

## Web 上での「叫ぶ」範囲選択方式の検討

山口亮輔<sup>†</sup> 金井敦<sup>†</sup>

近年,現実世界においてインフォーマル・コミュニケーションと呼ばれる,計画的でなく,カテゴリーや話題性が存在しない,偶発的なコミュニケーションが注目を集めている.現在 Web 世界においてもインフォーマル・コミュニケーションを実現するためにいくつかのサービスが提案されている.しかし,既存のサービスでは 1 つの Web ページに集まっているユーザ間を想定しており,意味的な範囲の広さが考慮されていない.そこで,本稿ではコミュニケーションを行う範囲を「叫ぶ」範囲と呼び,その範囲を定義し,その実現方式を検討する.実現方式としては,ハイパーリンク構造,URL のディレクトリ階層,ユーザの利用履歴を組み合わせた手法を提案する.

### A study on a methodology to select a “Shout” range on the Web

Ryosuke Yamaguchi<sup>†</sup> and Atsushi Kanai<sup>†</sup>

Recently, the importance of unscheduled, unarranged, accidental communication (Informal communication) has been recognized in real world. On the other hand, some service is suggested to realize informal communication on the Web. These existing services provide communication range based on only one page. However, it is necessary to communicate in more semantically wide range. In this paper, the range of web pages available to communicate is investigated, and range (We call the range “Shout” range) is defined. And the methodology to detect web page in the “Shout” range are proposed.

## 1. はじめに

近年,情報技術の発展に伴い Web を利用した様々なコミュニケーション形態が普及している.その背景には Web 上での情報の共有を目的としたサービスが急激な成長を遂げたことが関係している.一方,現実世界において計画的でなく,カテゴリーや話題が存在しない偶発的なコミュニケーションが注目を集めている.これはインフォーマル・コミュニケーションと呼ばれ[1],偶然の出会いから生まれるコミュニケーションである.現状の Web 世界ではこのインフォーマル・コミュニケーションは起こりえない.その理由として,現在の Web では誰かとコミュニケーションしたいときに,自分の意思でコミュニケーションできる機能をもった Web ページを訪れなければ,人と人が偶然に出会うことも,互いを認識することもできないからと考えられる.そこで,Web 上でインフォーマル・コミュニケーションを実現する方式として,今現在同一ページを閲覧している他者をインフォーマル・コミュニティとし,コミュニケーションすることができる wegru[2]や SPEEKY [3],Nesutu[4]が提案されている[5][6].

現在の wegru や SPEEKY,Nesutu では 1 つの Web ページに集まっているユーザを想定しており,ホームページの構成や内容によって異なる適切なコミュニケーションの範囲が考慮されておらず,インフォーマル・コミュニティの範囲に限られる.そこで,ある Web ページを中心とした Web ページグループを構成し,グループ内の他者を認識して,コミュニケーションを行うことを可能にする方式を検討する.

## 2. コミュニケーション範囲

現状の wegru には,ある 1 つの Web ページを閲覧しているユーザ同士が互いを認識でき,コミュニケーションできる機能が提案されている.しかし,Web ページのコンテンツによっては,現在閲覧している Web ページのみではなく,他の Web ページも考慮し,インフォーマル・コミュニティを形成することで,幅広いインフォーマル・コミュニティが形成できると考える.そこで,図 1 のように Web ページの適切なコミュニケーション範囲を想定する必要がある.

Web ページの範囲を考える上で現実世界の様子を考察すると,現実世界におけるコミュニケーションの範囲は人との距離や声量といった,物理的な制約が考えられる.より多く,より遠くにいる人とインフォーマル・コミュニティを形成するために大きな声を出す,密集するといった手法を取ることができる.それはコミュニケーションの範囲を自分で制限できるということである.

