

コミュニティウェブソフトウェアにおけるアクティブマイニング

石川 孝

日本工業大学工学部 〒345-8501 埼玉県南埼玉郡宮代町学園台 4-1

E-mail: tisikawa@nit.ac.jp

あらまし コミュニティウェブソフトウェアはインターネットを利用してコミュニティ活動を支援するためのウェブアプリケーションを構築するツールである。このツールによって構築されるコミュニティウェブは、メーリングリストと連動する掲示板によってコミュニケーションを支援するとともに、掲示板から抽出されたトピックを使ってコミュニティに関連するニュースをアクティブに検索して提供する。また、コミュニティウェブ同士のコラボレーションによってインターネット検索サイトに頼らないアクティブな分散情報検索を可能にする。

キーワード コミュニティウェブ, コミュニティ支援, アクティブ情報検索, アクティブコラボレーション, トピック抽出, 分散情報検索

Active Mining in Community Web Software

Takashi ISHIKAWA

Faculty of Engineering, Nippon Institute of Technology 4-1 Gakuendai, Miyashiro, Saitama, 345-8501 Japan

E-mail: tisikawa@nit.ac.jp

Abstract Community Web Software is a building tool for web applications to support activity of communities using the Internet. A community web built by the tool supports communication in the community via the bulletin board cooperating with a mailing list and provides news information actively retrieved with topics distilled from the bulletin board. Also community webs are able to execute active distributed information retrieval by the collaboration of them without relying Internet search sites.

Keyword community web, community support, active information retrieval, active collaboration, topic distillation, distributed information retrieval

1. はじめに

インターネットの HTTP プロトコルは、コミュニケーションと情報共有を同時に行うことができるウェブアプリケーションをもたらした。ウェブアプリケーションは、経済、社会、教育、文化、地域などにおけるさまざまなコミュニティ（共同体の意）にネットワークによる活動の場を提供する。このようなコミュニティのためのウェブアプリケーションを総称してコミュニティウェブと呼んでいる[1]。コミュニティウェブは、そのコミュニティの運営者によってコンテンツが管理されているため、コミュニティに対して適時に情報を提供するには運営者に対する支援が必要になる。コミュニティウェブにおける支援機能には、コンテンツ管理、情報収集配信、アーカイブ管理、ユーザ管理などがある。

本研究の目的は、コミュニティのメンバーと運営者に対する支援機能をもつコミュニティウェブを構築するためのコミュニティウェブソフトウェアの開発である[2]。このソフトウェアの目標は、ソフトウェアの専

門家でなくてもコミュニティウェブを開発し管理運営が容易にできるようにすることである。本論文は、コミュニティウェブソフトウェアにおけるアクティブマイニング技術の適用について述べる。

2. コミュニティウェブの課題

コミュニティウェブは、インターネット上のコミュニティに“コミュニケーションの場”を提供するウェブアプリケーションである。インターネット上のコミュニティは、メーリングリスト、掲示板などを使ってコミュニケーションを行っている。これらのコミュニケーション手段は、時間的空間的制約をもたないことによって、コミュニティ活動の場を拡大する。またコミュニティウェブは、そのコンテンツやアーカイブ機能によってコミュニティの“情報共有の場”となる。

コミュニティウェブは、実社会のさまざまな共同体が運営するウェブサイトや、インターネットプロバイダなどが運営するコミュニティサイトにおいて多数存在している。これらのサイトには管理者が存在し、ウ

ウェブページ上のコンテンツの管理とメーリングリストや掲示板の管理を行っている。このサイト管理者の役割は、ネットワークコミュニティの運営に重要である。コミュニティの参加者を維持、増大させるためには、コンテンツを更新してゆくことが不可欠であり、掲示板上の不適切な発言等を削除することも必要である。実社会の共同体が運営するコミュニティでは、実社会の規範が一般に存在するために、サイト運営に問題はあまり生じない。むしろサイト運営は共同体運営と一体といってもよい。しかし、コミュニティサイト上の個人を主体とするコミュニティウェブでは、サイト運営者がコミュニティの参加者（または管理者）ではないことによる多くの問題が生じている。例えば、ウェブページ上の共有情報であるコンテンツが適切に更新されなかったり、掲示板上の発言の評価基準が異なることによる発言妨害が起きたりすることなどである。コミュニティウェブの管理運営を適切に行うためには、そのコミュニティ活動支援機能が重要な課題となる。

