ソーシャルブックマークにおける時系列分析とグラフ構造解析を用いたスパマー検出

渡邊　桂太† 　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　〒 305–8577 茨城県つくば市天王台 1–1–1
E-mail: †ein_vogel@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 Web ページの情報源として注目を集めている．ソーシャルブックマークは現在までに，多数のユーザを獲得し，Web2.0 を代表するサービスとなったが，同時にスパマーが増加し，問題となっている．スパマーは，宣伝や SEO を目的に特定のページをソーシャルブックマークに登録することで，そのページへのソーシャルブックマークユーザのアクセスの誘引の効果と，Web 検索エンジンにおいて上位の検索結果を得る SEO の効果を得ている．このようなスパマーの増加によって，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，単一のユーザアカウントを利用してスパム行為を行っているシングル ID スパマーと，複数のユーザアカウントを利用してスパム行為を行っているマルチ ID スパマーが生成したソーシャルブックマークのユーザアカウントである，スパムアカウントを検出する手法を提案する．提案手法では，各スパマーのスパムアカウントの統計的特徴に着目し，グラフ構造解析と時系列分析を用いて，ユーザアカウントのスパマー度合いを評価し，高いスパマー度合いを持つユーザアカウントをスパムアカウントとして検出する．評価実験では，ソーシャルブックマークの実データを用いて，シングル ID スパマーとマルチ ID スパマー各々の検出性能を評価し，手法の有効性を検証した．

キーワード　ソーシャルブックマーク，スパマー検出

Detecting Spammers in Social Bookmarking Systems using Time Series and Graph Strucuture Analysis

Keita WATANABE†, 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, 305–8577 Japan
E-mail: †ein_vogel@kde.cs.tsukuba.ac.jp, ††{amagasa,kitagawa}@cs.tsukuba.ac.jp

1. 序　　論

Blog や SNS などのサービスの普及に伴い，Web の利用者は飛躍的に増加し，Web に蓄積されるコンテンツもまた，爆発的な増加を遂げた．情報が溢れかえった昨今の Web では，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強く認識されている．

一方，ソーシャルブックマーク (Social Bookmark: SBM) サービス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 SBM ユーザ（以降，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することによって，新しい Web 検索や情報抽出などを実現した研究が近年盛んである [1], [2].

SBM において，ブックマークを受ける Web ページは，ブックマークしたユーザが有用であると判断したページであると考えられているため，良質なページほど多くのユーザからブックマークを受けている傾向がある．そのため，ページの被ブックマーク数（ページをブックマークしているユーザアカウントの数）は，ページの質を表す指標として注目されている．

しかし, SBM ではブックマーク情報の生成が容易なことから，特定の Web ページの SEO や宣伝等を目的に，SBM を利用する悪意あるユーザ（スパマー）が増加している．これに伴い近年の SBM では，悪質なブックマーク情報が増加し，問題となっている．SBM におけるスパマーは，単一または複数の SBM のユーザアカウントを用いて，宣伝や SEO を行う対象のページをブックマークする．スパマーのブックマーク対象ページは，アフィリエイト広告や商品の販売を誘導するコンテンツを多く含んでいたり，スパマー自身が運営している EC サイトやブログであったりすることが多い．このようなスパマーの増加によって，SBM のブックマーク情報の品質の低下が懸念されている．

本研究では，グラフ構造解析および時系列分析を用いて，スパマーが生成したユーザアカウントである，スパムアカウントを検出する手法を提案する．グラフ構造解析では，スパマーが宣伝の対象としているページの被ブックマーク数が，一般のページに比べて少ないという特徴を用いて，ユーザアカウントのスパマー度合いを評価し，高いスパマー度合いをもつユーザアカウントをスパムアカウントとして検出する．時系列分析では，大量のスパムアカウントを用いているスパマーによって，一般のページと同等以上の被ブックマーク数を得ているスパムページを検出し，このようなスパマーによって生成されたスパムアカウントのスパマー度合いの評価に対応する．

評価実験では，SBM の実データを用いて，単一のスパムアカウントを用いているスパマーと，複数のスパムアカウントを用いているスパマーに対応した，スパムアカウント検出手法の性能評価を行った．いずれの手法も，高い適合率でスパムアカウントの検出できることを示した．また実験結果から，パラメータ設定の変更に伴い，性能が大きく変化することが確認されたため，各パラメータの変化に伴う適合率，再現率の変化の原因について検討する．

2. SBM におけるスパム行為

本研究では，特定の Web ページを不当に宣伝あるいは SEO を目的に SBM を利用するユーザをスパマーとし，彼らの行うスパム行為を以下のスパムブックマーク (Spam Bookmark) とスパムタギング (Spam Tagging) に分類する．

2.1 スパムブックマーク (Spam Bookmark)

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，または宣伝や SEO 目的に特定のサイトをブックマークする行為である．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，ブックマークの対象が，自営サイトのトップから末端のページまで及ぶ場合に，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，本研究ではブックマークスパマーと呼ぶ．

