

移動エントロピーによる  
動的ネットワーク化を用いた  
SNSと商品購買の相互関係の分析

尾崎研究室 天神 雄貴

# Twitterの投稿について

Twitter上では、気軽につぶやくことができるため、

「新しい服買ったー」

「ディズニーランド楽しかったー」

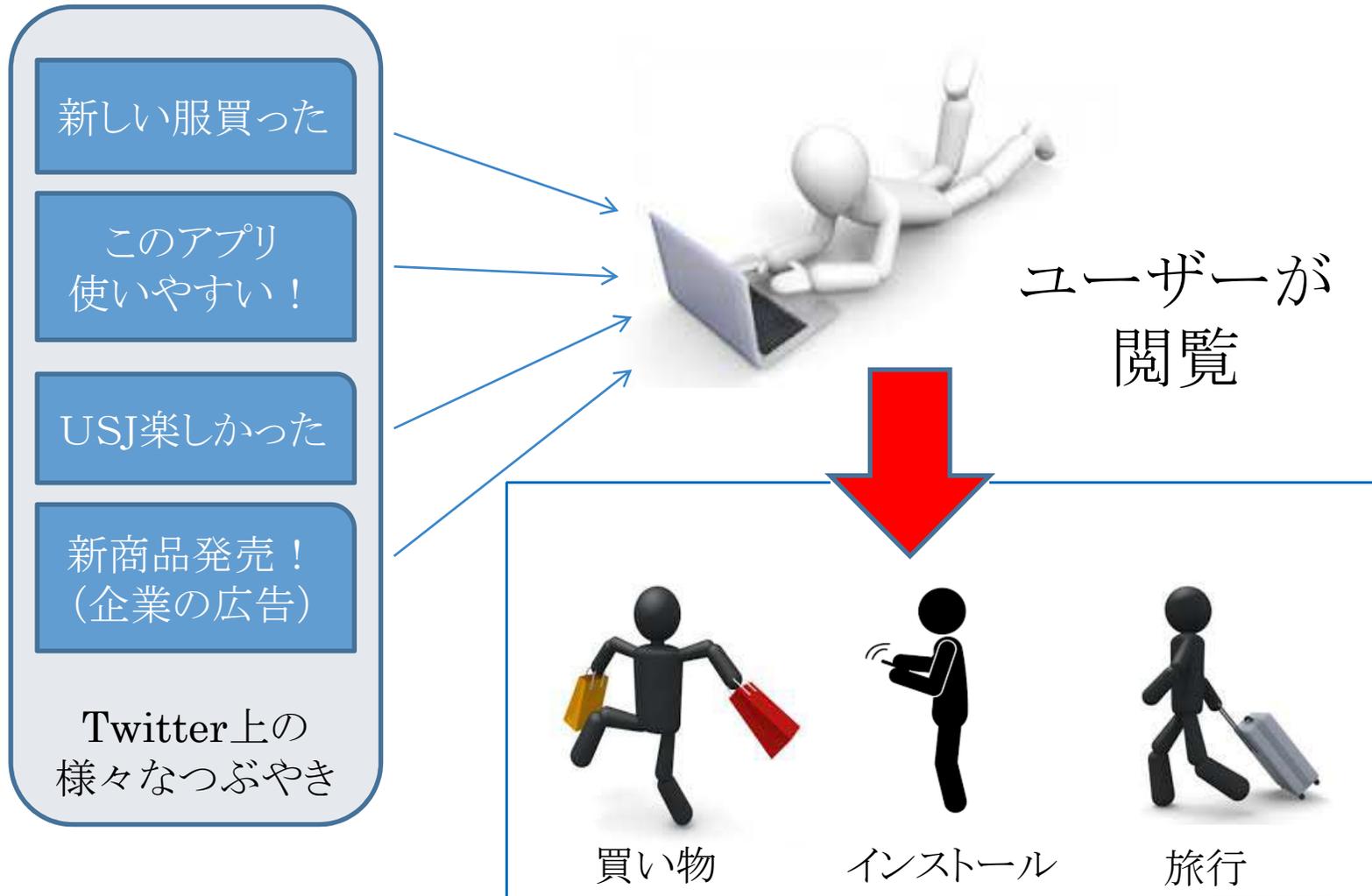
「ハワイ行ってきたー」

などユーザーは様々な投稿をする。



The image shows a screenshot of the Twitter web interface. At the top left is the 'twitter' logo. To the right are navigation links: 'ホーム', 'プロフィール', and '友'. Below the header is a tweet form. The text input field contains 'いまなにしてる?' followed by a character count '93'. Below the text is a link: '[hatena]はてなブックマークのサイト @ http://b.hatena.ne.jp/'. At the bottom left of the form, it says '最新のつぶやき: ほわっ 5秒以内前'. At the bottom right is a button labeled '投稿する'. A blue circle is visible on the right side of the page.

