Contrast Pattern-Based Classification for Bot Detection on Twitter

MD ASADUZZAMAN SARKER ANIK
ID: 23166031
GROUP: 12

Introduction

- Bots[1]
- Bots on twitter[2]
- Bot classifier
- Pattern extraction

Related Work

- Feature extraction
  - Tweet content
  - Tweet sentiment
  - Account information
  - Account usage
  - Social network structure
- Decorate by Lee et al[3] - 0.88
- Naive Bayes classifier by Wang [4] and Ahmed and Abulaish [5] - F1 > 0.90
- Random forest by Chu et al[6] and Yang et al. [7] - 96% and F1 -0.9
- DNA fingerprint by Cresci et al. [8] - f1 0.97

Methodology

- Creating a new feature model
    - Tweet content
    - Sentiment analysis
    - Association with other twitter accounts
- Inducing decision trees
- Selecting optimal contrast patterns
- Classifying query objects

Dataset

- 51,457 tweets
- 31,654 by Human
- 19,804 by bot

Result

- AUC of 0.90
- MCC of 0.91

Contrast Patterns

Class	ID	Items	Supp
Bot	CP_1	$[tweet\_source = \text{“}TweetAdderv4\text{”}]$	0.66
	CP_2	$[user\_friends\_count > 1054.50] \wedge [user\_favourites\_count \leq 2.50]$	0.62
	CP_3	$[tweet\_num\_mentions \leq 0.50] \wedge [user\_friends\_count > 1054.50] \wedge [user\_favourites\_count \leq 2.50]$	0.60
	CP_4	$[user\_favourites\_count \leq 10.50 \wedge [tweet\_year \leq 2014] \wedge [user\_friends\_count \in [1054.50, 2089.50]]$	0.59
	CP_5	$[user\_H16\_TweetDist > 6.50] \wedge [tweet\_year \leq 2014] \wedge [tweet\_retweet\_count \leq 0.50] \wedge [user\_favourites\_count \leq 4.50] \wedge [user\_H04\_TweetDist \leq 10.50] \wedge [tweet\_day > 1.50]$	0.50
Human	CP_6	$[tweet\_year > 2014]$	0.56
	CP_7	$[user\_followers\_count > 12.50] \wedge [user\_geo\_enabled = \text{“}1\text{”}] \wedge [user\_favourites\_count > 380.50] \wedge [user\_screen\_name \ != \text{“}\_iHATEMOON\text{”}] \wedge [user\_statuses\_count > 571.50]$	0.48
	CP_8	$[user\_geo\_enabled = \text{“}1\text{”}] \wedge [user\_favourites\_count > 380.50] \wedge [user\_profile\_text\_color \ != \text{“}F70A0A\text{”}] \wedge [user\_profile\_sidebar\_border\_color \ != \text{“}CC3366\text{”}]$	0.48
	CP_9	$[user\_geo\_enabled = \text{“}1\text{”}] \wedge [user\_time\_zone! = \text{“}EasternTime(US\&Canada)\text{”}] \wedge [user\_favourites\_count > 170.50]$	0.37
	CP_{10}	$[tweet\_num\_mentions > 0.50] \wedge [user\_descrip\_sa\_score\_tag = \text{“}NONE\text{”}] \wedge [tweet\_month \leq 5] \wedge [user\_H05\_TweetDistt \leq 1] \wedge [user\_favourites\_count > 208]$	0.26

Limitations

- Updated bots
- Language barrier
- No a general bot detector model

Future work

- Generate more high quality patterns
- Update for other social media
    - Instagram
    - Facebook
    - Google Plus

Reference

1. M. Debashi and P. Vickers, “Sonification of network traffic for detecting and learning about botnet behavior,” IEEE Access, vol. 6, pp. 33826–33839, 2018.
2. O. Varol, E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini. (2017). “Online human-bot interactions: Detection, estimation, and characteriza- tion.” [Online]. Available: https://arxiv.org/abs/1703.03107
3. K. Lee, J. Caverlee, and S. Webb, “Uncovering social spammers: Social honeypots+ machine learning,” in Proc. 33rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., 2010, pp. 435–442
4. A. H. Wang, “Detecting spam bots in online social networking sites: A machine learning approach,” in IFIP Annu. Conf. Data Appl. Secur. Privacy. Berlin, Germany: Springer, 2010, pp. 335–342.
5. F. Ahmed and M. Abulaish, “A generic statistical approach for spam detection in online social networks,” Comput. Commun., vol. 36, nos. 10–11, pp. 1120–1129, 2013.
6. Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, “Detecting automation of twitter accounts: Are You a Human, Bot, or Cyborg?” IEEE Trans. Dependable Secure Computing, vol. 9, no. 6, pp. 811–824, Nov. 2012.
7. [37] Z. Gilani, E. Kochmar, and J. Crowcroft, “Classification of twitter accounts into automated agents and human users,” in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining, Jul. 2017, pp. 489–496.
8. S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, “Dna-inspired online behavioral modeling and its application to spambot detection,” IEEE Intell. Syst., vol. 31, no. 5, pp. 58–64, Sep./Oct. 2016.