2.2 スパムタギング (Spam Tagging)

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．スパムタギングの例として，「通常のページに対して，悪質なタグを付与する行為」と「SEO 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，本研究ではタグスパマーと呼ぶ．

本研究では，スパマーのスパムアカウントのブックマークの特徴に着目してスパムアカウントの検出を行っているため，スパムタギングを行うタグスパマーのスパムアカウントは，検出の対象外としている．

3. 関 連 研 究

SBM のようなソーシャルメディアサービスにおけるスパムの問題を扱った研究として，Heymann と Koutrika らの研究が挙げられる [3], [4]．Heymann ら [3] 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 [4] 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，タグ付与におけるスパム行為のモデル化を行っている．また，ユーザが付与したタグを基にユーザの信頼度を決定し，スパムの影響を受けにくいタグ検索のランキング手法を提案した．Koutrika らは，タギングシステムのシミュレータを作成し，これを用いて彼女らが提案したタグ検索のランキング手法の有効性を検証した．結果，Koutrika らの手法は，単一のユーザアカウントからのスパム行為に対して有効であることが確認できたが，複数ユーザアカウントによる組織的なスパム行為に対しては，十分な対応がなされないことが確認された．

また，本研究と同じ SBM のスパマーの検出に関する研究として，宗方ら [7] と数原ら [8] の研究が挙げられる．

宗方らは，スパマーの以下のような特徴を示した．

（1） ブックマーク対象ページの被ブックマーク数が 1 の割合が高い．

（2） ブックマーク対象ページに，大量のタグを付与する．

（3） セルフブックマークを行っているスパマーは，ブック

マーク対象ページのドメインが一種類または数種類程度である．

（4） ブックマーク対象ページに付与するタグの数が一定である．

（5） ブックマーク対象ページに付与するコメントが、そのページのタイトルや本文からの引用である．

彼らはこれらの特徴の内，特徴 1, 特徴 2 を定量的に評価する特徴量を定義し，機械学習によってスパムアカウントの判別器を生成した．

数原らは，スパムアカウントと一般のユーザアカウントの間に，一定期間に付与するタグの量と，Chi [6] らが提案した条件付きエントロピーに，特徴的な差異があることを確認した．

前者の特徴は，ユーザアカウント毎に，あるブックマークの時刻と次のブックマークの時刻の差が k 分以下のとき，これら二つのブックマークは一つのセッションで行われたとし，複数のブックマークを一つのセッションとして扱う．このとき，ユーザアカウント毎に得られる複数のセッションを基に，以下の指標を，一般のユーザアカウントとスパムアカウントを区別する特徴量として用いる．

（1） タグ付与件数の最大値，平均，標準偏差

（2） ブックマーク数の最大値，平均，標準偏差

後者は，各ユーザアカウントが使用しているタグのエントロピーと，Chi ら [6] が定義した，タグを選択した際のコンテンツへの到達容易性や，タグ群がどれほどコンテンツをよく記述しているかを表す条件付きエントロピーを一般のユーザアカウントとスパムアカウントを判別する特徴量として用いる．

上述の特徴量を用い，彼らもまた，機械学習によるスパムアカウントの判別器を生成した．

4. 提 案 手 法

本研究では，シングル ID スパマーとマルチ ID スパマー，それぞれに対応したスパムアカウントの検出手法を提案する．本節では，まず初めに，SBM をユーザアカウントとブックマークされる Web ページの関係を SBM グラフという形でモデル化する．次に，シングル ID スパマーの特徴と，シングル ID スパマーがスパムブックマークの対象としているページを識別するために用いる指標 *ibf* について述べる．さらに，*ibf* を基に，ユーザアカウントのスパマー度合いを表す *LSS* の算出方法と，これを基にシングル ID スパマーが生成したスパムアカウントの検出手法について説明する．次に，マルチ ID スパマーに対応したスパマー度合い *aLSS* の算出と，*aLSS* を基にマルチ ID スパマーが生成したスパムアカウントの検出手法について述べる．また最後に，*aLSS* では対応が困難な，大量のスパムアカウントを用いている大規模マルチ ID スパマーのスパムブックマーク対象ページの検出手法について説明し，これを基に大規模マルチ ID スパマーに対応した，スパマー度合いの算出手法 $aLSS^*$ について述べる．

4.1 SBM グラフ

本研究では，SBM の構造を，図 1 のようなユーザアカウントとページをノード，エッジをブックマークの振舞を表す，図 1 のような二部グラフで考える．

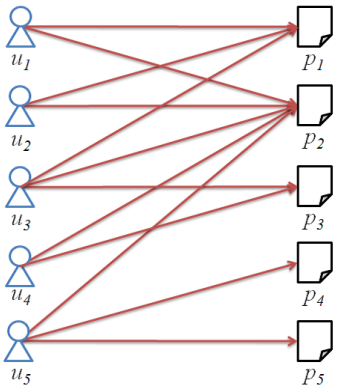

図 1 SBM グラフの例

図 1 は，ユーザアカウント数が 5，ページ数が 5，ブックマーク数が 12 で構成されている．

4.2 シングル ID スパマーの特徴と Inverse Bookmark Frequency(ibf)

