Multiple identities (sockpuppets) detection is a critical issue in social media. Previous studies (Kumar et al., 2017; Wang et al., 2018) have shown that sockpuppets are often applied to manipulate public opinion and heavily affect the user experience.

Recent studies focus on profile similarity based (Keselj et al., 2003; Li et al., 2017), verbal based (Solorio, Hasan, and Mizan, 2013) and behavior features based (Tiskerdekis and Zeadally, 2014) methods to detect the sockpuppets. Profile similarity based methods use the account profile, such as nickname (Keselj et al., 2003) and description of account (Li et al., 2017), for detection. Verbal based methods assume that sockpuppets have similar linguistic traits (Kumar et al., 2017) and extract text and semantic features from posts and comments. However, smart malicious users could change nicknames and writing style to disguise themselves.

Thus, behavior feature based methods are proposed to analyze account behaviors, such as the total number of bytes removed from all the revisions (Tiskerdekis et al., 2014), co-occurrence relationship (Liu et al., 2016). However, previous methods have high dependence on the particular platforms, thus lack generalization. For example, as a video platform, TikTok can hardly provide useful text to extract verbal features. In this paper, we propose a platform-independent interaction entropy method based on the social interaction activity. The method can be easily applied to many social media platforms.

The contribution of this work can be concluded as follows: (1) We apply the interaction activities, which are available on mainstream social platforms, to extract general features that are easily available from many platforms. (2) We propose a novel interaction entropy algorithm to depict the interaction activities of users for sockpuppets detection. (3) The experimental results achieve a significant performance over state-of-the-art methods and validate the effectiveness of our method.

Methodology

We extract the interaction activity from interaction logs and construct interaction tree as shown in Figure 1. Then we propose the interaction entropy based on the interaction tree inspired by the ideal of the graph entropy (Dehmer, 2008). Firstly, we define user interaction set $D_0 = \{d_u^i | u \in U, i \in M\}$, where $d_u^i$ is the interaction tree of account $u$ based on message $i$, $U$ is the set of users and $M$ denotes the messages in the social network. The function $|S(v, d_u^i)|$ denotes the depth of node $v$ in $d_u^i$ as shown in Figure 2. For a node $v \in d_u^i$, we define the node entropy function:

$$f(v) = |S(v, d_u^i)| \cdot \frac{C(v)}{C(d_u^i)}$$

(1)

where $C(v)$ is the number of account $v$ in $d_u^i$ (an account may appear many times in one tree) and $C(d_u^i)$ is the number of nodes in the tree $d_u^i$. Then we propose an function to represent the entropy of the interaction tree $d_u^i$.

$$H_p(d_u^i) = \sum_{i=0}^{C(d_u^i)} p(v, d_u^i) \cdot \log p(v, d_u^i),$$

(2)

where $p(v, d_u^i) = \frac{f(v)}{\sum_{i=0}^{C(d_u^i)} f(v)}$ is the probability mass function for the tree. The hyperparameters $\alpha$ is a shrinkage factor. According to the above definition, a tree has more diverse interactions when it has a lower entropy value.
Table 1: Multiple Identity Deception Detection

| Method                              | \( D_S \) | \( D_T \) |
|-------------------------------------|------------|------------|
|                                    | Precision  | Recall     | F1-Score   | Precision  | Recall     | F1-Score   |
| Verbal Based (Solorio et al., 2013) | 0.755      | 0.580      | 0.645      | 0.701      | 0.696      | 0.697      |
| Behavior Features Based (Tsikerdekis et al., 2014) | 0.700      | 0.564      | 0.622      | 0.766      | 0.765      | 0.764      |
| Behavior Features Based (Liu et al., 2016) | 0.680      | 0.700      | 0.690      | 0.760      | 0.580      | 0.660      |
| Profile Similarity Based (Li et al., 2017) | 0.746      | 0.686      | 0.707      | 0.710      | 0.672      | 0.690      |
| Behavior Based (Wang et al., 2018) | 0.840      | 0.870      | 0.850      | 0.830      | 0.800      | 0.820      |
| Our Method                          | 0.860      | 0.890      | 0.875      | 0.900      | 0.894      | 0.897      |

\( D_T \) dataset contains 991 users. Accounts are identified as sockpuppets according to rules, such as self-reported sentence matching like 'This is a sockpuppet of Mix'.

Experiments

In the experiments, we set the model parameters as: \( \alpha = 2.3, \beta = 0.6 \). The analyses of parameter sensitivity can be found in the supplemental materials. As shown in the Table 1, our model achieves significantly improvement over state-of-the-art methods on two datasets (+2.5-30.5% in terms of F1), which shows the effectiveness of the model. Furthermore, the interaction relations are easily available on mainstream social media platforms, such as Twitter, Facebook, Weibo, etc., thus the algorithm has wider applications compared with previous methods.

Conclusion

In this paper, we study the multiple identities deception detection problem, with the platform independent model based on social interaction structure. Then we raise a novel interaction entropy algorithm to depict the interaction activity. The experimental results validate the effectiveness of our method on two real-world datasets.

References

Dehmer, M. 2008. Information processing in complex networks: Graph entropy and information functionals. Applied Mathematics and Computation 201(1-2):82–94.

Keselj, V.; Peng, F.; Cercone, N.; and Thomas, C. 2003. N-gram-based author profiles for authorship attribution. In PAACLING’03, 255–264.

Kumar, S.; Cheng, J.; Leskovec, J.; and Subrahmanian, V. S. 2017. An army of me: Sockpuppets in online discussion communities. In WWW ’17, 857–866.

Li, Y.; Peng, Y.; Ji, W.; Zhang, Z.; and Xu, Q. 2017. User identification based on display names across online social networks. IEEE Access 5:17342–17353.

Liu, D.; Wu, Q.; Han, W.; and Zhou, B. 2016. Sockpuppet gang detection on social media sites. Frontiers Comput. Sci. 10(1):124–135.

Solorio, T.; Hasan, R.; and Mizan, M. 2013. A case study of sockpuppet detection in wikipedia. In Proceedings of the Workshop on Language Analysis in Social Media, 59–68.

Tsikerdekis, M., and Zeadally, S. 2014. Multiple account identity deception detection in social media using nonverbal behavior. IEEE Trans. Information Forensics and Security 9(8):1311–1321.

Wang, J.; Zhou, W.; Li, J.; Yan, Z.; Han, J.; and Hu, S. 2018. An online sockpuppet detection method based on subgraph similarity matching. In ISPA ’18, 391–398.

Datasets

In this paper, we conduct experiments on two real-world datasets \( D_S \) and \( D_T \) which are public available at (Wang et al., 2018). The \( D_S \) dataset contains 675 users. The

Experiments and Result