<sup>†</sup> 法政大学  
Hosei University

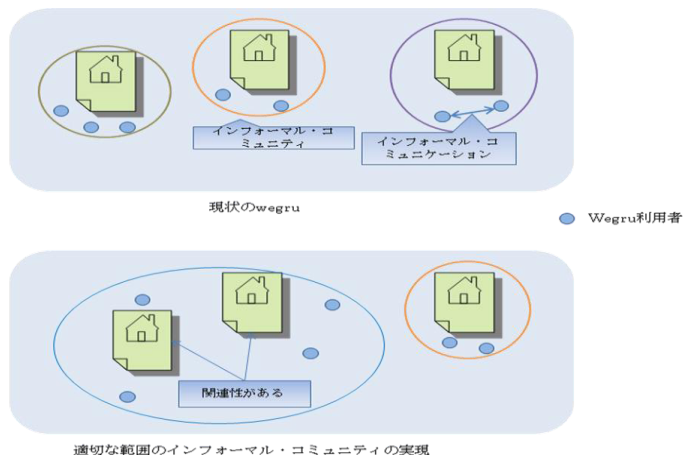


図 1 現在のインフォーマル・コミュニティの範囲と適切な範囲の提案

一方,Web 上では物理的制約がないのでコミュニケーションの範囲が規定できない.そこで,コミュニケーションの範囲を制限するためには Web ページ間の距離を定義する必要がある.Web 上での距離を基準としてコミュニケーションの範囲をコントロールし,ホームページをグループ化して,グループ内のユーザにコミュニケーションを行う「叫ぶ」機能を提案する.

### 3. 距離の規定

#### 3.1 距離の選定基準

「叫ぶ」機能を定義するにあたり,現実世界の物理的制約としての空間の代替として,Web ページ間の内容や構成の関連性を考慮した距離(空間距離とする)と,物理的制約のない仮想世界の特徴を利用して,過去にその Web ページを閲覧したユーザへの時間的な距離(時間距離とする)の 2 種類の選定基準を挙げる.

前者は,現在の Web ページのコンテンツに近い内容の Web ページを閲覧しているユーザとインフォーマル・コミュニティを形成することで,幅広い情報を得ることを目的としており,後者は,Web ページのコンテンツを過去に閲覧していたユーザとインフォーマル・コミュニティを形成することで,そのコンテンツについて人を介して深い情報を得ることを目的としている.

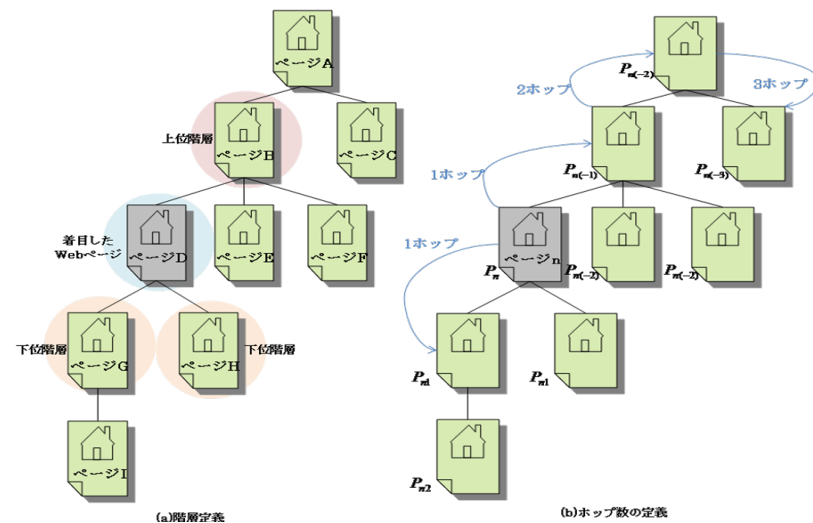


図 2 空間距離を基準としたホームページの階層とホップの定義

#### 3.2 空間距離を基準とした範囲

現在閲覧している Web ページの内容や,構造をもとに Web ページ間の距離を定義し,最も近い Web ページから順に「叫ぶ」範囲を調整する,Web ページをベースにしたインフォーマル・コミュニティの形成である.今回は,内容の類似性を把握することができないため,ホームページの構造をもとに距離を定義する.

