

2007年12月3日

(報道発表資料)

日本電信電話株式会社
NTTレゾナント株式会社

ブログの話題を仮想大陸で探索できる「BLOGRANGER TG」の 共同実験を開始

～ブログの話題を地形図として俯瞰する世界初の検索技術を適用～

日本電信電話株式会社（以下NTT、本社：東京都千代田区、代表取締役社長：三浦惺）と、ポータルサイト「goo」を運営するNTTレゾナント株式会社（以下NTTレゾナント、本社：東京都千代田区、代表取締役社長：和才博美）は、ブログ記事を分析し、仮想大陸の地形図上に、概念的に近いキーワードのタグ（※1）同士を配置することで、利用者の関心にマッチしたブログ記事を視覚的に見つけやすくした、まったく新しいブログ検索サービス「BLOGRANGER TG」の共同実験を開始します。

本サービスはポータルサイト「goo」上の実験サイト「gooラボ」（URL：<http://labs.goo.ne.jp/>）で本日12月3日から1年間の提供を予定しており、商用環境での実用性を検証します。

1. 「BLOGRANGER TG」の概要（別紙1、別紙2参照）

「BLOGRANGER TG」は、NTTのサイバーソリューション研究所とコミュニケーション科学基礎研究所（以下、NTTの研究所）が開発した「オートタギング技術」と「トピック空間可視化技術」を適用した新しいブログ検索サービスです。「BLOGRANGER TG」は、最新1ヶ月間に収集したすべてのブログ記事に、任意のテキストに対して適切なタグを推定する「オートタギング技術」を用いて、分類のキーワードであるタグを自動付与します。また、「トピック空間可視化技術」を用いて、概念的に関連するタグ同士の関係を二次元の仮想大陸の地形図（タグマップ）として自動的に可視化します。そうすることで、利用者は仮想大陸をマウスでスクロール（探索）して、仮想大陸上に配置されたキーワード（タグ）をクリックすることで関連したブログ記事を次々と閲覧することができます。

今回提供する「BLOGRANGER TG」で採用した、概念的な地図を利用したまったく新しい情報検索のスタイルは、ブログ検索として利用するだけでなく、ニュースや動画を含むインターネット上の様々なコンテンツを検索するインタフェースとして広く適用可能なものです。

さらに、このような情報検索スタイルの様々な分野への適用可能性を探るため、これらオートタギング機能と、トピック空間可視化技術によって

自動生成される仮想大陸の地形図（タグマップ）を配信する機能は、Web API（※2）としても提供します。これらにより、例えば、サイト構築者は自分のコンテンツをオートタギングし、仮想大陸上の適切な位置に配置してコンテンツに誘導するといったサービスが容易に提供可能になります。

2. 背景と目的

ブログによって配信される情報には、マスメディアでは取り上げられない情報や最新の話題、また商品・サービスの評判等の「消費者の生の声」が多く、ユーザのみならず、様々なWebサービス事業者からも有効な活用が望まれています。しかし、近年、ブログ記事の自動投稿ツールなどの普及により、広告を目的とした大量の記事が投稿され、単に、キーワードを含む記事を探すだけでは、利用者の関心にマッチしたブログ記事を見つけるのが困難な状況にありました。

「BLOGRANGER TG」では、最新1ヶ月間のブログから抽出された約5,000種類のタグを使ってブログ記事を適切なジャンルに自動分類するため、概念的に関心のあるジャンルのタグを選択することで、単にキーワードにマッチするだけではなく、関心のあるジャンルに絞ってブログ記事を探すことが容易になります。

例えば、「クリスマス」というキーワードでブログ記事を検索しても、料理、レストラン、インテリア、映画、コンサート等様々なジャンルでの話題がありますが、「BLOGRANGER TG」では、スクロール可能な地図として表現された仮想大陸の地形図をマウスでスクロールするだけで、様々なジャンルのクリスマスに関する記事を次々と閲覧することができます。それによりこれまでの画一的な検索結果画面で埋もれていた記事の発見が促され、楽しみながら視覚的に検索を行うことが可能になります。

