

A-5 ネット掲示板の議論の盛り上がりの可視化

中山 博文

ビジュアルインタフェース研究室

1. はじめに

現在、インターネットには多数の掲示板が存在し、膨大な情報がやり取りされている。ネット掲示板の議論が盛り上がっている状態(多くの人が議論に参加している状態)には、価値のある情報が含まれている可能性が高い。しかし、議論の盛り上がっている状態を探し出そうとすると、読むだけで非常に時間がかかるという問題点がある。

2. 本研究の提案

この問題点を解決するために、私はネット掲示板の議論の盛り上がりを目でわかるように可視化するソフトウェアを開発することにした。なお、本研究では掲示板の盛り上がり、単位時間あたりの書き込み数と定義する。

まず、横軸を時間、縦軸を書き込み数としたグラフによる可視化を提案する。これにより、グラフの傾きが議論の盛り上がりを表すことになり、傾きが大きければ議論が盛り上がっている状態、傾きが小さければ議論が盛り上がっていない状態であるということを表すことになる。

次に、掲示板サイト全体から1日のある一定時間における平均の書き込み数を取得し、そのデータを元に、ある一定時間における書き込み数が、平均より多いか少ないかを表すグラフによる可視化を提案する。これにより、前者の方法ではわからなかった掲示板サイト全体からみた議論の盛り上がりを表すことが可能になる。

今回は、月間利用者数930万人をほこる掲示板サイト「2ちゃんねる」[1]に対応したソフトウェアを作成した。

3. 盛り上がり可視化プログラムの作成

2ちゃんねるのスレッド(話題・トピック)の盛り上がりを可視化するプログラムを作成した。プログラムを起動し、そこに2ちゃんねるの書き込みデータであるdatファイルをドラッグアンドドロップすると図1のような画面が表示される。

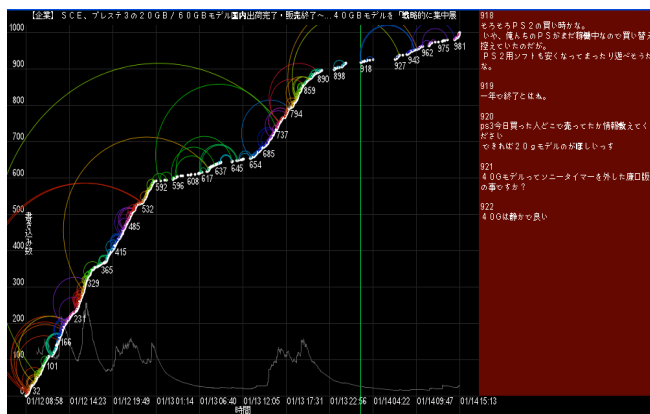


図1 可視化プログラムの実行画面

このグラフは横軸を時間、縦軸を書き込み数とし、どの時刻に何番目の書き込みがされたかがプロットされる。2つの点を結ぶ弧は、書き込みの返信関係を示す。

画面上をクリックすると、クリックされた部分の点を中心にグラフが拡大し、同時に点が示す書き込みの内容が右側に表示され、任意のキーを押すと最初のグラフ表示に戻る機能もつけた。

図1のグラフでは、スレッド開始から書き込み数590までの間と、書き込み数685から890までの間が、比較的盛り上がっていることがわかる。

4. 平均書き込み数との比較プログラムの作成

数週間にわたり、2ちゃんねるの書き込みデータを約20GB収集した。次にC言語を使用してデータファイルを板ごとに全て読み込み1スレッドあたりの10分間での平均書き込み数を計算するプログラムを作成した。

図2は、プログラムを使用し取得したある掲示板の平均データをグラフ表示した例である。このグラフでは20時から24時までの間に議論が盛り上がっていることがわかる。

図3は、図2の平均データと、ある特定のスレッドのデータを比較したグラフ表示の例である。このグラフは平均の書き込み数とスレッドの書き込み数の比を表示しており、あるスレッドが、平均の書き込みより何倍盛り上がっているか知ることができる。この図のグラフでは、スレッドが開始された直後に普段の30倍盛り上がっていることがわかる。

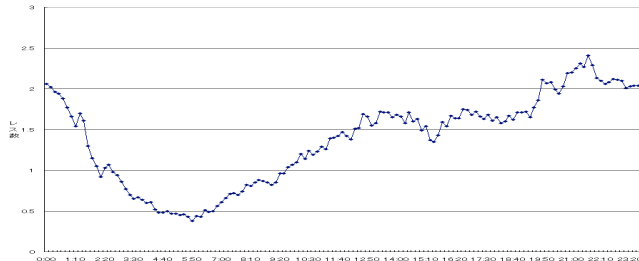


図2 サイト(板)の時刻ごとの平均書き込み数

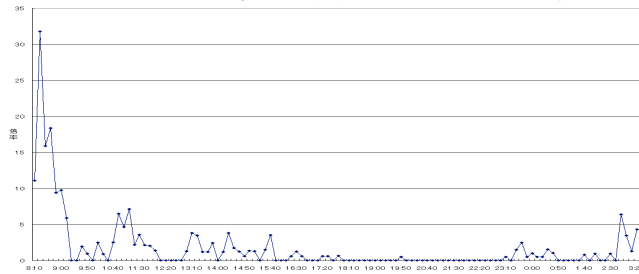


図3 平均書き込み数とスレッドの書き込み数の比

5. まとめ

実際にできあがったソフトウェアを使用してみて、どの時間帯に議論が盛り上がっているのかが一目でわかるようになった。今後、議論の盛り上がっている状態に価値のある情報が確実に含まれていることが検証されれば、本研究がより有益であったと言えるのではなからうか。

参考文献

[1] 2ちゃんねる, <http://www.2ch.net>, 1999年～。