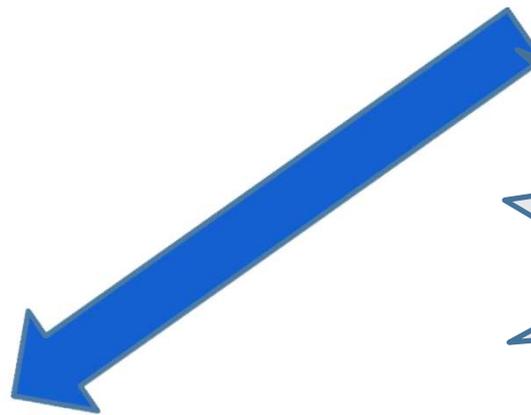
# つぶやきを見たユーザーへの影響



Twitter  
広告つぶやき



ユーザーが  
閲覧



買い物



インストール

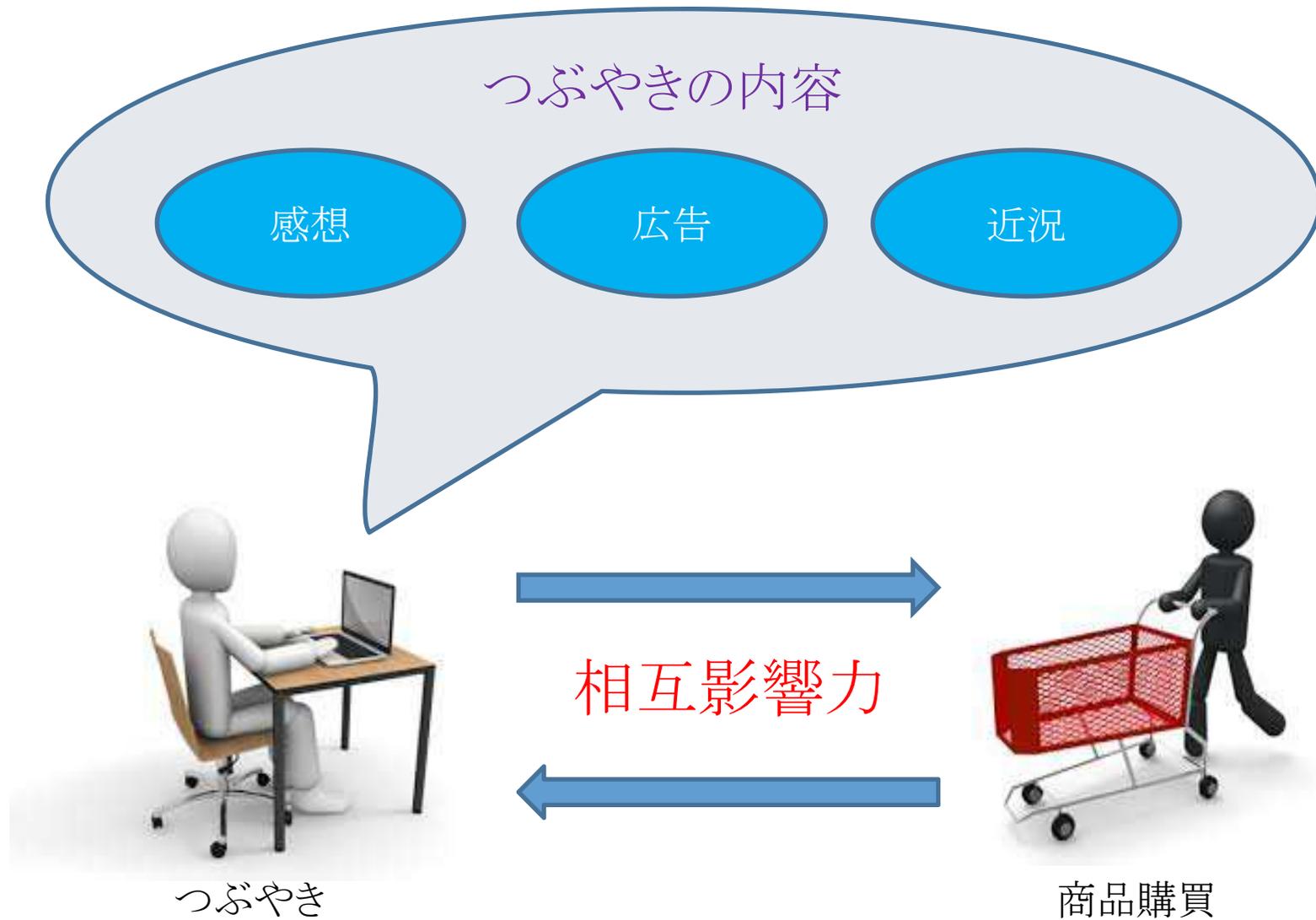


旅行



感想を投稿

# 本研究の目的

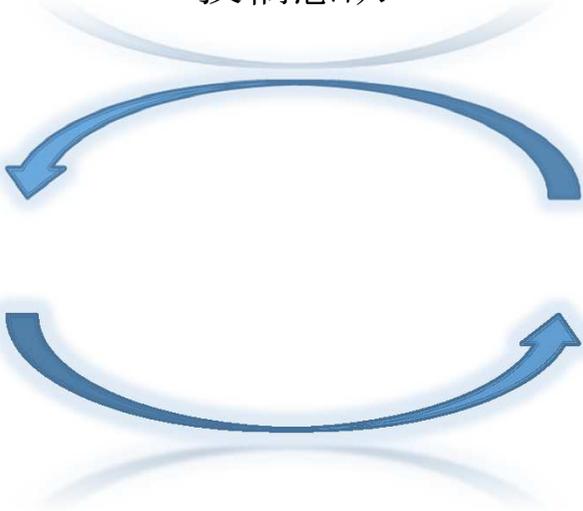


# 分析の概要

双方向の  
影響力を調べる



投稿意欲



Twitter上の単語



