

親近度を考慮した Weblog アグリゲーションの提案

福井 宏幸[†] 上田 真由美[‡] 上島 紳一[†]

[†] 関西大学総合情報学部 〒569-1095 大阪府高槻市霊仙寺町 2-1-1

[‡] 名古屋大学情報連携基盤センター 〒464-8601 名古屋市千種区不老町 1

E-mail: [†] {fa10382,ueshima}@edu.kansai-u.ac.jp, [‡] ueda@itc.nagoya-u.ac.jp

あらまし 近年, Weblog, Wiki などのコミュニケーションを可能にする対話型の情報資源が出現している. 対話型情報資源は, 利用者が個人の意見や評価を表現できるため, 知人ネットワークを構成しやすい特徴がある. これらの情報資源を対象として個人が必要な情報を探する場合, 知人ネットワークを用いて検索範囲を絞り, 次に, 内容のデータについて絞り込む 2 段階の検索手順を踏む方法が信頼性と効率性の観点から優れていると考えられる.

本稿では, 他の利用者が保有する情報を知人ネットワーク内で共有することにより, 利用者が必要とする個人的な情報を容易に入手可能とする手法について議論する. 知人ネットワークの形成には, RDF(Resource Description Framework)の応用型である FOAF(Friend of a friend)表記で記述した Weblog の作成者のプロフィール情報を用いる. プロフィール情報には, 各 Weblog 情報と共に, その興味領域において有益な情報を配信する Weblog の作成者に関する情報を記述する.

また, 情報の集約アプリケーションとしては, 我々が提案しているプライベート仮想ライブラリ PVL を用いて知人ネットワーク上の情報を動的に集約する. PVL は, 複数のサーチエンジンから得られたデータを融合するメタサーチエンジン機構を備えているため, 知人ネットワークで得られた情報を融合しやすい. 両者を融合的に用いることで, 利用者は自らを起点とする信頼関係でつながるネットワーク内で, 同じ趣向を持つ信頼できる人物を探し, 各利用者の専門領域における知識を共有することが可能であると考えられる. ここでは, 知人ネットワークの拡がり方向をもとに親近度を考慮しながら, 実例を用いて Weblog 情報の集約機構について議論する.

キーワード Web 利用技術, 半構造データ, Weblog

Proposing a Scheme for Weblog Aggregation Utilizing Affinity Network

Hiroyuki Fukui[†] Mayumi Ueda[‡] and Shinichi Ueshima[†]

[†] Faculty of Informatics, Kansai University 2-1-1 Ryozenji, Takatsuki-shi, Osaka, 569-1095 Japan

[‡] Information Technology Center, Nagoya University 1Furo-cho, Chigusa-ku,

Nagoya-shi, Nagoya, 464-8601 Japan

E-mail: [†] {fa10382,ueshima}@edu.kansai-u.ac.jp, [‡] ueda@itc.nagoya-u.ac.jp

Abstract A novel type of interactive information resources, such as Weblog, Wiki, that enable users to communicate over the Internet has gaining focuses recently. On such information resources, individuals can express their personal opinions and/or evaluations of relevant resources explicitly using interaction capabilities. Hence, users easily create their reliable affinity networks on various categorized themes utilizing these resources. In searching information stored in them, the following two-step retrieval method is considered to be appropriate from the points of reliability and efficiency: First navigate such resources using affinity networks restricting search range, and then squeeze those data obtained in the first step. In this paper we provide and discuss the framework to acquire such information that may be essentially necessary to users by sharing information assets over affinity networks. We employ FOAF(Friend of a friend), which is an application of RDF(Resource Description Framework) to depict both Weblog creators, who deliver beneficial information, and meta data of the Weblog site. As an interface for users, we use our PVL or Private Virtual Library, that works as a meta search engine. By combining these components, users can search and share reliable and essential knowledge out of the Internet, utilizing affinity networks originated from himself/herself. We introduce the degree of affinity among users over the affinity network, and describe each component using examples of aggregation of Weblog's.

Keyword web utilization technology, semi-structured data, weblog

1.はじめに

今日、ネットワーク上には膨大な情報が存在し、利用者が自由に利用することが可能である。最近では、利用者がテキストのみならず写真やビデオ映像なども用いて対話的に書き込むことのできる Weblog や、Wiki などのグループ間のグループウェアなども登場するなど、情報資源の種類は広がりを続けており、利用者ごとに適切な情報を提供する仕組みに関する研究はますます重要度を増している。

インターネット上に存在する情報が増加したことにより、個人にとって重要な情報を選別するのが困難となり、インターネット上の情報資源を個別化して利用するための枠組みに関する研究が盛んに行われている[1]。代表的な検索エンジンにおける個別化機構に関する研究としては、利用者の過去のアクセス履歴を検索にフィードバックして利用する手法や[2][3][4]、類似したアクセス形態を持つ他の利用者の履歴を基に WEB ページを推薦する機構の研究[5]などが挙げられる。これらの手法は一般的な WEB サイトに含まれる情報を対象としていることが多く、また、あらかじめロボットなどによりサーバーに蓄積されたデータを下個人に関連データを抽出していることが多い。