現在,多くのホームページが Web 上に存在しているが,多くの Web デザインマニュアルにおいて木構造が推奨されており,木構造に基づき研究されている[7]. よって今回はホームページがトップページからコンテンツごとに徐々に詳細化された Web ページに枝分かれする構造を持つと仮定し,これを意味的な木構造と定義する.その構造の中の階層の定義を行い,「叫ぶ」範囲を規定する.ホームページ内のある Web ページ  $n$  をユーザが閲覧していると, Web ページ  $n$  におけるコンテンツが詳細化された Web ページ  $n'$  を「下位階層」とする.例として図 2(a)のページ D を Web ページ  $n$  とすると,ページ G とページ H が Web ページ  $n'$  にあたり,ページ G,ページ H はページ D の下位階層となる.逆に Web ページ  $n'$  から見た Web ページ  $n$  を「上位階層」と定義する.図 2(a)を例にするとページ B はページ D の上位階層となる.

価格.com[8]を例に挙げると「ノート PC」のコンテンツが表示される Web ページをページ D とした場合,ページ B は「PC」全体のコンテンツが表示される Web ページと

なり、ページ E やページ F には「UMPC」や「デスクトップ PC」のコンテンツが表示される Web ページが考えられる。

現在の Web ページを中心に階層を 1 つ移り変わることを 1 ホップと定義したとき、「叫ぶ」範囲の大きさは、現在閲覧しているページ  $P_n$  からのホップ数を基準とする。図

2(b)に示すようにページ  $P_n$  より上位階層方向に  $i$  だけホップした Web ページを  $P_{n(-i)}$

とし、下位階層方向に  $i$  だけホップした Web ページを  $P_{ni}$  とする。「叫ぶ」範囲の広さを  $i$  で表す。

「叫ぶ」範囲の限界は、下位階層ならばある限りにホップ数を増やしても問題がないが、上位階層を制限なくホップしていくと、現在見ている Web ページのコンテンツとは関連性が徐々に薄くなっていくため、ここでは上位階層方向には 1 ホップを限界とする。 $i$  をマイナス方向に増やすと、意味的に広い範囲でインフォーマル・コミュニティを形成でき、 $i$  をプラス方向に増やすと意味的には変わらないが、多数のユーザとインフォーマル・コミュニティを形成できる。

### 3.3 時間距離を基準とした範囲

自分が現在閲覧している Web ページを過去に訪れたことがあり、現在 wegru を使用しているユーザを「叫ぶ」範囲とし、ユーザが Web ページを閲覧した時刻を過去に向かって遡るかたちで、距離を調整していく、ユーザをベースとしたインフォーマル・コミュニティの形成である。

図 3 に示すように、ある Web ページ A を訪れたユーザが最後に閲覧した時刻を記録し、現在を基準として、現在と最後に閲覧した時刻の差を距離と定義して、その距離の幅で「叫ぶ」範囲を調整する。「叫ぶ」範囲を現在から、一日以内、一週間以内、一ヶ月以内と過去に向かって広げていく。

## 4. 実際のホームページの構造

本稿では、時間距離を基準とした範囲の規定は取り扱わず、空間距離を基準とした範囲の規定について検討する。

3.2 節でホームページは木構造で構成されていると仮定したが、実際のホームページを検証してみると、意味的には木構造として構成されているが、つくりとして完璧な木構造で構成されているホームページは多くない。スタティックなホームページであっても、図 4 に示すように、ある Web ページから様々な Web ページに向かってハイパーリンクが貼られており、その中から下位階層を判断するのは難しい。

また、スタティックなホームページであれば下位階層は不変であるが、ダイナミック

なホームページである場合、ユーザの履歴を目安にリコメンドが形成されるケースや、ダイナミックに広告が形成されるケースがある。

しかし、ダイナミックに構造が変化するホームページのリコメンドや広告は、その Web ページの同位階層や別の階層であり、今回は 3.2 節で定義した通り、同位階層や別の階層は上位階層方向に 2 ホップ以上するため、下位階層と判断しない。