3. コミュニティ活動支援

コミュニティウェブにおけるコミュニティ活動支援は、コンテンツ管理およびアーカイブ管理などによって行われる[3]。コミュニティにとって有用な情報を適時に適切に提供することがコミュニティの活動を円滑にする。コミュニティウェブのコンテンツには、コミュニティ自体に関する記述、コミュニティが関心をもっている情報（ニュースや関連サイト）、コミュニケーションログ（メーリングリストのアーカイブや掲示板）、共有情報アーカイブのインデックスなどが含まれる。これらのコンテンツの情報源はそのほとんどがコミュニティのメンバーや運営者であるが、ニュースや関連サイトについてはウェブ上にも多くの情報源が存在する。ソフトウェアによってコミュニティの活動支援を行える1つの可能性はここにある。

コミュニティが関心をもっている情報を適時に適切に提供するソフトウェアには2つの大きな課題がある。その1つはコミュニティが関心をもっていることがらをどのように特徴づけるかである。もう1つはそのような関心に関連する情報をどのように収集するかである。コミュニティの関心はコミュニティの活動目的を中心としてメンバーの関心の変化やメンバーの変更によって時間的に変化する。メンバーの関心はコミュニケーションの内容として現れるので、コミュニケーションログからコミュニティの関心を捉えることが課題になる。また、関心をどのように表現するかも問題である。コミュニティの関心が捉えられたとして、それに関連する情報の方も情報源であるウェブにおいてたえず変化している。関心が一定であったとしても

関連する情報が新たに発生しうる。このように変化する関心と変化する情報源をどのようにアクティブにバランスさせるかが課題になる。

4. コミュニティウェブソフトウェア

コミュニティウェブソフトウェアはコミュニティウェブを構築するためのソフトウェアである。このソフトウェアは、インターネットからダウンロードし、ウェブアプリケーションを実行する環境にインストールして使用する。コミュニティウェブの構築および管理はインターネット上で管理ツールを使って行う。構築されたコミュニティウェブはインターネット上で不特定または特定の利用者が閲覧できる。利用者に提供される機能は管理ツールによって設定される。本予稿執筆時の機能評価版 (Ver. 0.8) は Jakarta プロジェクトの Struts フレームワーク [4] を使って実現されている。この版で構築されたコミュニティウェブのサンプル画面を図1に示す。

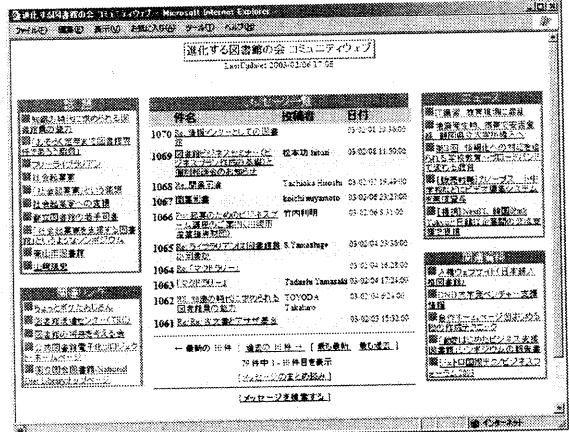


図1. コミュニティウェブのサンプル画面

コミュニティウェブソフトウェアによって構築されたコミュニティウェブは、つぎの機能を標準として提供する。

- (1) メーリングリストと連動する掲示板（メッセージリスト）

コミュニティウェブは標準のコミュニケーション手段としてメーリングリストによる電子メールを使用する。コミュニティのメンバーがメーリングリストに電子メールを送ると、その内容が掲示板のメッセージとして追加される。なおメーリングリストは、コミュニティウェブとは独立に管理される。

(2) 掲示板のメッセージを分類するトピックリスト

トピックリストは掲示板の中のメッセージから抽出されたトピックである [5]。トピックは設定によりキーワードのリストまたはキーワードを含む名詞句として表示される。表示されたトピックをクリックすると該当するメッセージの一覧が表示される。トピックの更新時刻は管理ツールによって設定される。