シングル ID スパマーの特徴は，宗方ら [7] の研究を基に,「シングル ID スパマーのスパムアカウントがブックマークするページは，高い割合で被ブックマーク数が 1 である」と考える．そこで，シングル ID スパマーのスパムブックマーク対象のページを識別する指標として，*ibf(Inverse Bookmark Frequency)* を定義する．

ibf は，TF-IDF 法における大域的重み *idf(Inverse Document Frequency)* の考えに基づくものである．*idf* 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, 特定の文書のみに偏って出現する語は, 特徴的な語と考え, 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得られる．ページ p の被ブックマーク数を, $R(p)$ としたとき，ページ p の *ibf* 値を，以下のような式で定義する．

$$ibf(p) = \frac{1}{\log_q(R(p) + q - 1)} \tag{1}$$

ただし q は，対数演算の底を表すパラメータである．以降の *ibf* の計算は，$q = 2$ で行うものとする．

4.3 Likelihood of SBM Spammer (LSS) とシングル ID スパマーのスパムアカウント検出

前節で述べた通り，シングル ID スパマーのスパムブックマーク対象ページは，高い割合で被ブックマーク数が少ないため，*ibf* 値が高い傾向にある．そこで本研究では，この特性に着目してユーザアカウントのスパマー度合いをスコアで表現する *Likelihood of SBM Spammer(LSS)* を提案する．$P(u)$ を，ユーザアカウント u がブックマークしているページの集合としたとき，ユーザアカウント u のスパマー度合い *LSS* は，以下のように定義する．

$$LSS(u) = \frac{1}{|P(u)|} \sum_{p \in P(u)} ibf(p) \tag{2}$$

LSS は，ユーザアカウントがブックマークしているページ全ての *ibf* 値を算出し，平均を取ったものである．そのため，*ibf* 値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アカウントの *LSS* は高くなるため，シングル ID スパマーのスパムアカウントの *LSS* もまた高くなることが考えられる．

LSS を用いたスパムアカウントの検出は，以下の手順で行う．

（1） 全ページの *ibf* を算出．

（2） (1) を基に，全ユーザアカウントの *LSS* を算出．

（3） 閾値よりも高い *LSS* を持つユーザアカウントを，スパムアカウントとして検出．

LSS の算出例を，図 2 に示す．この例では，p_4,p_5 をスパムブックマークの対象ページ，u_5 をシングル ID スパマーのスパムアカウントとしている．p_4,p_5 をブックマークしているユーザアカウントは，スパムアカウントである u_5 のみのため，p_4,p_5 の *ibf* 値は最大の 1 となっている．他のページは，複数のユーザアカウントからブックマークを受けているため，*ibf* 値は低くなる．この *ibf* 値を基に，各ユーザアカウントの *LSS* を算出した結果，u_5 に最も高い *LSS* が与えられた．今回は，この *LSS* に閾値を設定し，閾値以上の *LSS* を持つユーザアカウントをスパムアカウントとして検出する．

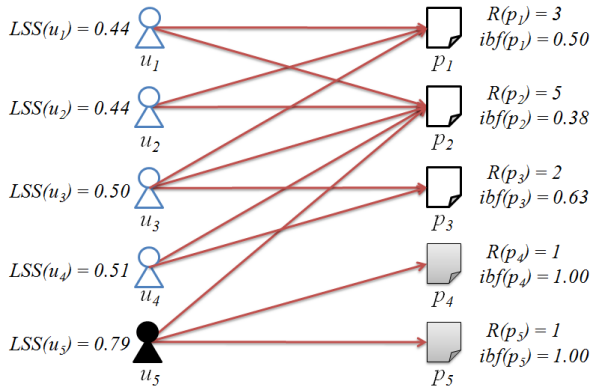

図 2　*LSS* の算出例

4.4　マルチ ID スパマーへの対応

4.3 節で提案した *LSS* では，マルチ ID スパマーへの対応は困難である．マルチ ID スパマーは，複数のスパムアカウントを用いて組織的にスパムブックマークを行うため，スパムブックマークの対象となるページは，被ブックマーク数は低くならず，高い *ibf* 値が得られない．そのため，図 3 のように，スパムページ p_4,p_5 をブックマークしているスパムアカウント u_5,u_6,u_7 は高い *LSS* が得られず，検出が困難となる．

