



















ings of other normal users. As a result, using this approach would potentially encourage users (especially the social capitalists) to show more discretion when establishing follow links, and thus prevent spammers from easily farming links in Twitter.

## 6. CONCLUSION

As Twitter emerges as a popular platform for sharing real-time information on the Web, it has become a target for spammers, who try to infiltrate its social network, gain influence, and promote their tweets by acquiring (farming) follower links. In this paper, we first investigated link farming activity in Twitter and then proposed approaches to deter the activity. Our analysis of link farming resulted in a surprising finding: a small number of legitimate, popular, and highly active Twitter users account for a majority of the link farming activity. These elite users unwittingly resort to link farming as they seek to amass social capital by indiscriminately following back any user who follows them. Spammers exploit their behavior to gain followers and reputation in the network. To discourage social capitalists from connecting to unknown users, we proposed a ranking scheme, where users are penalized for following spammers. Our evaluation shows that our ranking scheme effectively lowers the influence of spammers and their followers in the network.

**Acknowledgment** We thank the anonymous reviewers whose suggestions helped to improve the paper. This research was supported in part by a grant from the Indo-German Max Planck Centre for Computer Science (IMPECS).

## 7. REFERENCES

- [1] bitly blog - Spam and Malware Protection. <http://tinyurl.com/nv2oer>.
- [2] Klout | The Standard for Influence. <http://klout.com/home>.
- [3] There Are Now 155m Tweets Posted Per Day, Triple the Number a Year Ago. <http://rww.to/gv4VqA>, April 2011.
- [4] Twitter help center: The Twitter rules. <http://tinyurl.com/22obg56>, 2011.
- [5] L. Becchetti, C. Castillo, D. Donato, R. Baeza-Yates, and S. Leonardi. Link analysis for web spam detection. *ACM Transactions on the Web*, 2:1–42, March 2008.
- [6] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida. Detecting spammers on Twitter. In *Annual Collaboration, Electronic Messaging, Anti-Abuse and Spam Conference (CEAS)*, 2010.
- [7] K. Bharat and M. R. Henzinger. Improved algorithms for topic distillation in a hyperlinked environment. In *ACM Int'l Conference on Research and Development in Information Retrieval (SIGIR)*, 1998.
- [8] C. Castillo, D. Donato, A. Gionis, V. Murdock, and F. Silvestri. Know your neighbors: web spam detection using the web topology. In *ACM Int'l Conference on Research and Development in Information Retrieval (SIGIR)*, 2007.
- [9] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi. Measuring user influence in Twitter: the million follower fallacy. In *AAAI Int'l Conference on Weblogs and Social Media (ICWSM)*, 2010.
- [10] S. Chakrabarti. Integrating the document object model with hyperlinks for enhanced topic distillation and information extraction. In *ACM Int'l Conference on World Wide Web (WWW)*, 2001.
- [11] B. D. Davison. Recognizing nepotistic links on the web. In *AAAI Workshop on Artificial Intelligence for Web Search*, 2000.
- [12] D. Gayo-Avello and D. J. Brenes. Overcoming Spammers in Twitter - a tale of five algorithms. In *Spanish Conference on Information Retrieval (CERI)*, 2010.
- [13] C. Grier, K. Thomas, V. Paxson, and M. Zhang. @spam: the underground on 140 characters or less. In *ACM Int'l Conference on Computer and Communications Security (CCS)*, 2010.
- [14] Z. Gyöngyi and H. Garcia-Molina. Link spam alliances. In *Int'l Conference on Very Large Data Bases (VLDB)*, 2005.
- [15] Z. Gyöngyi, H. Garcia-Molina, and J. Pedersen. Combating web spam with trustrank. In *Int'l Conference on Very Large Data Bases (VLDB)*, 2004.
- [16] T. H. Haveliwala. Topic-sensitive pagerank. In *ACM Int'l Conference on World Wide Web (WWW)*, 2002.
- [17] US confirms it asked Twitter to stay open to help Iran protesters. <http://tinyurl.com/klv36p>.
- [18] H. Kwak, H. Chun, and S. Moon. Fragile online relationship: a first look at unfollow dynamics in Twitter. In *Annual Conference on Human Factors in Computing Systems (CHI)*, 2011.
- [19] K. Lee, J. Caverlee, and S. Webb. Uncovering social spammers: social honeypots + machine learning. In *ACM Int'l Conference on Research and Development in Information Retrieval (SIGIR)*, 2010.
- [20] K. Lee, B. D. Eoff, and J. Caverlee. Seven months with the devils: a long-term study of content polluters on Twitter. In *AAAI Int'l Conference on Weblogs and Social Media (ICWSM)*, 2011.
- [21] R. Lempel and S. Moran. The stochastic approach for link-structure analysis (SALSA) and the TKC effect. *Computer Networks*, 33:387–401, Jun 2000.
- [22] A. Ramachandran and N. Feamster. Understanding the network-level behavior of spammers. *SIGCOMM Computer Communication Review*, 36:291–302, Aug 2006.
- [23] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *ACM Int'l Conference on World Wide Web (WWW)*, 2010.
- [24] M. Sobek. Google PageRank - PR 0. <http://pr.efactory.de/e-pr0.shtml>.
- [25] D. Talbot. How Google Ranks Tweets. <http://www.technologyreview.in/web/24353/>.
- [26] J. Teevan, D. Ramage, and M. R. Morris. #TwitterSearch: a comparison of microblog search and web search. In *ACM Int'l Conference on Web Search and Data Mining (WSDM)*, 2011.
- [27] K. Thomas, C. Grier, V. Paxson, and D. Song. Suspended accounts in retrospect: an analysis of Twitter spam. In *ACM SIGCOMM Conference on Internet Measurement (IMC)*, 2011.
- [28] L. Rao, Twitter Seeing 90 Million Tweets Per Day, 25 Percent Contain Links, *TechCrunch*, 2010. <http://tinyurl.com/27x5cay>.
- [29] J. Weng, E.-P. Lim, J. Jiang, and Q. He. TwitterRank: finding topic-sensitive influential Twitterers. In *ACM Int'l Conference on Web Search and Data Mining (WSDM)*, 2010.
- [30] B. Wu and B. D. Davison. Identifying link farm spam pages. In *ACM Int'l Conference on World Wide Web (WWW)*, 2005.
- [31] B. Wu, V. Goel, and B. D. Davison. Propagating trust and distrust to demote web spam. In *Workshop on Models of Trust for the Web*, 2006.
- [32] S. Yardi, D. Romero, G. Schoenebeck, and D. M. Boyd. Detecting spam in a twitter network. *First Monday*, 15(1):1–13, Jan 2010.
- [33] C. M. Zhang and V. Paxson. Detecting and analyzing automated activity on Twitter. In *Int'l Conference on Passive and Active Measurement (PAM)*, 2011.