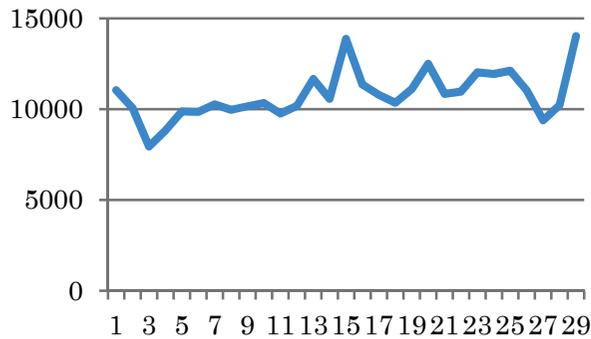
商品



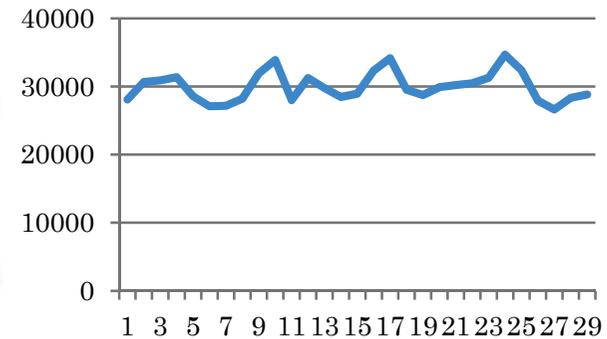
# 相互影響力の定量化

$$T_{y \rightarrow x} = \sum p(x_n, x_{n-1}^{(n-k)}, y_{n-1}^{(n-l)}) \log \frac{p(x_n | x_{n-1}^{(n-k)}, y_{n-1}^{(n-l)})}{p(x_n | x_{n-1}^{(n-k)})}$$

### Twitter単語出現回数



### 商品の売り上げ



相互影響度  
は??