マルチ ID スパマーが生成した複数のスパムアカウントは，「ブックマーク対象ページが高い割合で共通している」ということが確認されており，本研究ではこの特徴に着目し，マルチ ID スパマーに対応したスパムアカウントの検出を図る．まず，SBM グラフを用いて，ブックマーク対象が高い割合で共通しているユーザアカウントをクラスタリングし，ここで得られたクラスタを基に，組織的なスパムブックマークによって不当に増やされた被ブックマーク数に補正をかける．これを基に，マルチ ID スパマーが生成したスパムアカウントに対応したユーザアカウントのスパマー度合い *aLSS* を定義し，*aLSS* を用いてマルチ ID スパマーに対応したスパムアカウントの検出を実現する．

4.4.1　ユーザアカウントのクラスタリング

まず初めに，ブックマーク対象のページが高い割合で共通しているユーザアカウントのクラスタリングを行う．行をユーザアカウント，列をページとし，成分には対応する列のページの *ibf* 値を与えた隣接行列で SBM グラフを表現する（図 4）．隣接行列の成分に *ibf* 値を導入することによって，被ブックマーク数の低いページを共通にブックマークしているユーザアカウント同士が，同じクラスタに併合されやすくなる効果が得られる．

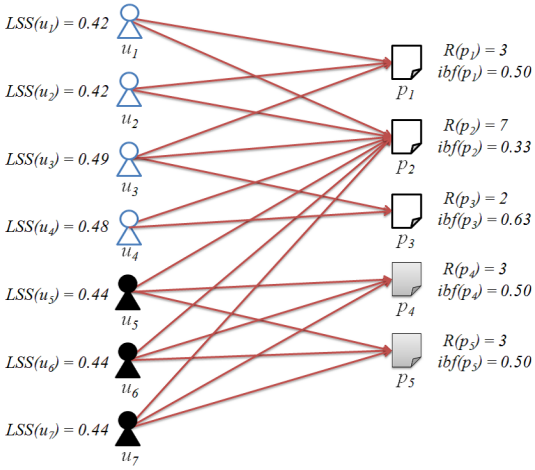

図 3　マルチ ID スパマーの例

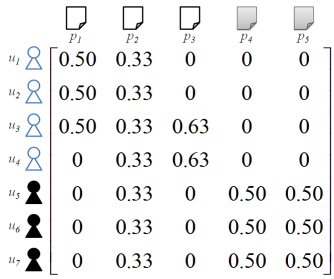

図 4　SBM グラフの隣接行列

4.4.2　組織的スパムブックマークの補正

本節では，複数のスパムアカウントを用いて組織的に増やされた被ブックマーク数の補正を行う．4.4.1 節で得られたクラスタを一つ選択し，選択したクラスタを一つの仮想ユーザアカウントとみなし，図 5 の手順で不正に増加させられた被ブックマーク数を補正する．

4.4.3　aLSS(adjusted LSS) とマルチ ID スパマーに対応したスパマー検出手法

4.4.2 節で得られた被ブックマーク数を基に，マルチ ID スパマーに対応したユーザアカウントのスパマー度合い，*aLSS(adjusted LSS)* の算出を行う．*aLSS* は，*LSS* と同様にユーザアカウントのブックマーク対象ページの *ibf* 値の平均であるが，ここで用いられる *ibf* は 4.4.2 節で補正をかけた被

ブックマーク数を基に算出される．

aLSS の算出とスパムアカウントの検出手順は，以下の通りである．

（1） ユーザアカウントをクラスタリング（4.4.1 節）．

（2） クラスタを一つ選択．

（3） （2）の選択クラスタに含まれるユーザアカウントのブックマーク対象ページの被ブックマーク数を補正．(4.4.2 節)

（4） （2）の選択クラスタに含まれるユーザアカウントのブックマーク対象ページの *ibf* 値を算出．

（5） （4）の *ibf* 値を基に，選択クラスタに含まれるユーザアカウントの *aLSS* を算出．

（6） 閾値以上の *aLSS* を持つユーザアカウントをスパムアカウントとして検出．

（7） （3）で補正した被ブックマーク数を元に戻し，(2）の手順に戻る．

図 3 の SBM グラフに，上述の検出手順を適用した例を以下に示す．

図 4 の隣接行列から，以下の 3 つのクラスタが得られたとする．

$$cluster_1 = \{u_1, u_2\}$$

$$cluster_2 = \{u_3, u_4\}$$

$$cluster_3 = \{u_5, u_6, u_7\}$$

最初に $cluster_1$ で行われた組織的ブックマークを，図 6 のように補正し，ここで得られた被ブックマーク数を基に，$cluster_1$ に含まれている u_1, u_2 の *aLSS* を以下のように算出する．

$$aLSS(u_1) = \frac{0.63 + 0.35}{2} \fallingdotseq 0.42$$

$$aLSS(u_2) = \frac{0.63 + 0.35}{2} \fallingdotseq 0.42$$

同様に，$cluster_2$ の組織的ブックマークを図 7 のように補正し，$cluster_2$ に含まれている u_3, u_4 の *aLSS* を算出する．

$$aLSS(u_3) = \frac{0.50 + 0.35 + 1.00}{3} \fallingdotseq 0.62$$

$$aLSS(u_4) = \frac{0.35 + 1.00}{2} \fallingdotseq 0.68$$

今回の例では，$cluster_1$ と $cluster_2$ は，それぞれ一般のユーザアカウントをグルーピングしたクラスタとしている．一般のユーザアカウントのブックマーク対象ページは，他のクラスタに含まれるユーザアカウントからもブックマークを受けているため，4.4.2 節の被ブックマーク数の補正を行っても *aLSS* は高くならないと考えられる．

