



Shaming Type	Event	Example Tweet
Whatabouterism (WA)	AK	Wifey #AamirKhan Rao wasnt scared when - AR Rahman was threatened by the Muslim Ulemas
Sarcasm/Joke (SJ)	AK	Just in..Agarwal Packers and Movers has sent a Friend Request to #AmirKhan on Facebook...
Referring to religion, ethnicity (RE)	AK	trending #IStandWithAamirKhan reflects besides pseudo secular a particular community trying to malign the sovereignty of hindustan.
Associating with negative character (AN)	TH	I liked a @YouTube video <a href="http://t.co/YpcoKEPbIu">http://t.co/YpcoKEPbIu</a> Phil Robertson Vs. Gays Vs. Justine Sacco
Abuses (AB)	TH	Better headline: "Non-Nobel winning Biologist Calls Tim Hunt a dipshit."
Passing judgment (PJ)	TH	Tim Hunt along with all his nose hair needs to lock himself in the basement and rot there.
Comparison with ideal (CI)	TH	Tim Hunt wouldn't recognize a good scientist if Marie Curie, Jane Goodall, Shirley Ann Jackson, and Sally Ride all kickâ€e
Irrelevant past tweet (IR)	JS	I had a sex dream about an autistic kid last night. #fml
False fact-ing (FF)	JS	Isn't Justine Sacco's father a billionaire business man in South Africa?
Questioning qualifications (QQ)	JS	Justine Sacco clearly knows nothing about media and PR. So how did she become a top PR executive?

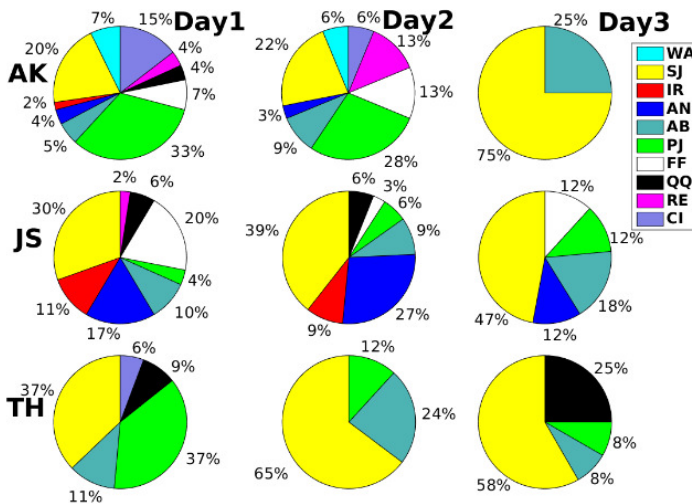


Figure 1: Shaming types for the first three days

gories - a problem that calls for more intricate use of NLP techniques and is left as future work.

## 4. DISCUSSION

Unlike rumors, whether detection and categorization of shaming tweets might be used to stop their spread is an open question as it could act as a two-edged sword - protecting the target from disproportionate punishment meted out without trial on OSN court vis-a-vis individual freedom of expression on OSN. Instead, we feel that our work can be used to study the nature of people who indulge in public shaming and determine their possible motive like one-upmanship, showing off righteousness, etc., based on past tweet history, number of followers, tendency to retweet and several other features that can be easily extracted. It can also find utility in the study of how a shaming target retaliates through his/her own tweets, be it in the form of

**Table 3: Significant features with mean and p-values.** HT: No. of hashtags, URL: urls, NNP: proper noun, PRP: personal pronoun, PRP\$: possessive pronoun, VBG: verb present participle, WRB: "wh" adverbs, SC: status, FLC: follower, FVC: favorited count

Feature	Non-Shaming Mean	Shaming Mean	p value
HT	0.41	0.50	0.06
URL	0.64	0.30	<0.001
NNP	3.71	3.42	0.03
PRP	0.55	0.85	<0.001
PRP\$	0.22	0.28	0.05
VBG	0.24	0.44	<0.001
WRB	0.10	0.15	0.02
SC	$3.81 \times 10^4$	$2.66 \times 10^4$	0.12
FLC	$1.40 \times 10^5$	$0.5 \times 10^5$	0.15
FVC	$2.86 \times 10^3$	$5.20 \times 10^3$	0.01

apologies or by direct confrontation. All these are challenging computational problems that we plan to work on.

## 5. REFERENCES

- [1] T. Takahashi and N. Igata. Rumor detection on Twitter. In *6th International Joint Conference on SCIS and ISIS*, pages 452–457. IEEE, 2012.
- [2] Z. Zhao, P. Resnick and Q. Mei. Enquiring minds: Early detection of rumors in social media from enquiry posts. In *24th International Conference on World Wide Web*, pages 1395–1405, 2015.
- [3] J. Ronson. *So You've Been Publicly Shamed*. Picador, 2015.
- [4] Y R Tausczik and J W Pennebaker. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1):24–54, 2010.