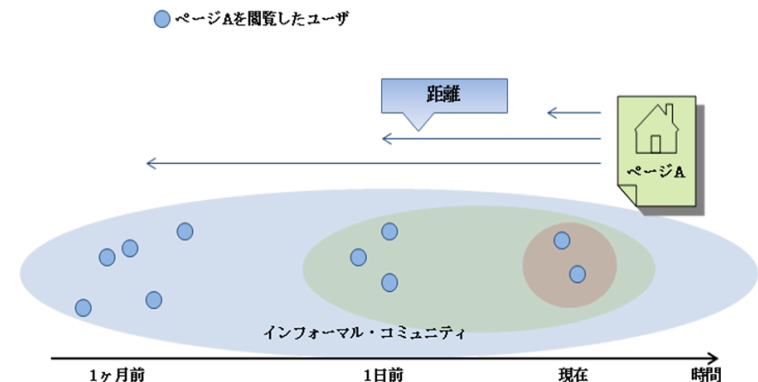


図 3 時間距離を基準としたインフォーマル・コミュニティの定義

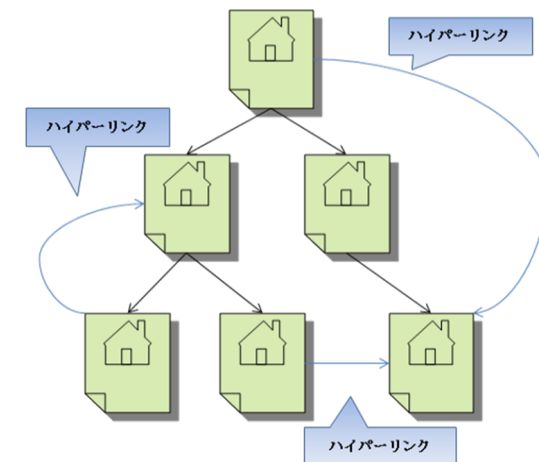


図 4 実際のホームページのつくり

## 5. ホームページ群選択方式の検討

本章では、ホームページの意味的な階層構造を判断するにあたり 3 つの選定方式を検討する。

- (1) ホームページのハイパーリンク(ハイパーリンク方式)
- (2) ホームページの URL のドメインやディレクトリ階層(ディレクトリ方式)
- (3) ユーザの履歴

以下にこれら 3 種類の選定方式案を実際のホームページを例に検討を行う。具体例として、評価を行うために極端なホームページ構成をもつ価格.com と Amazon[9]の 2 つのサイトを例に上記 3 方式の妥当性について検討する。

### 5.1 ホームページのハイパーリンク

ここではホームページのハイパーリンクを抽出し、リンク先をその Web ページの下位階層として判断する手法として検討する。ただし、リンク先が jpeg や gif 等の画像ファイルは除く。ホームページから抽出を行った結果、下位階層の Web ページを取得することができた。しかし、図 5 に示すように、あらゆる階層の Web ページを抽出してしまい、その中から意味的に下位階層である Web ページのみを選出することができないためこの手法では不完全である。

例えば、Amazon や価格.com では Web ページに、意味的な下位階層の Web ページと、ユーザごとのリコメンドや広告の Web ページのハイパーリンクが貼られており、それらを取得できるが、4 章で述べた通り、本来リコメンドは意味的な下位階層ではなく同位階層であり、下位階層と判断しない。しかし、この手法ではそれらを区別することができない。

### 5.2 ホームページの URL のドメインやディレクトリ階層

ここではホームページの URL のドメインやディレクトリ階層を抽出し、ホームページの階層構造を把握する手法を検討する。

価格.com では図 6 に示すように、URL ディレクトリの階層が、例えばパソコンの場合は <http://kakaku.com/pc/>、ノートパソコンの場合は <http://kakaku.com/pc/note-pc/>、そしてノートパソコンのカテゴリーvaio の場合 [http://kakaku.com/pc/note-pc/se\\_15\(vaio\)/](http://kakaku.com/pc/note-pc/se_15(vaio)/) となっており、意味的なホームページの階層と一致するので、構成の把握が比較的容易である。しかし、例外として製品 Web ページのディレクトリ構成が、例えば VAIO P シリーズ VGN-P92KS の場合 <http://kakaku.com/item/K0000125442/> となっており、URL ディレクトリが階層になっておらず独立しているため、この手法ではホームページの一部であっても意味的な階層と URL のディレクトリ階層が一致していないと構成全体を把握することが出来ない。