次に $cluster_3$ の組織的ブックマークを図 8 のように補正し，$cluster_3$ に含まれている u_5, u_6, u_7 の *aLSS* を算出する．

$$aLSS(u_5) = \frac{0.39 + 1.00 + 1.00}{3} \fallingdotseq 0.80$$

$$aLSS(u_6) = \frac{0.39 + 1.00 + 1.00}{3} \fallingdotseq 0.80$$

$$aLSS(u_7) = \frac{0.39 + 1.00 + 1.00}{3} \fallingdotseq 0.80$$

この例では，$cluster_3$ に含まれるユーザアカウントは，マルチ ID スパマーによって生成されたスパムアカウントとしている．$cluster_3$ に含まれるユーザアカウントは，p_4, p_5 を共通にブックマークしているが，これらは他のクラスタからブックマークの対象とされていない．そのため，図 8 のように補正された際，p_4, p_5 の被ブックマーク数は 1 となる．結果，u_5, u_6, u_7 の *aLSS* は,0.8 と高いスコアが算出される．

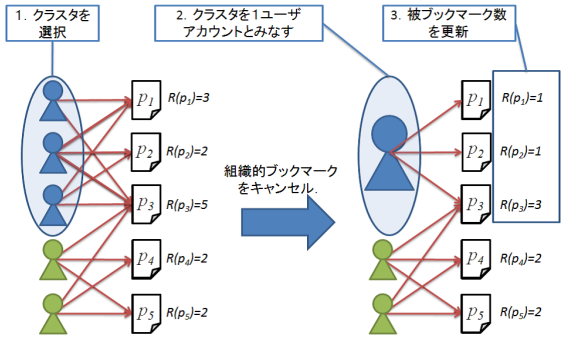

図 5　組織的ブックマークの補正手順

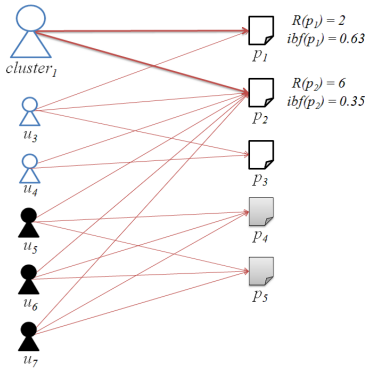

図 6　マルチ ID スパマーのスパムアカウント検出例：$cluster_1$.

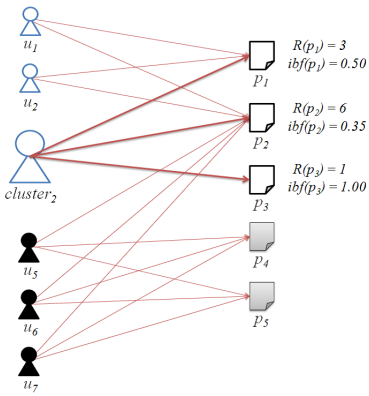

図 7　マルチ ID スパマーのスパムアカウント検出例：$cluster_2$.

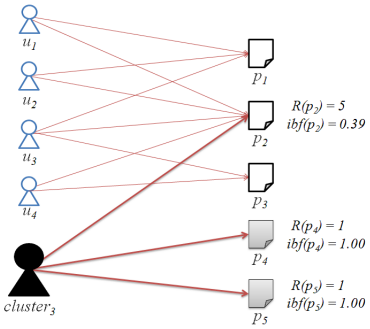

図 8　マルチ ID スパマーのスパムアカウント検出例：$cluster_3$.

4.5　時系列分析による大規模マルチ ID スパマーへの対応

4.4.3 節で述べた手法では，ユーザアカウント間で共通しているブックマーク対象のページの数と *ibf* 値によって，形成されるクラスタが決定され，被ブックマーク数の補正が行われる．そのため，大量のスパムアカウントを用いて，スパムブックマークの対象ページの被ブックマーク数を一般のページと同等以上にしているような，マルチ ID スパマーへの対応は困難である．以降，このように大量のスパムアカウントを利用してスパムブックマークを行っているマルチ ID スパマーを，大規模マルチ ID スパマーと呼ぶ．

前述の通り，大規模マルチ ID スパマーのスパムブックマーク対象ページは，被ブックマーク数が一般のページと同等以上となるため，*ibf* 値も一般のページと同等になってしまう．そのため，*ibf* では，一般のページとスパムブックマーク対象のページの識別が困難となり，4.4.1 節のクラスタリングの方法では，十分なユーザアカウントのグルーピングが行えない．その結果として，4.4.2 節で被ブックマーク数の補正が正常に行われず，大規模マルチ ID スパマーのスパムアカウントの $aLSS$ が一般のユーザアカウントとほとんど変わらなくなってしまうため，検出が困難となってしまう．

