CONCLUSION

In this paper, we have proposed a hybrid approach exploiting community-based features with *metadata-*, *content-*, and *interaction-based* features for detecting automated spammers in Twitter. Spammers are generally planted in OSNs for varied purposes, but absence of real-life identity hinders them to join the trust network of benign users. Therefore, spammers randomly follow a number of users, but rarely followed back by them, which results in low edge density among their *followers* and *followings*. This type of spammers interaction pattern can be exploited for the development of effective spammers detection systems. Unlike existing approaches of characterizing spammers based on their own profiles, the novelty of the proposed approach lies in the characterization of a spammer based on its neighboring nodes (especially, the followers) and their interaction network. This is mainly due to the fact that users can evade features that are related to their own activities,

but it is difficult to evade those that are based on their followers. On analysis, metadata-based features are found to be least effective as they can be easily evaded by the sophisticated spammers by using random number generator algorithms. On the other hand, both interaction- and community-based features are found to be the most discriminative for spammers detection.

Attaining perfect accuracy in spammers detection is extremely difficult, and accordingly any feature set can never be considered as complete and sound, as spammers keep on changing their operating behavior to evade detection mechanism.

Therefore, in addition to profile-based characterization,complete logs of spammers starting from their entry in the network to their detection, need to be analyzed to model the evolutionary behavior and phases of the life-cycles of spammers. But, generally spammers are detected when they are at very advanced stage, and it is difficult to get their past logs data. Moreover, it may happen that a user is operative in the network as a benign user, and later on, it start sillicit activities due to whatsoever reasons, and considered as spammer. In this circumstance, even analyzing log data may lead to wrong characterization.

Analysis of spammers network to unearth different types of coordinated spam campaigns run by the spambots seems one of the promising future directions of research. Moreover, analyzing the temporal evolution of spammers’ followers may reveal some interesting patterns that can be utilized for spammers characterization at different levels of granularity.