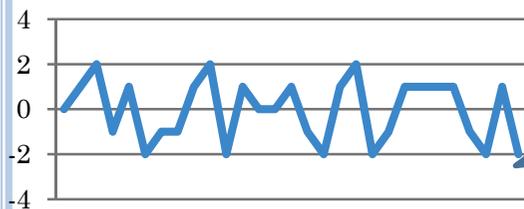
時系列の  
離散化



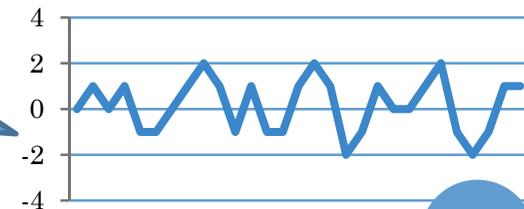
時系列の  
離散化



### Twitter離散時系列



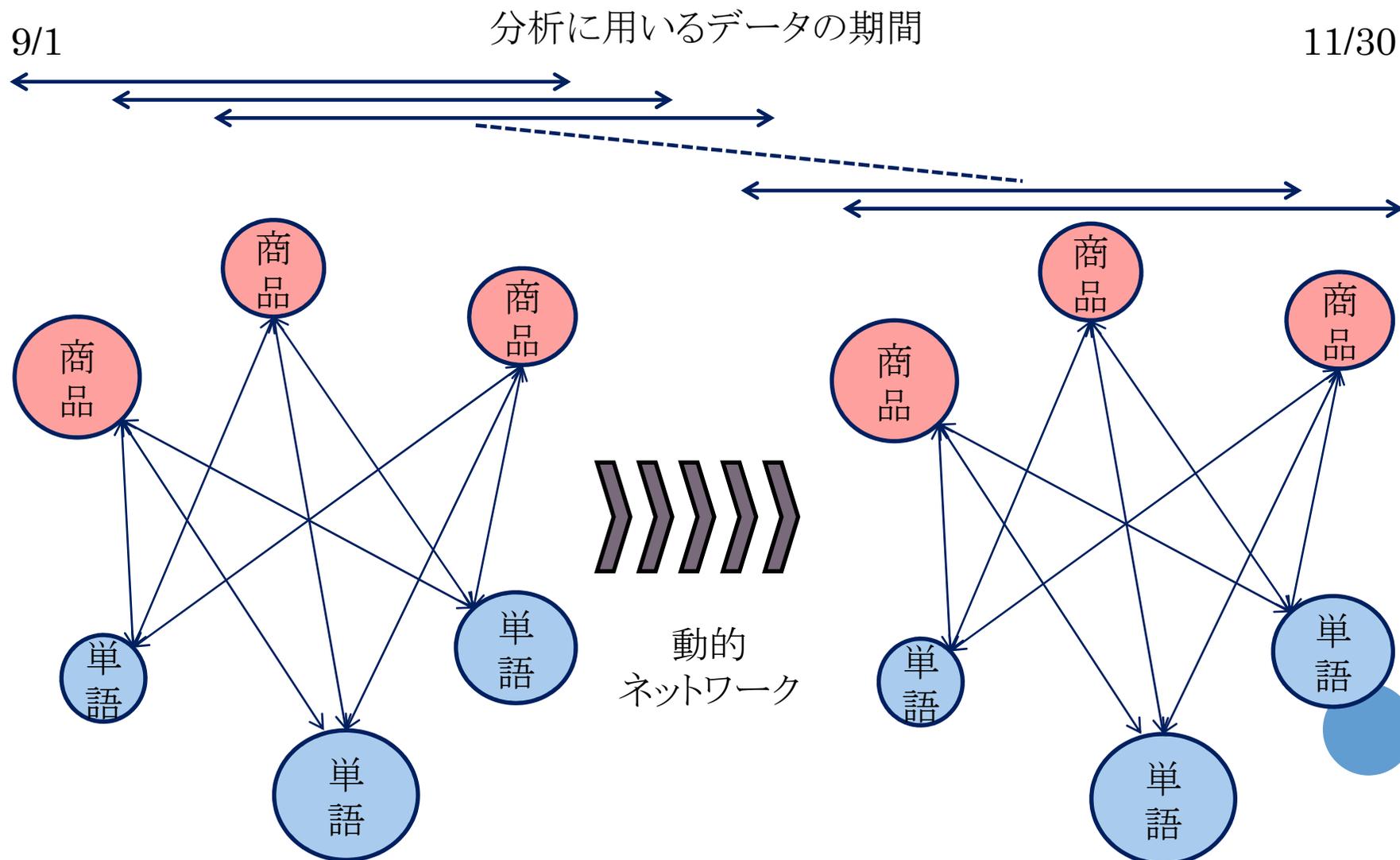
### 売り上げ離散時系列



離散化することで、  
移動エントロピーに  
よる定量化が可能



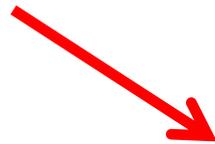
# 動的ネットワークの構築



# 分析に用いるデータ(楽天市場ランキング)

楽天市場ランキングAPIを用いて、  
「メンズファッションランキング」を  
毎日取得する。

ランキングは、1~1000位を取得  
ランキング中の商品に対して、  
ランキングの変化を時系列に



楽天ランキング時系列

商品名	9/1	9/2	9/3	.....
セーター	1位	2位	2位	.....
スウェット	35位	45位	44位	.....
ジャケット	124位	129位	250位	.....



# 分析に用いるデータ(Twitter)

「今日、新しいセーターを買った。」



形態素解析を用いて  
名詞、形容詞のみを抽出

「今日」、「新しい」、「セーター」

## 単語の出現回数の時系列化

単語	9/1	9/2	9/3	.....
今日	356回	380回	346回	.....
新しい	524回	534回	531回	.....
セーター	186回	187回	193回	.....



期間中のツイートに現れた  
すべての名詞、形容詞に対して、  
時系列を作成する



# 分析に用いるデータ(Twitter)

「感情表現辞典」を用いて、ツイート中の形態素解析した単語を下の表の10個の感情のどれか1個に結び付ける。

例: めでたい→喜  
てれる→恥 など

各感情に対応する単語が1日に出現した回数を時系列とする。

	喜	怒	哀	恐	恥	好	厭	昂	安	驚
登録 単語数	268 語	217 語	246 語	163 語	65 語	217 語	516 語	159 語	109 語	135 語



感情語グループ時系列

単語	9/1	9/2	9/3	.....