これに対して、Weblog や Wiki などの対話型の情報資源は、個人の日記替わりや複数の利用者が構成するコミュニティの掲示板的な役割を果たすため、既存の検索エンジンが対象とする Web サイトと比較して、情報の更新が頻繁で、情報交換が行われる範囲が狭いという特徴を持つ。これらの対話型の情報資源を対象として個人が必要な情報を探する場合、情報資源にアクセスしている利用者間が構成するネットワークを用いて検索範囲を絞り、次に、内容のデータについて絞り込む 2 段階の検索手順を踏む方法がよいと考えられる。実際、2 章で述べるような FOAF 形式などで記述されたサイトの特徴を RDF(Resource Description Framework)で記述し、各利用者が関連サイトを書き込むことで、いわばインターネットの上に知人ネットワークを構成して情報交換の範囲を指定する手法も出現している。

本稿では、他の利用者が保有する情報を知人ネットワーク内で共有することにより、利用者が必要とする個人的な情報を容易に入手可能とする枠組みについて議論する。まず、知人ネットワークの形成には RDF の応用型である FOAF を用い、人々の流動的な表現を包含する Weblog を個人の情報ベースと捉え、知人ネットワーク内で共有する FOAF により各利用者の興味領域でつながるネットワークを辿ることが可能であるため、信頼できる情報を提供することを実現する。

また、このような情報資源に対しての情報の集約アプリケーションとしては、我々が提案しているプライベート仮想ライブラリ PVL を用いて知人ネットワーク上の情報を動的に集約する方法がある。PVL は、複数のサーチエンジンから得られたデータを融合するメタサーチエンジン機構を備えているため、知人ネットワークで得られた情報を融合しやすい。両者を融合的に用いることで、利用者は自らを起点とする信頼関係でつながるネットワーク内で、信頼できる人物を探し、各利用者の専門領域における知識を共有することが可能であると考えられる。

2.研究動機

2.1 対話型情報資源と知人ネットワーク

現在、Web 上には Weblog や Wiki などネットワーク上でのコミュニケーションを可能にする対話型情報資源が急増している。これらの対話型情報資源は、各分野において利用者が自分の経験や考えに基づいてデータを提供しており、提供される情報は膨大な Web 情報資源に対して利用者によるフィルタを掛けて抽出していることに相当する。また、そこに書き込まれる情報はある分野や話題に偏ることが多くなり、一般の Web サイトと比較して、情報が専門化されていたり、新しい情報である場合が多い。

また、昨今、“ソーシャルネットワーキング”の概念に基づく様々なサービスが提供されている[5]。利用者は一様に、自らの個人情報をそのコミュニティ内で開示することで、求職や友達作りといった目的に役立てている。参加するためにはすでに参加している人の招待が必要になることから、知人関係でつながる信頼できるコミュニティが形成される。

社会学では、知人間の情報共有は Stanley Milgram の「六次の隔たり論」[6]や、Mark Granovetter の「弱い絆の強さ論」[7]から有効であると考えられており、信頼できる知人関係で繋がる知人ネットワーク上で情報を共有することで情報の信頼性が向上することが示されている[8]。

2.2 P2P 型の情報集約

対話型情報資源に蓄積されている情報を集約する場合、既存の検索エンジンのようにロボットにより情報収集し、サーバーで全文索引を構成する手法よりも、この知人ネットワークを用いてネットワーク上の検索範囲を絞り、次に内容のデータについて絞り込む 2 段階の検索手順を踏む方法がよいと考えられる。この場合、各利用者をピアとみなし、リクエスト発行者を起点として利用者が自律的にリクエストの転送先を決定しながら動的にネットワークをツリー型に生成する P2P ネットワーキング技術が適切であると考えられる。

また、このような対話型情報資源においては、RSS によりサイトの要約を記述して関連サイトのメタ情報を提供したり、知人ネットワークを明示的に指定している場合が多く、P2P と RSS サイト記述の融合的利用や、次節で述べる FOAF を用いた知人の明示的指定などの方法が有効であると考えられる。特に Weblog は利用者が Web を巡航することで得られた知見や日常生活の記録場所として、自由に、対話的に書き込める掲示板の役割を持つ新しい対話型情報資源である。既存の HTML 言語などを用いずに簡単に利用できるため、現在では広く利用され、その数は増加の一途を辿っている。

2.3 セマンティック Web

日々増大する Web 資源を対象に機械可読なセマンティックタグを付与することで、情報検索やサービスの提供に応用するいわゆるセマンティック Web 技術がある。セマンティック Web は次世代 Web と言われており[9][10]、Web 上の資源に対してメタデータやオントロジーを用いて“意味”を付与することにより、エージェントがネットワーク上の資源の意味を解釈し、資源の組み合わせや処理を自動化することを可能にする技術で、W3C によって標準化が進められている。セマンティック Web を実現するための様々な階層で標準仕様が整い、ツール群が作成され、セマ