本節では，このような大規模マルチ ID スパマーのスパムブックマーク対象ページを時系列分析によって検出し，これを基に大規模マルチ ID スパマーの検出に対応したユーザアカウントのスパマー度合い，$aLSS^*$ の算出手法を提案する．

大規模マルチ ID スパマーのスパムブックマーク対象のページの多くは，SBM サービスの人気エントリへの掲載を図っていることが多いため，短い時間で大量のユーザアカウントからブックマークを受けている傾向がある．そこで本手法では，ページがブックマークを受ける時間間隔に着目し，短い間隔で大量にブックマークを受けているページを，大規模マルチ ID スパマーのスパムブックマーク対象ページとして検出する．ここで検出したページの *ibf* 値を，被ブックマーク数に関係なく最大値である 1 とすることで，大規模マルチ ID スパマーに対応したユーザアカウントのスパマー度合い $aLSS^*$ を実現する．

4.5.1　大規模マルチ ID スパマーのスパムブックマーク対象ページの特徴分析

多くの SBM サービスには人気エントリという，最近に多くのユーザアカウントからブックマークを受けたページをリストアップして，SBM サービスのトップページ等に掲載するような機能がある．これに掲載されるページは，多くのユーザの閲覧を誘導できるため，高い宣伝効果が得られる．そのため，多くの大規模マルチ ID スパマーは，大量のスパムアカウントを用いて，スパムブックマーク対象ページが人気エントリに掲載されることを図っていると考えられる．

大規模マルチ ID スパマーのブックマーク対象ページは，ブックマークを受ける時間間隔（以降，ブックマーク間隔と呼ぶ）が短いことが多いと考えられ，他の一般のページに比べ，ブックマーク間隔が高い割合で短いページは，大規模マルチ ID スパマーによってスパムブックマークを受けたページであると考えられる．そこで本節では，ページのブックマーク間隔から得られる標準偏差を基に上述のようなページを識別する特徴の分析を行った．

被ブックマーク数が 30 以上のページをランダムに 100 ページ抽出し，目視で一般のページとスパムページにラベル付けを行い，これらの間で標準偏差に差異があるか確認した．各ページのブックマーク間隔を昇順にしたときに得られる上位 n %のブックマーク間隔から，標準偏差を算出した結果を図 9 に示す．

図 9 より，スパムページのブックマーク間隔から得られる標準偏差は，低い値に集中していることが確認できた．

そこで本研究では，n とブックマーク間隔の標準偏差の閾値 σ_{std} を設定した際，あるページの昇順のブックマーク間隔の上位 n %で得られる標準偏差が，閾値 σ_{std} 以下の場合は，そのページをスパムページと判定する．以上の方法で，スパムページと判定したページの *ibf* 値を 1 とすることで，大規模マルチ ID スパマーに対応したスパマー度合い $aLSS^*$ の算出を実現する．

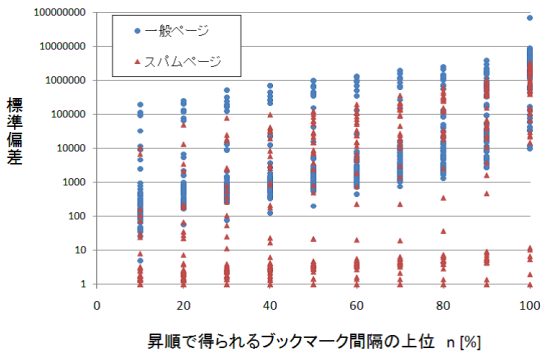

図 9　被ブックマーク数が 30 以上のページのブックマーク間隔から得られる一般ページとスパムページの標準偏差の分布

5.　評 価 実 験

本節では，シングル ID スパマーに対応した LSS を用いたスパムアカウント検出手法の性能と，マルチ，大規模マルチ ID スパマーに対応した $aLSS$, $aLSS^*$ を用いたスパムアカウント検出手法の性能評価を行い，その有効性を検証する．

5.1　*LSS* によるスパムアカウント検出実験

5.1.1　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(注1)の SBM データを用いて 4.3 節のシングル ID スパマーに対応したスパムアカウント検出手法の性能を評価する．実験データの詳細は以下の通りである（表 1）．