一方、Amazon ではドメインやディレクトリに法則性がないため、この手法で意味的な構成を把握することはできない。

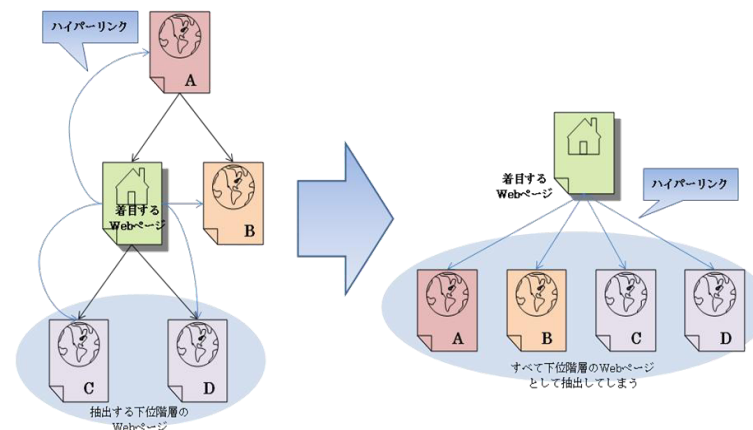


図 5 ホームページのハイパーリンクを用いた手法

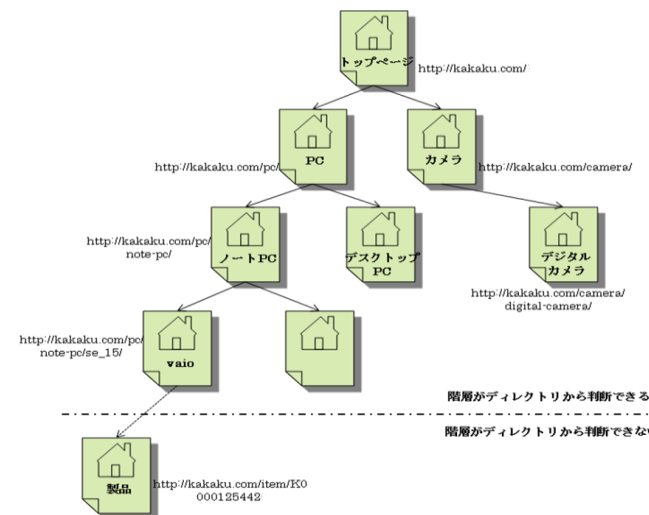


図 6 ホームページの URL のドメインやディレクトリ階層を用いた手法



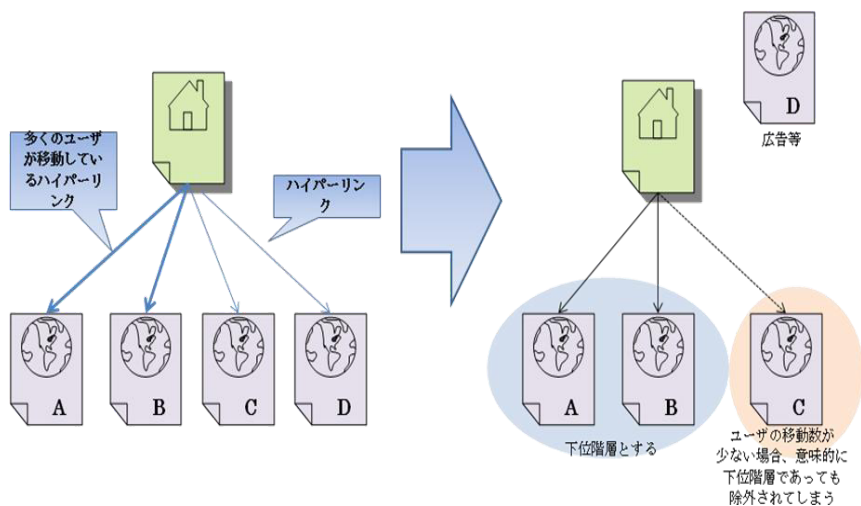


図 7 ホームページ内のユーザーの履歴を用いた手法

### 5.3 ユーザの履歴

ここではユーザーがホームページ内で移動した履歴からホームページの階層構造を把握する手法を検討する。ユーザーがホームページで Web ページ間を移動した履歴を抽出し、他のユーザーと照らし合わせることで、ある Web ページ  $n$  から多くのユーザーが移動した Web ページ  $n'$  を Web ページ  $n$  の下位階層とする手法である。この手法はユーザー履歴のデータを取得しなければならないため現段階では仮説となるが、ユーザーがホームページ内を移動する傾向として、1 番目に下位階層、2 番目に上位階層、最後に広告等に進むと想定される。この手法ならば、ホームページのハイパーリンクで取得したときに問題となった、下位階層以外の Web ページを、取得するユーザー数が多いほど遮断できる。その理由として価格.com、Amazon によく見られる、広告などの一過性の Web ページへのハイパーリンクの抽出を大幅に削減できるからである。しかし、この手法はユーザーの履歴の蓄積を前提としているため、ユーザー数の少ない初期の段階では効果をあげにくい特性がある。また、図 7 に示すように、移動数が少ないと、意味的に下位階層であっても判断されず、ホームページを見る傾向としてトップページや上位の階層に戻るユーザーが多いと、階層が逆転してしまう可能性がある。

