

Why You Follow: A Classification Scheme for Twitter Follow Links

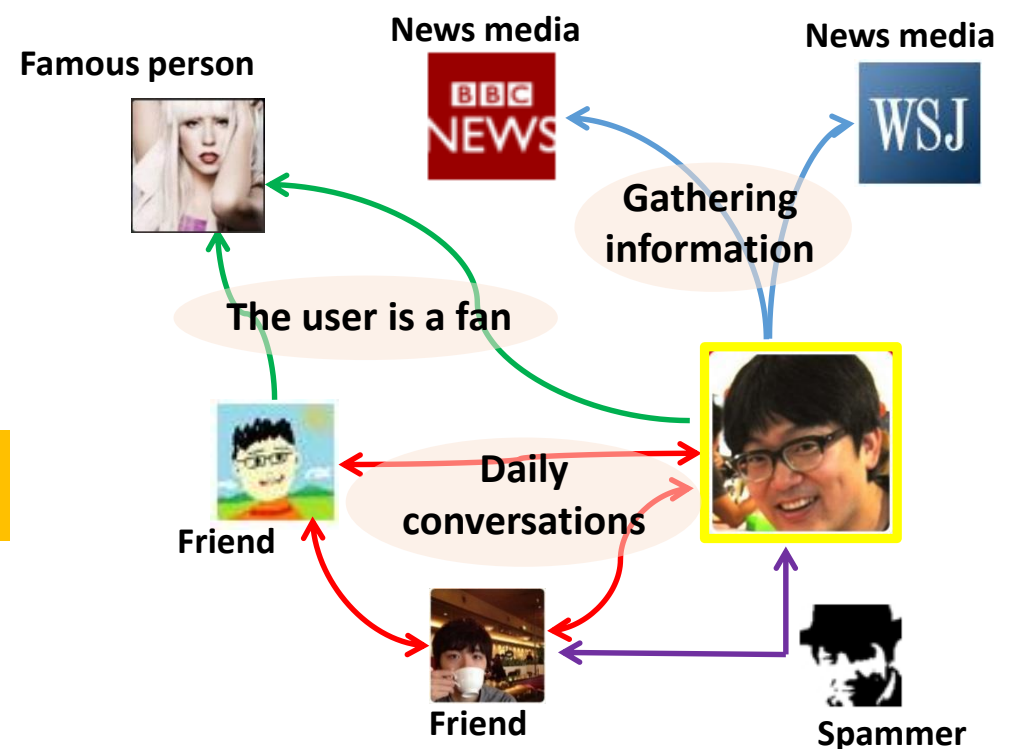
Atsushi Tanaka
a.tanaka@mbs.co.jp

Hikaru Takemura
takemura@dl.kuis.kyoto-u.ac.jp

Keishi Tajima
tajima@i.kyoto-u.ac.jp

Various types of follow links in Twitter:

- links for information gathering,
- links to celebrities from their fans,
- links for daily conversation with friends,
- links for participating public discussions,
- links by spammers, ... and more.

