

Twitter におけるフォロー推奨システム

研究系卒研

0821068 黒柳 智士

(指導教員 速水 治夫 教授)

1. はじめに

Twitter とは^[1]2006 年に Twitter 社が始めたサービスであり、マイクロブログあるいはミニブログと呼ばれる。

他のユーザをフォローすることで、会話のようなコミュニケーションを行うことができるが、どのようなユーザをフォローすればいいか迷うことがある。Twitter そこで本研究では会話数、類似度、キーワード情報の提示によってフォロー推奨ユーザを提案する。

2. 提案システム

本システムは Apache, MySQL, PHP, TwitterAPI^[2], MeCab を用いた Web アプリケーションである。

機能として、ユーザ認証部、推奨ユーザ探索・分析部、推奨ユーザ表示部などがある。以下に各部詳細を示す。本システムの画面遷移図を図 1 に示す。

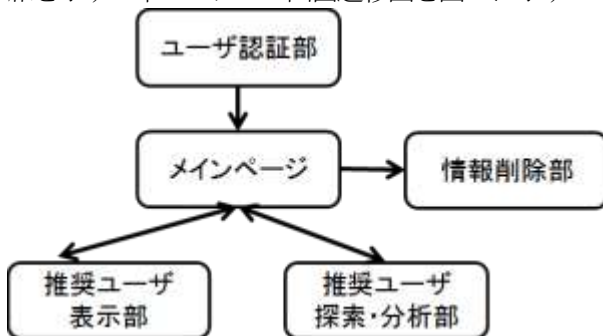


図 1 画面遷移図

2.1 ユーザ認証部

利用者の Twitter アカウントと本システムを連携させる処理をここで行う。連携することで本システムの利用が可能となる。

2.2 推奨ユーザ探索・分析部

推奨ユーザを探索する処理を行う。

処理の流れは、初めにシステム利用者の Tweet を取得し、Tweet から返信情報の取得・ハッシュタグの保存、MeCab での形態素解析を行い Wikipedia ワード・はてなキーワードをユーザと関連付けて保存する、その後、どのユーザにどのくらい返信があるか確認する。その後、返信が一番多いユーザから Tweet を取得し、システム利用者と同様の処理を行う、以上の処理を利用者が指定した階層数、ユーザに対する上限数を指定した分だけ処理を行う。

2.3 推奨ユーザ表示部

探索・分析部での処理終了後、推奨するユーザの表示を行う、以下に表示の一例を図 2 に示す。

2次元の散布図表記とし、縦軸をカテゴリ、横軸をそれぞれ類似度、共通ワード数、共通ハッシュタグ数、円のサイズを前の階層のユーザとの会話数の量による分類を行う。

3. 評価実験

Twitter 利用者 6 名に実際にシステムを利用してもらい、アンケートを実施した。以下、表 1 に結果を示す。

表 1 評価実験アンケート結果

質 問	低 → 高					平均
	1	2	3	4	5	
類似度が近いほどおススメ度が上だったか?		1	1	4		3.50
カテゴリが近いほどおススメ度が上だったか?		1	2	2	1	3.50
自分が思ってもいなかった意外なユーザは見つかりましたか?				3	3	4.50
共通ワードが多いほどおススメ度が上だったか?	1	1	1	2	1	3.17
共通ハッシュタグが多いほどおススメ度が上だったか?		2	1	1	1	3.20
提示されたユーザーのツイート内に含まれていたワードは有用だったか?		1	2	1	2	3.67
提示されたユーザーのツイート内に含まれていたハッシュタグは有用だったか?	1	1	2		1	2.80

- ハッシュタグを利用していなかったためハッシュタグの評価が出来なかった。
- 既存のサービスで表示されないユーザが出てきた
- 提示ユーザの tweet ワードが多すぎる

4. おわりに

評価実験の結果より、全体的な評価は低かった、これはアンケートより、会話が多い事は必ずしもユーザを推奨するに結びくが確認することができたが、意外なユーザが推奨されたユーザとして出現されたことに高評価を頂けた。

今後は類似度の評価や適切なワードの取捨選択による情報の提示などが挙げられる。

参考文献

- [1] Twitter
<http://twitter.com>
- [2] 山本裕介：TwitterAPI ポケットリファレンス (2011) 技術評論社