そこで、Web ページ  $n'$  に最も多くのユーザーが移動してきた Web ページ  $n$  を Web ページ  $n'$  の上位階層とするという条件を加え、上記の条件と組み合わせることでそれらの問題を解決する。例として Web ページ  $n$  からユーザーの移動数が少ない Web ページ  $n'$  は

上記の条件だけでは下位階層と判断されないが、Web ページ  $n'$  に最も多く移動してきた Web ページが Web ページ  $n$  であるならば Web ページ  $n'$  を Web ページ  $n$  の下位階層と判断するというものである。

### 5.4 3 種の選択方式の組み合わせた手法の検討

ここでは上記で示した 3 種類の方式を組み合わせた手法を検討する。

まずハイパーリンク方式を用いて、ホームページ内でユーザーが閲覧している Web ページ  $n$  のハイパーリンクをすべて抽出する。抽出した Web ページ群を Web ページ  $N$  とする。

次にディレクトリ方式を使用する。ハイパーリンク先の Web ページ  $N$  に着目し、Web ページ  $N$  の URL のディレクトリ構成が、Web ページ  $n$  の URL のディレクトリ構成からパスが 1 つ増えていれば Web ページ  $N$  を Web ページ  $n$  の下位階層と判断する。

最後に URL のディレクトリ構成からでは下位階層と判断できない場合は、ユーザーの履歴の方式を利用する。多くのユーザーが Web ページ  $n$  から Web ページ  $N$  に移動していれば、Web ページ  $N$  を Web ページ  $n$  の下位階層と判断する。また、Web ページ  $n$  からユーザーの移動数が少ない Web ページ  $N$  についても、Web ページ  $N$  に最も多くのユーザーが移動してくる Web ページが Web ページ  $n$  であれば、Web ページ  $N$  を Web ページ  $n$  の下位階層とする。上記の条件以外の Web ページ  $N$  は Web ページ  $n$  の下位階層と判断しない。以上のアルゴリズムで汎用的にホームページの構成を把握できると考える。