表 1　*LSS* によるスパムアカウント検出実験データ

ユーザアカウント数	3,735
ページ数	1,265,208
ブックマーク（エッジ）数	3,991,955

5.1.2　*LSS* によるスパムアカウント検出性能の評価

LSS を用いたスパムアカウントの検出性能の評価は，以下の手順で行う．

（1）　実験データからランダムに 100 ユーザアカウントを抽出．

（2）　抽出したユーザアカウントの *LSS* を算出．

（3）　*LSS* の閾値 σ を設定し，$LSS(u) \geqq \sigma$ となるユーザアカウント u をスパムアカウントとして検出．

（4）　検出したユーザアカウントがスパムアカウントかどうかを目視で判別し，σ を変化させたときに得られる適合率，再現率を算出する．

以上の実験を 4 回試行し，各試行で得られた適合率と再現率の平均を図 10 に示す．結果，*LSS* の閾値 σ の増加に伴い，再現

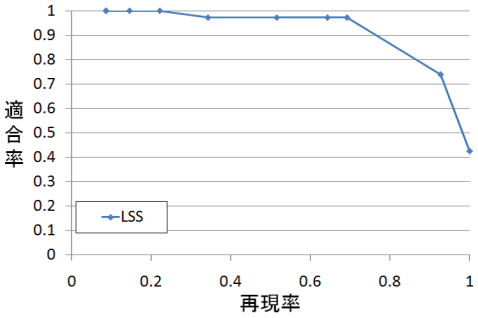

図 10　*LSS* のスパムアカウント検出性能．

率が減少し，適合率が高くなる傾向が確認できた．再現率の減少の原因として，一部のスパマーが生成したスパムアカウントがスパムブックマーク対象のページだけでなく，一般のページもブックマークしているため，そのスパムアカウントの *LSS* が低くなってしまうことが挙げられる．また，スパムブックマーク対象ページの被ブックマーク数を高くしているマルチ ID スパマーのスパムアカウントは，*LSS* が低くなっているため，こちらも再現率減少の原因として挙げられる．

LSS では，被ブックマーク数の少ないページを多くブックマークしているユーザアカウントをスパムアカウントと捉えているため，本手法では，後者のような組織的なスパムブックマークによってブックマーク対象のスパムページに高い被ブックマーク数を与えているマルチ ID スパマーに対応したスパムアカウント検出は困難であることが確認できた．

(注1)：http;//b.hatena.ne,jp/

5.2　*aLSS* と *aLSS によるスパムアカウント検出実験**

本節では，マルチ，大規模マルチ ID スパマーが生成したスパムアカウントを対象に，4.4.3 節と 4.5.1 節で示した *aLSS*，*aLSS** を用いたスパムアカウント検出手法の性能評価を行う．

5.2.1　実験データの概要

本実験では，livedoor クリップ(注2)から提供されている研究用データセットを用いて，スパムアカウントの検出性能を評価した．実験データは，2010 年 6 月に公開されたものを用い，詳細を表 2 に示す．

表 2　マルチ ID スパマー検出の実験データ

ユーザ数	55,274
ページ数（被ブックマーク数が 3 以上）	410,048
ブックマーク（エッジ）数	3,005,069

本実験データは，被ブックマーク数が 3 未満のページは含まない仕様となっているため，シングル ID スパマーは含まれていない．

5.2.2　*aLSS* と *aLSS** によるスパムアカウント検出性能の評価

aLSS と *aLSS** によるスパムアカウント検出性能の評価は，以下の手順で行う．

（1）　生成するクラスタ数 k を設定し，k 個のユーザアカウントのクラスタを抽出．

（2）　(1) で抽出したクラスタを基に，ユーザアカウントの *aLSS* と *aLSS** を算出．

（3）　ランダムに 100 ユーザアカウントを選択．

（4）　(3) で得られたユーザアカウントが，それぞれスパムアカウントかどうかを目視で判別し，σ を変化させたときに，このユーザアカウント集合から得られる適合率，再現率を算出．

また，4.4.1 節で用いるクラスタリング手法は，クラスタリングツール bayon(注3)による bisecting k-means を用いた．

以上の実験を，クラスタ数 $k = 10$〜1000 でそれぞれ 4 回試行し，各試行で得られた適合率と再現率の平均を図 11〜16 に示す．また，今回の実験では，*aLSS** のパラメータは，$n = 20, \sigma_{std} = 10$ と設定した．

図 11〜14 より，*aLSS*，*aLSS** ともに閾値 σ の増加に伴い，再現率が減少し，適合率が高くなる傾向が確認できた．この再現率の減少は *LSS* と同様に，一部のスパマーが生成したユーザアカウントが，スパムブックマーク対象のページだけでなく，一般のページもブックマークしていることが原因であると考えられる．

図 15,16 に，クラスタ数 k の変化に伴う $\sigma = 0.6$ で得られる *aLSS*，*aLSS** の性能を示す．これらから，同じ適合率で *aLSS** は，*aLSS* より高い再現率が得られることが確認できた．また，k を高くするにつれて，*aLSS*，*aLSS** ともに適合率は上昇，再現率が減少することが確認できた．この原因として，クラスタ数が小さければ，規模の大きいマルチ ID スパマー