ンティック Web 技術を用いた様々な研究活動も行われている[11]。文献[12]では RDF 形式のメタデータに着目し、企業内の人や文章、スケジュールなど様々な資源の連携について述べられている。また、DAML S によるプロセス記述も解釈し、Web サービス連携を実装したエージェントを自動的に生成する機構に関する研究などが行われている[13]。

しかし、対象とするデータが Web 情報資源全般に及ぶ場合、その内容を正確に把握する必要があるため、その全てにタグを付与することは現実的ではない。また、タグ付けによる効果を即座に享受することができないため、Web 情報資源に対して積極的なタグ付けが行われるとは考えがたい。一方、RDF の応用型である Web サイトの要約を記述するための RSS は文法が簡易であることから広く利用されつつある。また、サイトへの書き込みに特定のフォーマットを用いる Weblog はそのコンテンツの RSS を自動生成・配信するツールを有する。今後、セマンティック Web 技術が普及するには、Weblog のような利用者が意識せずタグ付けを行える仕組みや、本稿で述べるような積極的なタグ付けを促進するためにその恩恵を即座に享受できる仕組みが必要であると考えられる。

3.FOAF ネットワーク

3.1 人に関する情報を表現する FOAF

本章では、FOAF 表現を用いた知人ネットワークの形成について述べる。FOAF(Friend of a friend)は、RDF を利用し、個人情報とその関係を記述するための語彙が定義されている。FOAF の記述例を図 1 に示す。FOAF は(1)RDF の宣言、(2)個人情報の記述、(3)知人情報を記述する部位の 3 つのブロックからなる。(1)では、この XML で使用する名前空間の URI を指定する。図 1 の例では RDF 自体の語彙と RDFS、FOAF で規定された語彙を定義している。(2)には myself@example.com というメールアドレスを持つ人物は、名前が myself でホームページを http://www.myselfhp.com/ に持ち、weblog を http://www.myself.blog.com/ に、仕事用のホームページを http://www.myworkplace.com/ に持つという情報が記述されている。また、関心を持つことに、http://www.kantei.go.jp/ と http://www.scei.co.jp/ が記述されている。前者は内閣府、後者はソニーコンピュータエンタテインメントのホームページの URL である。FOAF ではこのように興味を持つことを表現するために、その嗜好を反映する Web サイトのホームページで表現する。さらに(3)には foaf:knows によって囲まれている部分に知人である User A の情報が記述されている。ここでは RDF スキーマの rdfs:seeAlso を用いて、知人の FOAF ファイルの所在である http://userA.blog.com/foaf.rdf を記述している。エージェントはこの記述から知人として記されている人物の FOAF ファイルを把握することができる。

3.2 FOAF ネットワークの構造

FOAF による知人関係の表現は、J・Moreno (1934)の集団の内部構造を理解するための“ソシオメトリ理論で分析した人脈図(ソシオグラム)でいう“連鎖型”の知人関係に相当する。本研究では図 2 に示すように知人、知人の知人、さらなる知人というように知り合いの繋がりが一本の線のように展開した連鎖型の知人関係を用いて知人ネットワークを形成している。また、現実世界で人が複数の知り合いを持つように、FOAF においても複数の人物を知人として紹介することが出来る。そのため、表現される知人

```

<rdf:RDF (1)
  xmlns: rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns: rdfs = "http://www.w3.org/2000/01/rdf-schema#"
  xmlns: foaf = "http://xmlns.com/foaf/0.1/">

<foaf:Person (2)
  <foaf:name> Myself </foaf:name>
  <foaf:mbox rdf:resource = "mailto:myself@example.com"/>
  <foaf:homepage rdf:resource = "http://myself.hp.com/">
  <foaf:weblog rdf:resource = "http://myself.blog.com/">
  <foaf:workplaceHomepage rdf:resource = "http://myworkplace.com/">
  <foaf:interest rdf:resource = "http://www.kantei.go.jp/">
  <foaf:interest rdf:resource = "http://www.scei.co.jp/">

<foaf:knows (3)
  <foaf:Person>
  <foaf:name> User A </foaf:name>
  <foaf:mbox rdf:resource = "mailto:userA@example2.com"/>
  <rdfs:seeAlso rdf:resource = "http://userA.blog.com/foaf.rdf"/>
  </foaf:Person>
</foaf:knows>
</foaf:Person>
</rdf:RDF>