(3) トピックによって検索されたニュースリスト

ニュースリストは、管理ツールによって設定されたニュースサイトから、トピックを表すキーワードの組 (AND 検索を意味する) を使って検索されたニュースの見出しを表示する。この見出しをクリックするとリンク先のニュースが表示される。新しいニュースが検索されるとニュースリストに追加される。ニュースリストの更新時刻は管理ツールによって設定される。

(4) 管理者によって設定された関連サイトおよび関連資料リスト

これらのリストはコミュニティの情報共有のためのアーカイブとして利用される。その設定は管理ツールによって行う。

コミュニティウェブで表示されるリストは、各項目への利用者のアクセス数によってその表示順序を変更することができる (c.f. [6])。例えば、ニュースリストにおいて、アクセス数がより多い見出しは上位に移動するので、新しいニュースが追加されてもリストの中に長く留まるようになる。また、関連サイトや関連資料のリストにおいてもリスト中の順位によってメンバーの利用頻度を表現することができるようになる。この機能はコミュニティウェブのコンテンツに対する利用者の協調的な選択を可能にする。リストの表示順序の変更に対する設定は管理ツールによって行う。

5. アクティブニュース検索

コミュニティウェブにおいてメーリングリストのメッセージから抽出されたトピックは、コミュニティの関心 (その表現がトピックである) に関連するニュースなどの情報をアクティブに検索することに利用される (図 2)。この機能をアクティブニュース検索と呼ぶ。この場合の情報検索は 2 つの意味でアクティブである。1 つはコミュニティに新しいトピックが持ち込まれたときに検索を行うことである。コミュニティウェブのトピックリストは定期的に更新されるので、新しいトピックはキーワードの新しい組み合わせとして認識される。もう 1 つはニュースサイトに新しいニュースが掲載されたときに検索を行うことである。ニュ

ースサイトは一定間隔で更新されるので、その都度、トピックに関連する新しいニュースを発見する可能性がある。このコミュニティとインターネットの両方の変化に対応することは、コミュニティウェブのアクティブな情報検索の特徴である。

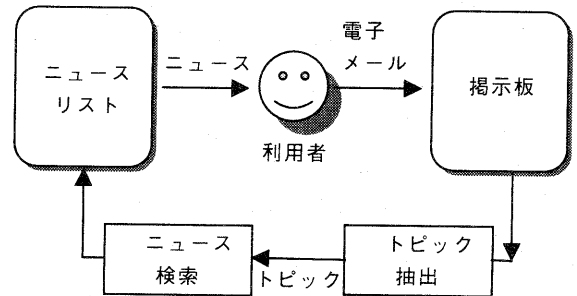


図 2. アクティブニュース検索

アクティブニュース検索は、つぎのようにしてコミュニティの活動を間接的に支援することが期待される。第 1 にコミュニティウェブがコミュニティのトピックに関連するニュースをいち早く提供することである。これによってコミュニティの活動がその環境の変化に適時に対応できるようになる。第 2 にコミュニティのメンバーが陽には意識していないトピックに関連するニュースを発見できるようにすることである。これによって、関連するニュースを媒介としてコミュニティ活動の焦点化が促される。第 3 にコミュニティウェブ上での関連するニュースへのアクセス頻度によってニュースの注目度を評価し、見出しの並べ替えを行うことである。これによって、注目度の高いニュースがより長く表示されることで活動の安定化が図られる。

6. アクティブコラボレーション

コミュニティウェブを構築するツールであるコミュニティウェブソフトウェアは、このツールで構築されたすべてのコミュニティウェブからダウンロードできる。コミュニティウェブをインストールすると、そのコミュニティにウェブソフトウェアをダウンロードした元のコミュニティウェブの URL が記録されるとともに、新しいコミュニティウェブに自動的に登録される。これによってすべてのコミュニティウェブ同士が相互に通信できるようになる (大元のダウンロードサイトは 1 つとする)。通信はコミュニティウェブの中のコラボレーションエージェントによって実行される [7]。コラボレーションエージェントはインターネットの電子