驚	11718回	11045回	10063回	.....
恐	18738回	17097回	15143回	.....
.....				



# PageRank

	商品、単語 間ネットワーク			商品、感情語 間ネットワーク		
期間	1位	2位	3位	1位	2位	3位
9/1~9/30	かわいかつ	エア	おお	驚	怒	ラルフローレン
9/11~10/10	ケーブルニット	チノパンツ	カジュアルシャツ	恥	驚	スーツ
9/21~10/20	カーディガン	チノパンツ	デイリー	恥	ワークパンツ	カーディガン
10/1~10/31	チノパンツ	岡山	理	恥	驚	怒
10/11~11/10	いところ	たこ焼き	神谷	恥	パーカー	テーパード
10/21~11/20	チノパンツ	おいしかつ	スキニーパンツ	恥	驚	キネティクス
11/1~11/30	ケーブルニット	期待	ダッフルコート	驚	恥	ジャケット

# Jaccard類似度 (楽天ノード) $\frac{A \cap B}{A \cup B}$

期間	9/11~ 10/10	9/21~ 10/20	10/1~ 10/31	10/11~ 11/10	10/21~ 11/20	11/1~ 11/30
9/1~9/30	0.71	0.55	0.46	0.39	0.35	0.32
9/11~10/10		0.76	0.61	0.49	0.42	0.36
9/21~10/20			0.76	0.57	0.50	0.40
10/1~10/31				0.71	0.56	0.44
10/11~11/10					0.75	0.56
10/21~11/20						0.70

# Jaccard類似度 (Twitterノード)

$$\frac{A \cap B}{A \cup B}$$

期間	9/11~ 10/10	9/21~ 10/20	10/1~ 10/31	10/11~ 11/10	10/21~ 11/20	11/1~ 11/30
9/1~9/30	0.52	0.48	0.40	0.39	0.39	0.37
9/11~10/10		0.55	0.46	0.43	0.43	0.41
9/21~10/20			0.54	0.48	0.46	0.44
10/1~10/31				0.56	0.49	0.43
10/11~11/10					0.55	0.44
10/21~11/20						0.51