマーの問題の概観を述べ，スパム問題に取り組むための三つの手法を提案した．一つ目はスパムコンテンツの検出，二つ目は緩和のためのスパムの降格，三つ目は，ユーザが付与できるタグの数を制限するなどの，ユーザの行動を制限することによるスパムの防止である．

Madkour ら[6]は，ソーシャルブックマークスパマーを判別するための必要条件を考察し，いくつかのクラス分類モデルを提案した．現在考えているユーザとトレーニングセットに含まれるスパマーもしくは非スパマーの共通しているリソースやタグの数を利用した．スパマーの特徴付けのために，タグのブラックリストも用いた．

Koutrika ら[14]は，ユーザが保存されたコンテンツに対して自由にタグを付与できるタギングシステムにおける，スパム検出問題について考察した．コンテンツとタグの正しい対応は既知であるという仮定のもとで，タグのスパム度

ⓒ2012 Information Processing Society of Japan

合いを評価した

SBM スパマーの検出という問題に関して，Krause ら[8]は，教師あり学習手法を提案した．ラベル（スパマーか非スパマー）を含むトレーニングデータセットが与えられ，ユーザがブックマークしているコンテンツやタグをどのくらいスパマーと共有しているかという特徴に基づき，トレーニングデータを用いてクラス分類器を学習させる．

上記の関連研究は，本稿で取り組んだ複数のユーザアカウントによって作成される集中的なブックマークを扱っていない．また，提案手法は，学習を用いないので，多くの場合入手することが困難であるトレーニングデータセットを必要としない．

6. まとめ

本稿では，複数のユーザアカウントを用いた集中的ブックマーク（IBMA）の検出手法を提案した．ブックマークしている URL 集合もしくはサイト URL 集合が類似している SBM アカウントのクラスタリングを行い，同じ（サイト）URL をブックマークしているユーザアカウントをスパムと判定する．この目的のために，三種類の類似度，URL に基づく類似度，サイト URL に基づく類似度（重複なし/重複あり），を提案した．評価実験によって，提案手法は IBMA を正しく検出することを示した．また，提案した類似度は, 異なる種類の IBMA スパマーを検出した. これは，より優れたスパム検出のためには，それらの特徴を組み合わせることが重要であることを示している．

今後の課題として，提案手法の性能を示すために，異なるデータセットに適用する予定である．他の課題として，再現率という観点からの性能評価がある．また，タグなどの他の特徴との組み合わせも興味深い課題である．

謝辞 本研究の実験データは渡邊桂太氏に提供していただいた．ここに記して謝意を示す．

参考文献

1) C. Dugan, M. J. Muller, D. R. Millen, W. Geyer, B. Brownholtz, and M. Moore: The dogear game: a social bookmark recommender system, in Proceedings of International Conference on Supporting Group Work, 2007, pp. 387–390.
2) T. Takahashi, H. Kitagawa, and K. Watanabe: Social bookmarking induced active page ranking, IEICE Transactions, vol. 93-D, pp. 1403–1413, 2010.
3) Y. Yanbe, A. Jatowt, S. Nakamura, and K. Tanaka: Towards improving web search by utilizing social bookmarks, in Proceedings of 7th International Conference on Web Engineering, 2007, pp. 343–357.
4) M. Akram, I. Sohail, S. Hayat, M. I. Shafi, and U. Saeed: Search engine optimization techniques practiced in organizations: A study of four organizations, Computing Research Repository, 2010.
5) R. Wetzker, C. Zimmermann, and C. Bauckhage: Analyzing social bookmarking systems: A del.icio.us cookbook, in Proceedings of the ECAI 2008 Mining Social Data Workshop, 2008, pp. 26–30.
6) A. Madkour, T. Hefni, A. Hefny, and K. S. Refaat: Using semantic features to detect spamming in social bookmarking systems, ECML PKDD Discovery Challenge 2008, 2008.
7) B. Markines, C. Cattuto, and F. Menczer: Social spam detection, in Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web, ser. AIRWeb ’09, 2009, pp. 41–48.
8) B. Krause, C. Schmitz, A. Hotho, and G. Stumme: The anti-social tagger: detecting spam in social bookmarking systems, in Proceedings of the 4th international workshop on Adversarial information retrieval on the web, ser. AIRWeb ’08, 2008, pp. 61–68.
9) H. Saito, M. Toyoda, M. Kitsuregawa, and K. Aihara: large-scale study of link spam detection by graph algorithms, in Proceedings of the 3rd international workshop on Adversarial information retrieval on the web, ser. AIRWeb ’07, 2007, pp. 45–48.
10) Livedoor clip, http://clip.livedoor.com/
11) R, http://www.r-project.org/
12) Wayback machine, http://www.archive.org/web/web.php
13) P. Heymann, G. Koutrika, and H. Garcia-Molina: Fighting spam on social web sites: A survey of approaches and future, IEEE Internet Computing, vol. 6, pp. 36–45, 2007.
14) G. Koutrika, F. A. Effendi, Z. Gy¨ongyi, P. Heymann, and H. Garcia-Molina: Combating spam in tagging systems, in Proceedings of the 3rd International Workshop on Adversarial Information Retrieval on the Web, 2007, pp. 57–64.

ⓒ2012 Information Processing Society of Japan