```

図 1 FOAF の記述例

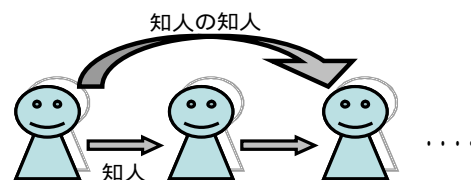


図 2 連鎖型の知人関係

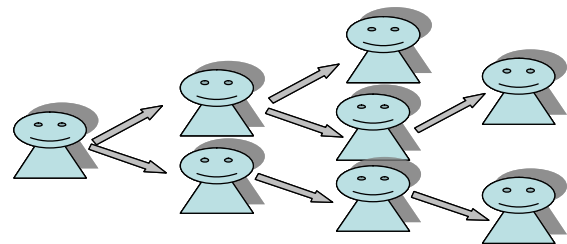


図 3 ツリー型の知人ネットワーク

の連鎖は図 3 に示すように、利用者本人を親ノード、利用者本人の知人を子、知人の知人を孫とする木構造を形成する。

3.3 FOAF ネットワークの構成要素

FOAF ネットワークを構成する要素は、第一に Weblog を所有する人物でなければならない。本稿においては Weblog に関する情報を FOAF ネットワーク内で共有することを目的としている。そのため、FOAF は人々が所有する Weblog 中の 1 ページとして記述されていることが望ましい。第二に、自身の FOAF が自分以外の第三者によって紹介されなければならない。FOAF ネットワークは、Weblog を所有する世界中のありとあらゆる人物の FOAF を繋ぐことを目的とするものではなく、様々な興味領域において有益な情報を配信する Weblog を所有する人物の FOAF をその興味領域で繋ぐことを目的としている。そのため、Weblog が重要な情報源であるかどうかは、人々の Web 巡航において、趣向に合致し、またその内容も有益であると判断される必要がある。さらにその内容が自らの FOAF を用いて、その Weblog を所有する人物の FOAF を紹介することで、初めて各利用者の興味領域で繋がる

FOAF ネットワークの一要素となる。

3.4 Weblog と FOAF の関連性

Weblog で提供される情報は、タイトルとその内容の記述に分離したフォーマットを用いるため、その記事の要約を RSS フィードとして自動配信することが可能である。また、RSS 自体は簡易な文法を用いているため、セマンティック Web を実現するための様々な種類のコンテンツに対するコンテナとしての可能性を秘めている。

RSS フィードは RSS アグリゲーションソフトウェアと組み合わせることにより、新着情報の取得など必要とする情報のプッシュ配信といった検索支援にも利用できる。しかし、Weblog の利用が拡大するにつれて、ネットワーク上で膨大な量の RSS フィードが提供されることになり、現在の Web から利用者が望む情報を取得する困難さと同様に、利用者が望む RSS フィードを取得することが困難である。さらに、Weblog は電子メールやメッセージといったコミュニケーションツールの1つとして利用することが可能である。しかし、Weblog を介した情報交換はトラックバック機能により、ある程度、限定された人物間での情報の交換に留まるなど、狭いコミュニティ間での利用に限定されているのが現状である。そこで、Weblog でより広範囲に情報の共有・交換を行うためには、(1)有益な情報を配信する Weblog の所在を人を通じて意識することなく知ることができる仕組み(2)自らが推薦する Weblog に関する情報をネットワーク上で共有可能とする仕組みを構築することが1つの手段となりえらる。 (1)では、人を介することで、自らが推薦する人がさらに推薦する人、またその人が推薦する人など信頼できる知人の連鎖から情報を収集することができるため、情報の信頼性が高まると考えられ、(2)の実現のためには、利用者が提供される情報を過不足なく収集することを可能にするため、統一化された記法に従って作成される必要があると考えられる。これは前述した FOAF を用いて記述することで解決することができる。

既存のアプリケーションと FOAF を融合的に利用することで、より広範囲の情報交換や密なコミュニティの形成を支援することが可能であると考えられる。しかし、現在の Web 上には FOAF は普及しているとは言えず、また、その利用は Weblog と同様に狭いコミュニティ間で相互に紹介しあうに留まっているのが現状である。FOAF には人の情報を表すための様々なタグが用意されているため、RSS のように利用者が意図することなく生成される仕組みを構築することは困難であると考えられる。今後、FOAF が普及するためには、FOAF-a-Matic[14]のようにメタデータの記述を支援するアプリケーションの利用や、特定の人に対してメタデータの作成を促す機構が必要になると考えられる。

4. 親近度を用いた情報検索

4.1 親近度の定義

FOAF は foaf:knows タグを用いて FOAF ファイルの在り処を示す URL を利用者の知人として紹介することができる。同様に紹介された知人もその興味領域において推薦する知人を紹介するため、知人ネットワークは利用者を起点とする木構造を形成する。

本稿では、信頼のおける知人間で情報を共有する枠組みを構築することによって、信頼性の高い情報を収集することを目的としているが、より信頼性の高い情報を優先して利

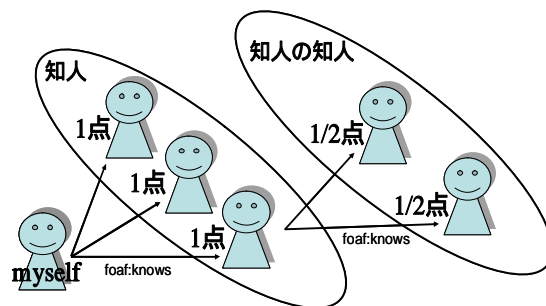


図 4 知人の重要度

用者に提供するために、何人の知人を介して利用者本人と繋がるのかで情報を提供する知人のランク付けを行う。利用者本人が紹介する知人の重要度を 1 とし、その知人がさらに紹介する知人(利用者にとっての知人の知人)の重要度を 1/2 とする。以下、知人を隔てる数に反比例する重要度をそれぞれの知人に与えていく(図 4)。本稿では、知人の隔たり次数で算出される知人の重要度を利用者に対する親近度と定義する。