(注2)：http://clip.livedoor.com/

(注3)：http://code.google.com/p/bayon/

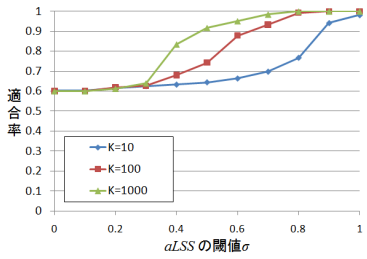

図 11　閾値 σ の変化に伴う $aLSS$ の適合率

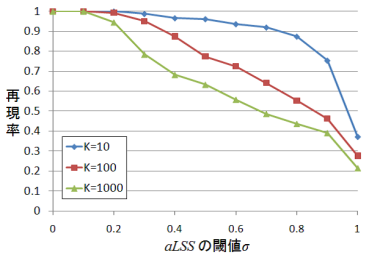

図 12　閾値 σ の変化に伴う $aLSS$ の再現率

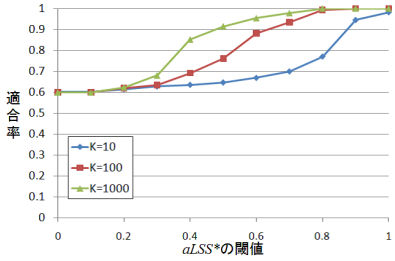

図 13　閾値 σ の変化に伴う $aLSS^*$ の適合率

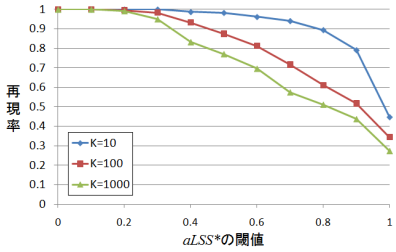

図 14　閾値 σ の変化に伴う $aLSS^*$ の再現率

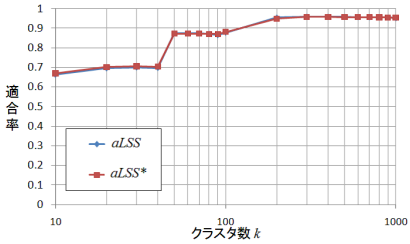

図 15　閾値 $\sigma = 0.6$ のときのクラスタ数 k の変化に伴う $aLSS$, $aLSS^*$ の適合率

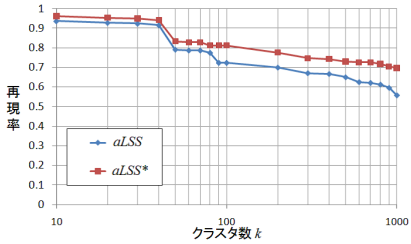

図 16　閾値 $\sigma = 0.6$ のときのクラスタ数 k の変化に伴う $aLSS$, $aLSS^*$ の再現率

に対応できるが，一般のユーザアカウントも含めて検出してしまうことが考えられる．そのため，誤検出が増えるが高い再現率が得られるといった結果が得られた．逆にクラスタ数が多ければ，規模の小さいマルチ ID スパマーに対応できるが，規模の大きいマルチ ID スパマーへの対応が不十分となることが考えられる．そのため，誤検出は減るが再現率は減少してしまうといった結果が得られた．

6.　まとめと今後の課題

本研究では，SBM サービスを対象にグラフ構造解析と時系列分析を用いて，スパムアカウントを検出する手法を提案した．SBM のユーザアカウントがブックマークしているページの被ブックマーク数を基に，そのユーザアカウントのスパマーらしさを評価する指標 LSS を定義し，閾値以上の LSS を持つユーザアカウントを，シングル ID スパマーが生成したスパムアカウントとして検出する手法を提案した．また，複数のスパムアカウントを用いてスパムブックマーク対象の被ブックマーク数を操作するマルチ ID スパマーに対応したスパマー度合いの算出手法の $aLSS$ と，大量のユーザアカウントを利用している大規模マルチ ID スパマーに対応したスパマー度合いの算出手法の $aLSS^*$ を提案した．さらに，評価実験より，いずれのスパマーのスパムアカウントも，パラメータの設定によって高い適合率で検出が可能であることを示した．

$aLSS^*$ は，$aLSS$ より高い検出性能が得られることを示したが，設定するパラメータが増えてしまったため，この手法についてはより深い考察を行い，設定パラメータの削減を行うことが，今後の課題として挙げられる．

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] Y. Yanbe, A. Jatowt, S. Nakamura and K. Tanaka, "Towards Improving Web Search by Utilizing Social Bookmarks", In Proc. of 7th International Conference on Web Engineering, pp.343-357, 2007.

[2] 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.

[3] 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.

[4] 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, ACM, Banff, Alberta, CANADA, 2007.

[5] Beate Krause, Christoph Schmitz, Andreas Hotho, and Gerd 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*, pages 61–68, ACM, Beijing, CHINA, 2008.

[6] Ed H. Chi and Todd Mytkowicz. Understanding the efficiency of social tagging systems using information theory. In Proceedings of the nineteenth ACM conference on Hypertext and hypermedia, pp. 81–88, 2008.

[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.