価格.com では 5.2 節で示した通り、ある程度 URL のディレクトリ階層と意味的な階層が同じ構成になっているので、製品の Web ページ以外をディレクトリ階層で構成した後、ハイパーリンクを抽出して製品ページの階層を補うことで全体の構成を把握できる。しかし、広告等の一過性のハイパーリンクは、あらゆる階層に貼られるケースが多く、本来意味的に最下層に位置づけられるが、上位層の直下にも収められてしまうパターンが発生する。そこでユーザー履歴を利用した手法を組み合わせ、一時的なハイパーリンクを除外することで全体の構成を把握することができる。

一方、Amazon では 5.2 節で検討した通り、ドメインやディレクトリに法則性がないため、(2)の手法で意味的な構成を把握することはできないので、ホームページのハイパーリンクとユーザーの履歴の 2 種類の組み合わせの手法を検討することになる。ユーザー履歴を用いた手法を組み合わせることで、ハイパーリンクの手法の問題であった下位階層とそれ以外の区別をつけることができる。

上で述べた 3 つの手法と 3 種類の方式を組み合わせた手法の特徴を表 1 に示す。

表 1 提案した各手法と 3 種類の方式を組み合わせた手法の特徴

|             | 下位階層の取り<br>こぼしがない | 下位階層とそれ<br>以外の区別 | 履歴蓄積の<br>必要性 |
|-------------|-------------------|------------------|--------------|
| 方式(1)       | ○                 | ×                | 必要なし         |
| 方式(2)       | ×                 | △(*)             | 必要なし         |
| 方式(3)       | △                 | ○                | 必要           |
| (1)+(2)+(3) | ○                 | ○                | 必要           |

\*URL がディレクトリ階層構造をもつホームページに限る

### 5.5 考察

ユーザ履歴については、本来実際のユーザ履歴を用いて評価、検証しなければならない。本稿では、ユーザの使い方を想定しユーザの流れを仮定したが、想定したユーザの移動の傾向と、実際のユーザが移動する傾向が異なる場合期待した効果が得られない可能性がある。

別の問題として、Web ページの下位階層のハイパーリンクが少ない場合は、下位階層の Web ページへの移動数は全体の移動数のかなりの割合を占めるため判別が容易であると思われる。しかし、下位階層のハイパーリンクが大量に存在する場合、ユーザの移動数が薄められて広告との区別が困難になり、判別が不可能な場合が想定される。このように「多くのユーザが移動している」という基準や適応範囲を明確にする必要がある。

## 6. おわりに

本稿では、ホームページ内で適切なインフォーマル・コミュニティを形成するにあたり、ホームページの意味的な構成を把握するため方式を提案して検討した。

今後の課題としてユーザ履歴の実際のデータを収集、評価し、実際にコーディングして有効性を確認する必要がある。また、今回取り扱わなかった時間距離を使った範囲の検討と評価、そして、実際のつくりに関わらず、汎用的にホームページの構成を把握できるか検証し、ホームページの枠を超えた「叫ぶ」機能の適切な範囲と実装形式も追求していく。

## 参考文献

1) Kraut, R. E., Fish, R. S., Root, R. W., and Chalfonte, B. L.: Informal communication in organizations: Form, function, and technology, In S. Oskamp & S. Spacapan(Eds.), Peoples's Reactions to Technology, pp.145-199, Sage Publications London,1990.

2) 後藤慎弥：Web 上でのインフォーマル・コミュニケーション実現方式の提案,第 74 回グループウェアとネットワークサービス研究発表会

3) SPEEKY, <http://specky.jp>

4) Nesutu, <http://adgjm.net/nesutu.htm>

5) Saul Greenberg,Michael Rounding. "The notification collage: posting information to public and personal displays". SIGCHI,pp.515-521,2001.

6) 西本一志：インフォーマル・コミュニケーションによる知識共創場の構築，共創と複雑系シンポジウム予稿集,計測自動制御学会システムインテグレーション部門共創システム部会, pp.17-26,2006.

7) 熊谷知久,寺田実：Web サイト改良支援のための構造と履歴の視覚化システム, <http://www.interaction-ipsj.org/archives/paper2002/pdf2002/122paper.pdf>

8) 価格.com, <http://kakaku.com/>

9) Amazon, <http://www.amazon.co.jp/>