4.2 Weblog 情報の特徴

Weblog はホームページのようにある特定のテーマにおける詳細な情報を記述するものではなく、各々が Web 巡航で興味のあるコンテンツに出会えば、そのコンテンツを指すリンクを貼り付けておくことで利用者間の情報共有を行うことを目的とする。利用者は各々の価値観で膨大な Web 資源から情報を選別する役割を果たす。しかし、既存の検索エンジンは入力されたキーワードを含む Web ページを形態素解析、TF/IDF などの自然言語処理技術や、リンク構造の密度により、コンテンツをランク付けし、評価が高いものから順に優先的に検索結果として表示している。Weblog 自体は密なリンク構造を有するもの、コンテンツそのものに関する情報を提供するものではなく、有益な情報源へのリンクとコンテンツに対する自他の主張を提供するものである。そのため、既存の検索エンジンに利用されるスコアリング手法では、Weblog で提供される有益な情報の重要度が低く設定されてしまうことがある。また、さして重要でないものがその密なリンク構造により重要度が高く設定されてしまうこともある。

現在、Weblog を対象とした検索サービスを提供するポータルサイトは多数存在するが、そのほとんどは利用者が指定するキーワードからその意図を満足に反映した結果の表示が行えてはいない。さらに、その記事内に貼られているリンクが示すコンテンツについても不問である。こうした背景から、知人ネットワーク内で共有する情報を Weblog とし、各々の興味領域において有益な情報を配信する Weblog を人々が共有できるような仕組みを構築することは有用であると考えられる。

4.3 複数の資源の統合手順

本章では、知人ネットワーク内で共有する情報資源を統合する手順について述べる。図 5 に示すように、利用者は自らの FOAF で紹介する知人の FOAF ファイルの所在を示す Knows タグ配下の SeeAlso タグを参照し、知人の FOAF ファイルを取得する。次に、取得した FOAF ファイルから Weblog タグを参照して、知人が所有する Weblog 情報の URL を取得する。さらに、知人が紹介する知人(利用者本人の知人の知人)の FOAF ファイルの所在を参照し、さらなる知人へとネットワークを辿り、Weblog 情報の収

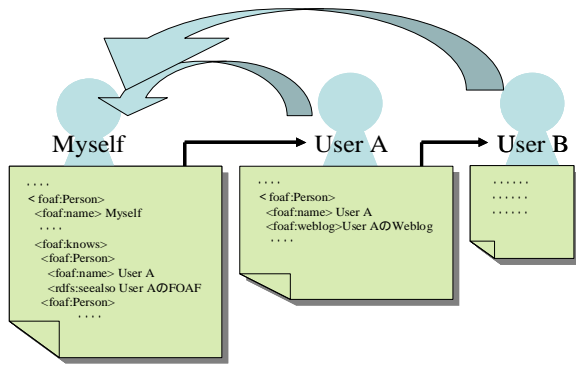


図 5 資源の統合手順

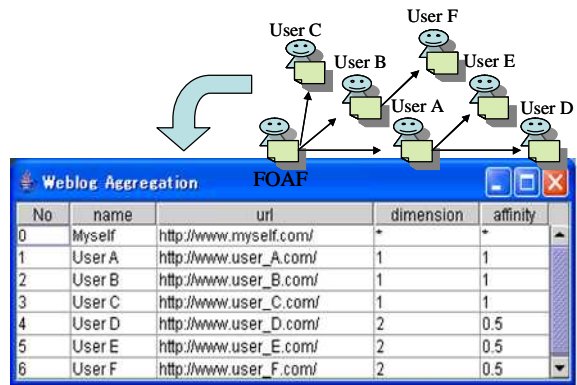


図 6 Weblog 情報の集約

集を行っていく。

4.1 節で述べたように、知人ネットワークは利用者を基点とする木構造を形成するため、幅優先探索でそのネットワークを辿ることにより、利用者と親近度の密な順に Weblog の情報を収集することができる。

図 6 は本研究のプロトタイプシステムを用いて、Myself を起点とする FOAF ネットワーク内で Weblog 情報を収集し、利用者本人に対する親近度をそれぞれ付与した例である。ここで、Myself はそれぞれ User A, User B, User C をこの順番で知人として紹介し、さらに User A・User B の両人は Myself の知人の知人である User D・E と User F をそれぞれ紹介している。本システムでは、知人の名前(name)、Weblog の URL(url)、隔たりの回数(dimension)、親近度(affinity)を提供する。

name	key	url	affinity
liris	Eclipse	http://www.liris.org/zope/blog/450	1
liris	Eclipse	http://www.liris.org/zope/blog/447	1
tsu	Eclipse	http://tsu.readymade.jp/archives/2004/1...	1
tsu	Eclipse	http://tsu.readymade.jp/archives/2004/1...	1
norakona	Eclipse	http://d.hatena.ne.jp/norakona/20041205	0.5
norakona	Eclipse	http://d.hatena.ne.jp/norakona/20041203	0.5
k-sakabe	Eclipse	http://blog-mac.k-sakabe.com/archives/...	0.5