3. 技術のポイント

(1) オートタギング技術

ブログには、その記事のジャンルを表現するキーワード（タグ）を設定する機能がありますが、そのキーワードはブログの著者によって自由に設定されるため、たとえば「日記」「携帯から投稿」といった内容を表わさないタグや、「アニメ」などの内容を表わすものであっても、

「アニメのこと」「アニメ日記」「最新アニメ」など多くの異なった表記が用いられます。このため、ブログに付与されているタグをそのまま利用しても必ずしも適切な記事が見つけれらるとは限りません。

NTTでは、最新の約1,000万のブログ記事からタグが付与されている約600万記事（付与されているタグは約50万種類）を分析し、ブログ記事の内容の特徴がよく表されている約5,000種類のタグを抽出しました。さらにタグが付与されている記事に含まれる特徴的な単語の統計量を抽出

し、それを基に、任意のテキストに対して高い精度で自動的にタグを付与するオートタギング技術を開発しました。「BLOGRANGER TG」では、このオートタギング機能により、収集したすべてのブログ記事に、再度タグを付与し直し、高精度の検索ができるようにしました。

(2) トピック空間可視化技術 (Topigraphy)

近年、矩形領域にタグをあいうえお順などで並べて提示するタグクラウドとよぶ情報選択のインタフェースが広く普及しています。しかし、タグクラウドでは、表示するタグの数が増えると、利用者が求めるジャンルを表わすタグを見つけることが困難になるという問題がありました。

そこでNTTは、タグ同士の概念的な距離を高精度で算出し、より関連するタグ同士が二次元平面上で近くなるように配置し、視覚的にわかりやすく検索することを可能にする「トピック空間可視化技術」を開発しました。また、タグの文字列の長さは多様であり、フォントサイズは記事数を表現しているため、タグの占める領域の面積は様々ですが、タグをお互いに重ならないように配置することもできます。

今回、タグマップの背景画像には、起伏のある地形図のデザインを導入し、多くのタグに類似する中心的な概念（例えば、「スポーツ」は「野球」、「サッカー」、「バレーボール」のいずれにも類似する中心的な概念）に高い高度（Z座標）を与えています。これにより、尾根の繋がりでトピックの繋がりが直観的に把握できます。例えば、バレーボールというタグが、マップ上ではスポーツと政治との等距離に配置された場合でも、背景画像の尾根の繋がりでスポーツに近いと表現されます。このようにトピックを地形図で表現しようとする試みは、これまでにないまったく新しい試みです。今後は、ブログ検索のみにとどまらず、様々な分野に対するこの技術の適用可能性を探っていく予定です。

4. 各社の役割と今後の予定

(1) NTT

本実験サービスに必要な技術を提供し、評価とフィードバックを行うことで、商用環境における同機能の有用性および有効性を検証します。

(2) NTTレゾナント

実験終了後の商用サービス導入に向け、利用動向や利用者のご意見を基にビジネス性を評価・検討します。

5. 利用方法

- ・「goo」トップページ (<http://www.goo.ne.jp/>) から「gooラボ」をクリック
- ⇒ 「BLOGRANGER TG」のバナーをクリック

※右記のURL (<http://ranger.labs.goo.ne.jp/>) から直接利用することも可能です。

[用語解説]

(※1) タグ

主に写真や記事などの情報コンテンツに対して、利用者が設定した内容をよく表現するキーワード。

(※2) Web API

Web上に公開している情報資源を、ネットワークを介してプログラムから利用するためのインタフェース。

- [\(別紙1\) トピック空間可視化技術](#)
- [\(別紙2\) ブログ空間の可視](#)

<本件に関する問合せ先>

日本電信電話株式会社
サイバーコミュニケーション総合研究所
企画部広報担当 川名
Tel : 046-859-2032
E-mail : randd@lab.ntt.co.jp

NTTレゾナント株式会社
ポータル事業本部 田畑
Tel : 03-5224-5500
E-mail : pr@nttr.co.jp

NTT ニュースリリース 