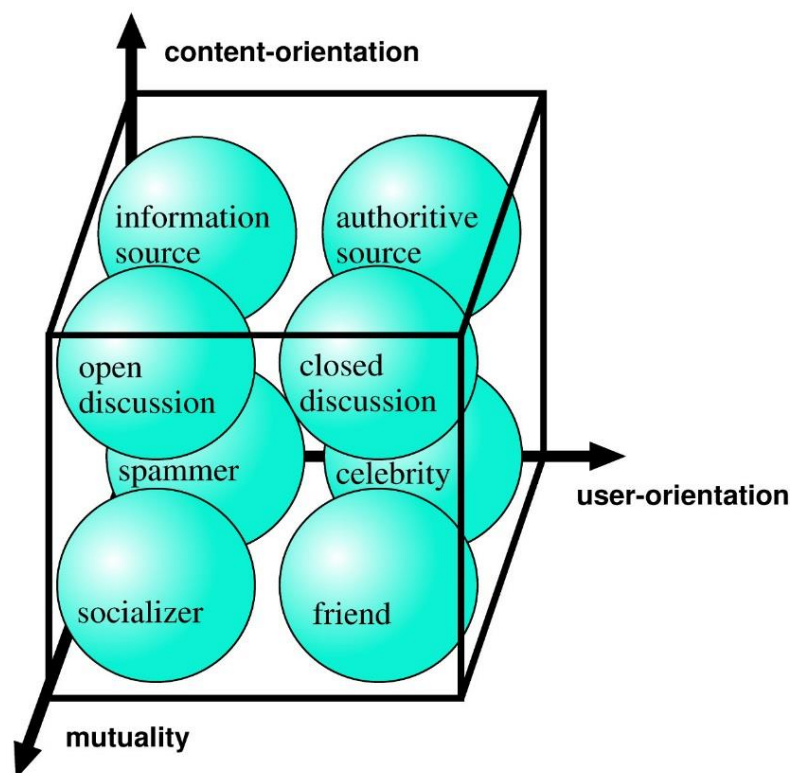


We propose classification axes for follow links

Requirements for good classification axes:

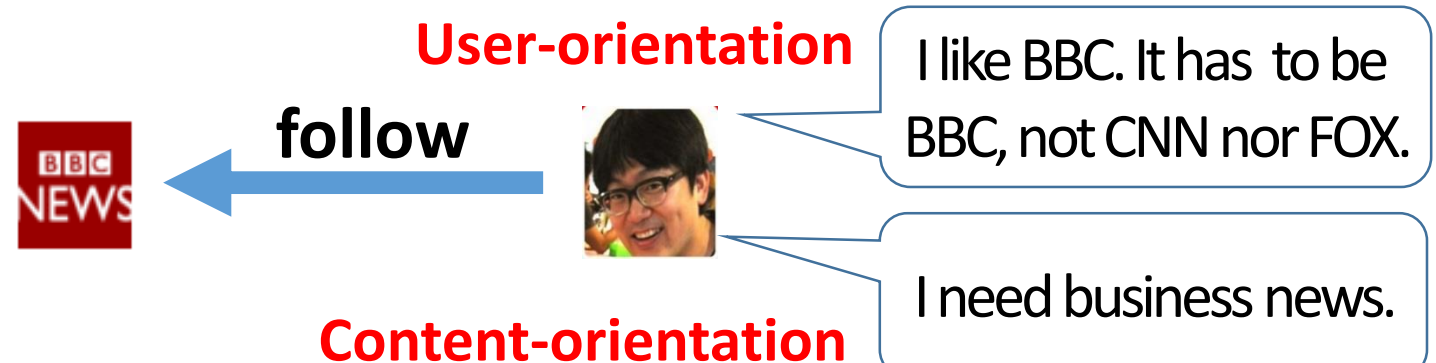
1. primitive
2. have clear intuitive meanings
3. independent from one another, and
4. can classify most typical types of follow links.

Our three axes for follow link classification



1. **User-orientation:** the follower is interested in the followee user itself, and it cannot be replaced with another user with very similar tweets.
2. **Content-orientation:** the follower is interested in specific topic, and no reason to follow the followee if she/he stops tweeting about the topic.
3. **Mutuality:** the follower expects to have mutual communication.

User-orientation and Content-orientation are not exclusive



Experiment

Dataset

1253 links collected by a questionnaire to 44 Twitter users on a crowdsourcing service.

mutuality = 0	user		
	1	0	
content	1	410	244
	0	149	232

mutuality = 1	user		
	1	0	
content	1	126	23
	0	40	29

Classifier Construction

1. SVM
2. decision tree

×

1. 3 binary classifiers
2. single 8-class classifier

=

4 approaches

Classification Features

category	features
(A) followee	ratio of information lists/community lists including the followee, those listed in (B)
(B) follower	number of followees, followers, reciprocal follows, lists, reciprocal follower ratio, reciprocal followee ratio
(C) relation	number of lists including both followee and follower, reciprocity, frequency of replies, frequency of RT

Conclusion

Result

features used	A	B	C	A+B	A+C	B+C	A+B+C	majority
3 binary SVMs combined	30.97	36.95	33.52	36.95	34.08	50.20	42.70	30.63
3 binary decision trees combined	25.30	48.36	33.52	43.58	27.29	54.91	43.26	30.63
8-class SVM	38.07	43.81	32.72	46.93	37.51	43.26	50.28	32.72
8-class decision tree	28.01	51.32	35.67	49.16	29.93	55.87	50.12	32.72

1. Our axes are independent.
2. The type of a link does not solely depend on the follower nor the followee.