図 7 プロトタイプの実行情例

5. プロトタイプシステム

5.1 Weblog 検索エンジンとの結合

本章では、Infoseek の Weblog 検索エンジンを利用し、特定のキーワードの検索結果を FOAF ネットワーク内で得られる情報でフィルタリングする手法を述べる。

1. Infoseek の Weblog 検索エンジンから特定のキーワードの検索結果の一覧を HTTP 通信で取得し、検索結果の HTML 文章をファイルに保存しておく。
2. 正規表現を用いて、保存された HTML 文章に含まれるリンク集合を取り出しておく。
3. 取り出されたリンク集合と、FOAF ネットワーク内で共有される有益な情報を配信する Weblog 情報の URL とを親近度順に照合する。
4. 適合するリンクを、その作者の情報と共に本システムにおける結果として利用者に提供する。

以上より、インターネット上から無作為に収集された Weblog 情報を、FOAF ネットワークから得られる“信頼性”が保証された情報源でフィルタリングし、不要な情報をふるい落とすことにより、利用者にとって有益と思われる情報を集約している。

本システムでは、“Eclipse”をキーワードとした Infoseek の検索結果を、Myself とその知人関係から成る 6 名の知人ネットワークを擬似的に構築し、その中で共有される Weblog 情報でフィルタリングを行った。図 7 に示すように、Myself と知人関係にある liris, tsu, norakona, k-sakabe, の 4 名が提供する記事を集約していることがわかる。

6. PVL と FOAF の融合による推薦機構

本章では、我々が提案しているプライベート仮想ライブ

ラリ PVL と FOAF を用いた知人ネットワークの融合について述べる。

6.1 PVL とは

我々はインターネット上の情報を本に見立てて本棚に格納し、本棚の個別化を進めるプライベート仮想ライブラリ PVL を提案している[15]。

PVL では“本棚”のメタファーを用いて現実の本棚と仮想の本棚の利点を融合した夢の本棚の実現を目指している。すなわち、我々の考える本棚メタファーは、インターネット上にある自分好みの情報を、自分の本棚にある本のように自由に扱うことができるような仮想の本棚を用いたメタファーで個人向けの夢の仮想ライブラリと考えている。

このような仮想ライブラリの実現には、検索機構が重要であると考え、PVL ではエージェント群を用いて、既存の検索エンジンが提供する情報を統合利用するメタ検索エンジン機構を備えている。本検索機構は、(1)各検索エンジンの偏りを生かすことで網羅性を高め、(2)重複性を生かすことで提要情報の信頼性を高めることを目指している。本検索機構では、インターネット上に存在する複数の検索エンジンを利用して問い合わせを行い、利用者プロフィールを用いて情報の集約・個別化して利用者毎に順序付けて情報の提供を行う。

PVL では、各検索エンジンが提供する情報を PVL で利用する形式に統一するためのラッパーの働きを備えたアクセスエージェントを、既存の検索エンジンごとに備えている。アクセスエージェントが収集した情報は、メディアータが集約し、個別化処理を行い利用者に提供する。利用者は提供された情報の中から、要・不要を判断し、必要と判断した情報を後日利用するために、本棚に格納することが可能である。

6.2 P2P 型推薦機構

PVL では、他の利用者が繰り返し PVL を利用し、個人

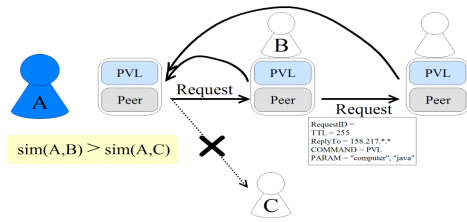


図 9 リクエストの伝播と各ピアの動作

利用を目的に本棚に蓄えた情報を資源とした P2P 型推薦機構を備えている。本推薦機構は、ネットワーク上に分散した不特定多数の PVL 利用者を対象に情報収集・提供を可能にするため、P2P 型ネットワークを通信手段に用いる。P2P 型ネットワークを通信手段に用いることにより、不特定多数の PVL 利用者を対象とした PVL コミュニティを形成することが可能である。

さらに、利用者間の類似度を考慮して形成した小規模コミュニティ内でリクエストの伝播を行う個人向けの推薦機構も備えている。全 PVL 利用者を対象とした上記の推薦機構はアライアンス表に記述されたリクエストの伝播先に推薦要求を送っていたが、小規模コミュニティ内での個人向け推薦機構はアライアンス表と利用者間の類似度によって推薦要求の伝播先を決定する。類似度を考慮することにより、個人の嗜好にあった情報を入手することが可能である[16]。ただし、利用者間の類似度は、PVL 利用時に生じるアクセスエージェントの得点で構成される各利用者の特徴ベクトルを用いてコサイン尺度により算出する(式(1))。

$$\text{sim}(v_a, v_b) = \frac{(v_a, v_b)}{|v_a| |v_b|} \quad (1)$$

ここで、 v_a 、 v_b は利用者 A、B の特徴ベクトルを表し、 $\text{sim}(v_a, v_b)$ は利用者 A、B の類似度をあらわす。

図 9 では、推薦要求者となる利用者 A にとって、利用者 B、C との類似度を式(1)で算出した際 $\text{sim}(A, B) > \text{sim}(A, C)$ となり、利用者 B との類似度が高いため、利用者 A はリクエストを利用者 B に伝播し、利用者 C には伝播しない様子を示している。リクエストを受け取った利用者 B は、更に次の利用者にリクエストを伝播している。このように各ピアで類似度に基づき転送先を切り替えて、A を起点とするツリー型のネットワークを構成している。

