$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}$$

For a particular user under suspicion, we find the deviation of its distribution based on the same three parameters from that of the random sample and define *suspect_ratio* as the inverse of the deviation. Therefore, smaller *suspect_ratio* indicates higher suspicion towards the perceived social reputation (follower count) of the user. We then conducted an in-the-wild experiment over 1% Twitter stream data and labeled a user as suspicious if her *suspect_ratio* was significantly low below a certain threshold. So far we have been able to label over 56,000 users as suspicious. A small sample of the users detected by our proposed methodology can be seen at `http://bit.ly/FakeFollowProj`.

6. CONCLUSION AND FUTURE WORK

This work aims to detect and measure the deviation from perceived social reputation of an OSN user. We start by landscaping the sources of such manipulation like blackmarkets and scratch-back services. Preliminary results bring out the underlying structure of blackmarket which can be helpful to uncover the market leaders. Eliminating or hindering their operations can significantly bring down crowdsourced manipulation of social reputation. Initial results also show that a robust and adaptive technique can be built to detect social reputation manipulation. However, our proposed framework is at a very nascent stage and needs much more improvement and rigorous evaluation. Much work yet remains to leverage this framework to build an alternate social reputation system and measure the effects of social reputation manipulation on OSN's ecosystem.

7. REFERENCES

[1] A. Aggarwal and P. Kumaraguru. What they do in shadows: Twitter underground follower market. In *Privacy, Security and Trust (PST)*, 2015.

[2] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida. Detecting spammers on twitter. In *Collaboration, electronic messaging, anti-abuse and spam conference (CEAS)*, volume 6, page 12, 2010.

[3] Z. Chu, I. Widjaja, and H. Wang. Detecting social spam campaigns on twitter. In *Applied Cryptography and Network Security*, pages 455–472. Springer, 2012.

[4] D. B. Clark. The bot bubble. `https://newrepublic.com/article/121551/bot-bubble-click-farms-have-inflated-social-media`, April 2015.

[5] DailyMail. More than 2 million of hillary clinton's twitter followers are fake or never tweet. `http://www.dailymail.co.uk/news/article-3038621/More-2-MILLION-Hillary-Clinton-s-Twitter-followers`, April 2015.

[6] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao. Detecting and characterizing social spam campaigns. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pages 35–47. ACM, 2010.

[7] A. Kapravelos, C. Grier, N. Chachra, C. Kruegel, G. Vigna, and V. Paxson. Hulk: Eliciting malicious behavior in browser extensions. In *Proceedings of the 23rd Usenix Security Symposium*, 2014.

[8] K. Lee, S. Webb, and H. Ge. Characterizing and automatically detecting crowdturfing in fiverr and twitter. *Social Network Analysis and Mining*, 5(1):1–16, 2015.

[9] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw. Detecting product review spammers using rating behaviors. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 939–948. ACM, 2010.

[10] Microsoft. Trojan:js/febipos.a. `www.microsoft.com/security/portal/threat/encyclopedia/entry.aspx?Name=Trojan:JS/Febipos.A`, August 2013.

[11] Microsoft. Trojan:js/kilim.a. `https://www.microsoft.com/security/portal/threat/encyclopedia/entry.aspx?Name=JS/Kilim`, June 2013.

[12] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker. Dirty jobs: The role of freelance labor in web service abuse. In *Proceedings of the 20th USENIX conference on Security*, pages 14–14. USENIX Association, 2011.

[13] NYTimes. A rave, a pan, or just a fake? `http://www.nytimes.com/2011/05/22/your-money/22haggler.html`, May 2011.

[14] NYTimes. Fake twitter followers become multimillion-dollar business. `http://bits.blogs.nytimes.com/2013/04/05/fake-twitter-followers-becomes-multimillion-dollar-business/`, April 2013.

[15] NYTimes. All the product reviews money can buy. http://www.nytimes.com/2015/12/06/your-money/all-the-product-reviews-money-can-buy.html, December 2015.

[16] G. Stringhini, G. Wang, M. Egele, C. Kruegel, G. Vigna, H. Zheng, and B. Y. Zhao. Follow the green: growth and dynamics in twitter follower markets. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 163–176. ACM, 2013.

[17] K. Thomas, D. McCoy, C. Grier, A. Kolcz, and V. Paxson. Trafficking fraudulent accounts: The role of the underground market in twitter spam and abuse. In *USENIX Security*, pages 195–210. Citeseer, 2013.

[18] B. Viswanath, M. A. Bashir, M. Crovella, S. Guha, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Towards detecting anomalous user behavior in online social networks. In *Proceedings of the 23rd USENIX Security Symposium (USENIX Security)*, 2014.

[19] B. Viswanath, M. A. Bashir, M. B. Zafar, S. Bouget, S. Guha, K. P. Gummadi, A. Kate, and A. Mislove. Strength in numbers: Robust tamper detection in crowd computations. In *Proceedings of the 2015 ACM on Conference on Online Social Networks*, pages 113–124. ACM, 2015.

[20] G. Wang, T. Wang, H. Zheng, and B. Y. Zhao. Man vs. machine: Practical adversarial detection of malicious crowdsourcing workers. In *23rd USENIX Security Symposium, USENIX Association, CA*, 2014.

[21] G. Wang, C. Wilson, X. Zhao, Y. Zhu, M. Mohanlal, H. Zheng, and B. Y. Zhao. Serf and turf: crowdturfing for fun and profit. In *Proceedings of the 21st international conference on World Wide Web*, pages 679–688. ACM, 2012.