メールと同様のメカニズムによってすべてのコミュニティウェブ間での情報交換を可能にする(図3)。

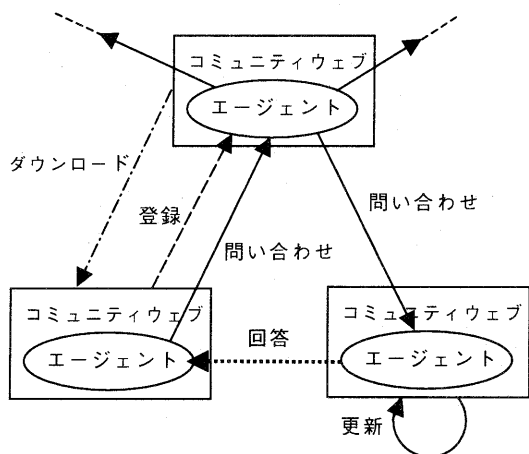


図3. アクティブコラボレーション

コミュニティウェブ間での情報交換機能を使うと、あるコミュニティウェブ内でのメッセージやニュースの検索を他のすべてのコミュニティウェブに対して行うことができるようになる。さらに、他のコミュニティウェブから一度検索の問い合わせがあると、その問い合わせを記録しておき、あとでそのコミュニティウェブに該当する情報が発生したときに自動的に回答を送ることができる。この機能をアクティブコラボレーションと呼ぶ。アクティブコラボレーションによって、他のコミュニティウェブが興味をもっている情報をアクティブに発信できるようになる。また、この機能は情報源が変化する環境でのダイナミックな分散情報検索を可能にする。

7. 今後の課題

コミュニティウェブソフトウェアは、ソフトウェアの専門家でない利用者を目標としているために、高い完成度が求められる。このため、コミュニティウェブとしての基本機能とコミュニティ活動支援のために必要最小限のアクティブマイニング機能の実現を図っている。機能評価版(Ver. 0.8)ではアクティブニュース検索が実現されているが、つぎの公開評価版(Ver. 0.9)ではインターネットでのダウンロード機能を追加してアクティブコラボレーションを実現する計画である。アクティブコラボレーションが可能になると、コミュニティウェブソフトウェアの自動更新も可能になる。

アクティブニュース検索などのアクティブ情報検

索については、抽出するトピックがコミュニティの関心をより適切に捉えるようにすることが改良の課題である。今後の研究の方向としては、キーワードの重要度やクラスタリングのアルゴリズムの改良だけでなく、メッセージの応答関係を反映するようにメッセージを分類することを検討している。また、検索されたニュースなどの情報をより適切なものに絞り込むことも課題である。アクティブコラボレーションについては、その基盤となるコラボレーションエージェントの実現が当面の課題である。その上で、分散情報検索アルゴリズムの効率を改良していくことが今後の課題になる。

8. おわりに

この論文は、インターネットを利用してコミュニティ活動を支援するコミュニティウェブにおけるアクティブニュース検索とアクティブコラボレーションについて述べた。アクティブマイニング技術のコミュニティウェブへの適用は、エージェントの自律性と知能性の向上によって、ますます適用領域が拡大することが予想される。特に、分散情報検索はウェブ全体を統合された知識ベースとして利用する可能性をもたらす。

文 献

- [1] Christophides, V. et al., Community Webs (C-Webs): Technological Assessment and System Architecture, <http://citeseer.nj.nec.com/christophides00community.html>, 2000.
- [2] 石川 孝, コミュニティウェブを支援する InterMediator エージェントの開発, 第29回知能システムシンポジウム, 2002.
- [3] 梅木秀雄, 下郡信宏, 横田健彦, ネットワークコミュニティ形成支援, 情報処理学会研究報告, 情報メディア 37-5, pp.25-30, 2000.
- [4] ウェリントン・ダシルバ, JSP & タグライブラリー Struts による Web 開発シナリオ, 高森郁子訳, ピアソン・エデュケーション, 2002.
- [5] 余 東明, 石川 孝, コミュニティウェブにおける掲示板からのトピック抽出, 第1回情報科学技術フォーラム, 一般講演論文集, 第2分冊, E-17, pp.115-116, 2002.
- [6] Perkowitz, M. and Etzioni, O., Towards adaptive Web sites: Conceptual framework and case study, *Artificial Intelligence*, Vol. 118, pp.245-275, 2000.
- [7] Thanachai Wongvibulsin, Takashi Ishikawa, Open Collaborative Agents to support Community Webs, In *Proceedings of FIT2002*, L-16, pp.33-34, 2002.