図 10 では、推薦要求者 A が自身の持つアライアンス表に持つピア集合 B、C、D から、特徴ベクトル間の類似度を比較して B に転送している様子を示す。更に、B は自身の持つアライアンス表に持つピア集合 E、F から F を選択し、リクエストを転送している。

同様にして、F から J、J から N へとリクエストが伝播されている。このリクエストの伝播については、A を起点とするツリー型ネットワークの中の 1 本の経路が使用されたことを示す。

6.3 FOAF を用いた P2P 型推薦機構

6.2 節で述べた個人向け推薦機構は、利用者間の類似度を考慮した個人向けの情報を提供する推薦機構である。しかし、類似度の算出にアクセスエージェントで構成する利用者の特徴ベクトルを用いるため、個人の嗜好を正確に反映しているとは言い難い。

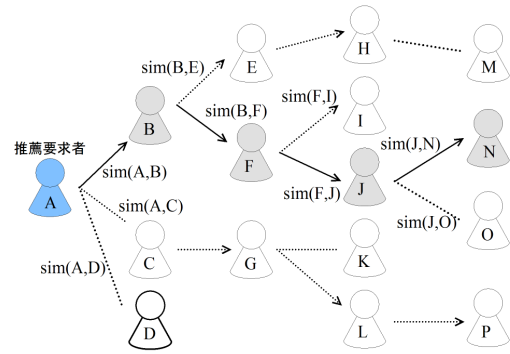


図 10 推薦要求の伝播

そこで、ここでは小規模コミュニティの形成に FOAF 表記を用いた知人ネットワークを用いる。知人ネットワークを用いることにより、6.2 節の手法で実現した利用傾向による利用者間の類似度のみでなく、知人関係による信頼性と、FOAF 表記によって明確に記述された個人の興味領域を考慮したコミュニティを形成することが可能となり、より個人の嗜好にあった情報を提供することができる。

7. メタデータ記述の支援機構

3.4 節で述べたように、現在の Web 上には FOAF が普及しているとは言い難く、一部の人が興味本位に記述したり、知った仲間同士で互いを紹介しあったりといった利用に限られている感がある。FOAF と既存のコミュニケーションツールとを組み合わせることで、本稿で提案するような知人間で共有される情報を検索支援に利用したり、Weblog でより密なコミュニティを形成することを可能にしたりするなどの利点が考えられるものの、FOAF 自体の利用に関してはまだ明確なビジョンがあるわけではなく、FOAF もある程度まとまった数がなければ、それを作成することによる恩恵を満足に享受することはできないと考えられる。また、FOAF にはその人が関心を持つことを表現するために foaf:interest タグや、その人物の勤務先を示す foaf:workplacehomepage タグなど、人を描写するための様々なタグが用意されているが、そのリソースには URI を用いるため、紹介する知人がどのようなことに関心がある人物で、その専門度はどのくらいであるのかといった細やかな情報をプログラムが判断することは困難である。

本稿で述べた情報共有手法を実現するためには相当数の FOAF が必要になるし、また、詳細な人物描写を本システムの利用者間でのみ識別ができるような記法で表現しておけばプログラムで扱うことが容易になり、知人ネットワークで共有される情報を個人の興味領域で取捨選択することができるため、得られる情報の信頼性がより高まると考えられる。そこで、例えば特定の Weblog を何度も閲覧したり、積極的にコメントを書き込んだりする人々に対して、FOAF を容易に作成できるような GUI を提供するなどしてその普及に貢献したり、また本システムの利用者に対しては foaf:knows で紹介する知人ごとにその興味領域や専門度をプログラムが理解できるような形式で FOAF の記述を行ってもらうための仕組みを提供するなどが考えられる。

8. おわりに

本稿では人のつながりを用いて形成される FOAF ネットワーク上で情報を共有し、既存の検索手法に融合する手法を提案した。2.1 節で述べたソーシャルネットワーキングサービスやポータルサイトが提供する Weblog サービスでは、個人の情報を記述するのに独自のフォーマットを用いているため汎用性に欠ける。そこで FOAF のような統一化されたフォーマットを用いることで、今まで狭いコミュニティ内での利用に限定されていた情報を広範囲に公開したり共有したりすることが可能になる。また、人に関する情報は既存の検索手法やサービスに応用することで、新たなサービス産出を助長するものと考えられる。

FOAF には人に関する情報を細かく表すためのタグが多数用意されている。これらのタグのリソースは Web 資源を対象としており、興味と合致する内容が記述されている Web ページの URL を用いて人が興味を持つことを表現する。これらのタグを有効的に利用することで FOAF ネットワーク上で共有される情報の中で個人の興味領域に合致するもののみを収集するなど情報の取捨選択が行えるが、Web 上にはそれこそ膨大な量の Web ページが存在するため、URL で表現されるこれらの情報はプログラムが利用者の要求に合わせて適切に判断するのは容易ではない。そのため、利用者があるパラダイムや分野に属する情報を得たい場合、FOAF 単体を利用したの実現は困難であると考えられるため、他のツールと FOAF を融合的に利用する必要がある。その際、利用者の多様な要求を充足する仕組みを構築することは必要不可欠である。

本稿では、既存のアプリケーションをフィルタすることで信頼性の高い情報を提供することを目的としたが、有用性の面を考えた場合、利用者の問い合わせに適合する情報のみが必ずしも役に立つとは言えない。文章量の少ない Weblog はなおさらでそうである。そこで、特定の興味領域で繋がる FOAF ネットワークを形成することができれば、同一の分野に属する情報を容易に収集することが可能となり、利用者が有用とする情報を得る手助けとなろう。

最後に Web 上のオープンな環境で個人情報を開示する行為は個人のプライバシーに関わる問題である。例えば、FOAF では人物を一意的に識別するためにメールアドレスを ID として用いている。しかし、メールアドレスをそのまま公開するとスパムの対象となるため、ハッシュ関数で暗号化することでスパムから保護している。そこで、個人が保有する情報を情報検索やサービスの提供に利用する際、世間に公開しても良い情報と非公開にしておく情報はどのように区別するのかといったことや、信頼のおける人物間でのみ共有したい情報をどのように外界から保護するのかといった問題があり、今後議論される必要がある。

謝辞

本研究の一部は文部科学省私立大学学術研究高度化推進事業オープン・リサーチ・センター整備事業（平成 15 年度-20 年度）並びに同サイバーキャンパス整備事業（平成 14 年度-16 年度）の一環として、また、同省の「知的資産の電子的な保存・活用を支援するソフトウェア基盤技術の構築」プロジェクトの支援により行われた。

参考文献

- [1] 清光英成, 竹内淳記, “ Web データの個別化と環境適応”, 情報処理学会論文誌: データベース, Vol.42, No.SIG 8(TOD10), pp.185 -194, 2001
- [2] 仲川こころ, 木下敦史, 高田義朗, 関浩之, “ ユーザによる文書ランキングの調整が可能な対話的 WWW 検索手法の提案”, 情報処理学会第 63 回全国大会論文集(3), pp.131 -132, 2001
- [3] Google: <http://www.yahoo.co.jp/>
- [4] Yahoo: <http://www.google.co.jp/>
- [5] 武田英明, 大向一輝 “ Weblog の現在と展望 - セマンティック Web およびソーシャルネットワーキングの基盤として - ”, 情報処理学会誌, 45 巻, 6 号, pp.586 - 593, June 2004.
- [6] Stanley Milgram : <http://stanleymilgram.com/>
- [7] Mark Granovetter : <http://www.stanford.edu/dept/soc/people/faculty/granovetter/granovet.html>
- [8] TysonCondie, SepandarD.Kamvar, HectorGarcia-MoliNa, “ AdaptivePeer-to-Peer Topologies:”, Proc.4th International Conference on Peer-to-Peer Computing (2004)
- [9] The Semantic Web, Tim Berners-Lee, James Handler and Ora Lassia, Scientific America, pp.28-37, May 2001
- [10] 浦本直彦: “ Semantic Web の可能性 ”, 電子情報通信学会言語理解とコミュニケーション研究会シンポジウム, Sep. 2001 .
- [11] 荻野達也編集, “ セマンティック Web 特集 ”, 情報処理学会誌, Vol.43, No.7, July 2002 .
- [12] 片山佳則, 小櫻文彦, 井形信之, 渡辺勇, 津田宏: “ セマンティックグループウェア WorkWare++ と KnowWho 検索への応用 ”, 情報処理学会研究報告, Vol.DBS-130, FI-71, pp.9-16, 2003 .
- [13] 福田直樹, 和泉憲明, 山口高平: “ DAML-S に基づくサービス連携エージェントの実現について ”, 人工知能学会研究会資料, SIG-SW&ONT-A301-04, June 2003
- [14] Foaf-a-Matic : <http://www.ldodds.com/foaf/foaf-a-matic.ja.html>
- [15] 上田真由美, 上島紳一, “ エージェントの得点付け手法により個別化を行うプライベート仮想ライブラリ PVL の提案と評価 ”, 情報処理学会論文誌: データベース Vol.43 No.SIG2(TOD16), pp.103--114, 2002
- [16] 上田真由美, 上島紳一, “ PVL のための個人向け推薦機構に関する検討 ”, 電子情報通信学会 第 14 回データ工学ワークショップ DEWS2003 オンライン論文集 5-P-4, 2003
- [17] Masahide Kanzaki : <http://www.kanzaki.com/docs/sw/>
- [18] FOAF Project : <http://rdfweb.org/topic/FoafProject>