Integrating Semantic, Social, and Spatial Dimensions for Inductive Malicious User Detection in Social Networks

(Discussion Paper)

Francesco Benedetti[1,2,3,*], Antonio Pellicani[1,3], Gianvito Pio[1,3] and Michelangelo Ceci[1,3,4]

[1]Dept. of Computer Science, University of Bari, Via Edoardo Orabona 4, Bari, 70125, Italy
[2]Dept. of Computer Science, University of Pisa, Largo Bruno Pontecorvo 3, Pisa, 56127, Italy
[3]Data Science Lab, National Interuniversity Consortium for Informatics (CINI), Via Volturno, 58, 00185 Roma, Italy
[4]Jožef Stefan Institute, Jamova Cesta 39, 1000 Ljubljana, Slovenia

Abstract

Social networks have become central platforms for shaping public discourse, influencing opinions, and facilitating communication. However, these platforms also increasingly serve as breeding grounds for radicalization and the dissemination of hateful or criminal ideologies. With the exponential growth of users and content on social networks, an effective monitoring and detection of harmful actors have become critical for both societal well-being and security. In this discussion paper, we introduce IMMENSE, a novel machine learning-based system for detecting malicious social media accounts. Our framework leverages a multi-perspective approach that integrates three complementary dimensions to classify users: the semantics of the content they generate, the topology of their social relationship network, and the spatial information derived from their geographical position. The key innovation of our system lies in its inductive architecture, which enables generalization to previously unseen users or entirely new networks without requiring retraining, thus achieving significant advancement in both efficiency and practical applicability. We validate IMMENSE against a state-of-the-art transductive system using two diverse datasets extracted from the X social network, demonstrating competitive performance despite the inherent additional challenges introduced by the inductive setting.

Keywords

Social network analysis, Malicious user classification, Inductive learning, Multi-perspective classification

1. Introduction

Social networks are online platforms that facilitate interpersonal connections, communication, and interest-sharing. Over the last decade, they have become increasingly popular, emerging as one of the primary media for communication, information dissemination, and entertainment for a significant portion of the population. Their social impact is multifaceted, as their proliferation also came with a substantial increase of malicious activities such as cyberbullying, spam attacks, misinformation propagation, extreme political or religious views, and recruitment for illicit purposes. This increase in harmful behaviors highlights the need for robust detection mechanisms to identify potentially dangerous users and ensure the safety and integrity of these platforms.

Social networks can formally be modeled as graphs, where nodes represent users and edges denote the relationships between them, such as *following* or *friendship*. Accordingly, the detection of dangerous users falls under the node classification umbrella, that corresponds to labeling each user as either risky or safe. In the literature, several works attempted to solve this task using different approaches. Methods that classify users based solely on the posted content are commonly referred to as *content-based*, and are mostly employed to classify individual posts rather than users. An example is [1], where the authors

SEBD 2025: 33rd Symposium on Advanced Database System, June 16-19, 2025, Ischia, Italy
*Corresponding author.
francesco.benedetti@phd.unipi.it (F. Benedetti); antonio.pellicani@uniba.it (A. Pellicani); gianvito.pio@uniba.it (G. Pio); michelangelo.ceci@uniba.it (M. Ceci)
0009-0009-6610-9846 (F. Benedetti); 0000-0002-4193-3486 (A. Pellicani); 0000-0003-2520-3616 (G. Pio); 0000-0002-6690-7583 (M. Ceci)

aim to detect bot accounts. The authors of [2] further extend this work by evaluating various classifiers that exploit content and metadata from user profiles to identify spambots and fake followers on X.

On the other hand, approaches that focus on the structure of the network of users relationships are commonly called *topology-based*. A relevant example is SybilRank [3], that exploits early-terminated random walks to detect fake users in a social network. Similarly, [4] applies random walk-based techniques to identify scammers in Web3 transaction graphs.

Finally, *hybrid methods* aim to integrate multiple aspects. Ribeiro et al. [5] analyze both user-generated contents and social relationship graphs, comparing a GNN-based model against gradient-boosted tree classifiers for identifying hateful users. Wang et al. [6] consider user's demographic features, as well as the social connections, the generated content, and dynamic features in the Momo social network to detect malicious accounts. Similarly, a hybrid approach was implemented for detecting malicious users on GitHub in [7], which evaluates profile characteristics, user activity patterns, and interactions with both other users and repositories. Hybrid strategies have proven to be the most effective for detecting malicious users/accounts, as they consider multiple perspectives and are more difficult to deceive. These approaches often analyze different aspects of social networks using separate specialized modules, with a decision maker providing the final classification. This approach is adopted in [7] and [8], although in the latter the authors employed this strategy to analyze textual (i.e., posts authored by the users and received comments) and non-textual (user characteristics and social relationships) attributes. Another example of a hybrid approach is SAIRUS [9], a multi-view user classification framework that integrates three distinct perspectives: user-generated content, social relationships, and geospatial proximity among users. For network representation, SAIRUS adopts Node2Vec, which, as a random walk-based embedding technique, operates in a transductive setting: SAIRUS requires that all users (even the unlabeled ones) need to be known during the training phase. This limitation makes it inapplicable in dynamic real-world environments, where the classification of a novel user (unseen during the training of the model) would require a full re-training of the model.

To fill this gap, in this paper we discuss our method IMMENSE, a hybrid, multi-perspective, inductive system for the detection of malicious/risky users. The adoption of inductive techniques makes IMMENSE capable of learning models that can generalize to new, unseen users. In the following sections, we provide a detailed description of IMMENSE and evaluate its effectiveness on two real-world datasets.

2. The proposed method

Before describing our method, we formalize some key aspects. A social network can be represented as a graph defined by a triple $\langle N, C, E_T \rangle$, where:

- N is the set of nodes, with each node representing a user.
- C is the set of textual contents, i.e., posts created by each user. Each post can possibly be associated with the geographical position in which the user was located when such a post was generated.
- $E_T \subseteq N \times N$ is the set of topological relationships among users. Without loss of generality, social relationships are represented as directed links.

IMMENSE employs three specialized modules, respectively, for the semantic analysis of the textual content, for the analysis of social relationships, and for the analysis of spatial relationships among users, followed by a model that combines their contributions to make the final decision. A graphical view of IMMENSE is provided in Figure 1, while in the following subsections we describe in detail each module.

2.1. Semantic Content analysis

The goal of this module is to provide an initial classification of users considering only the semantics of the content they posted. For this purpose, we initially perform a preprocessing step that, for each user, concatenates the posts into a unified text and performs tokenization, stopword removal, and stemming. Then, a Word2Vec [10] model is trained on the resulting corpus. In particular, for each

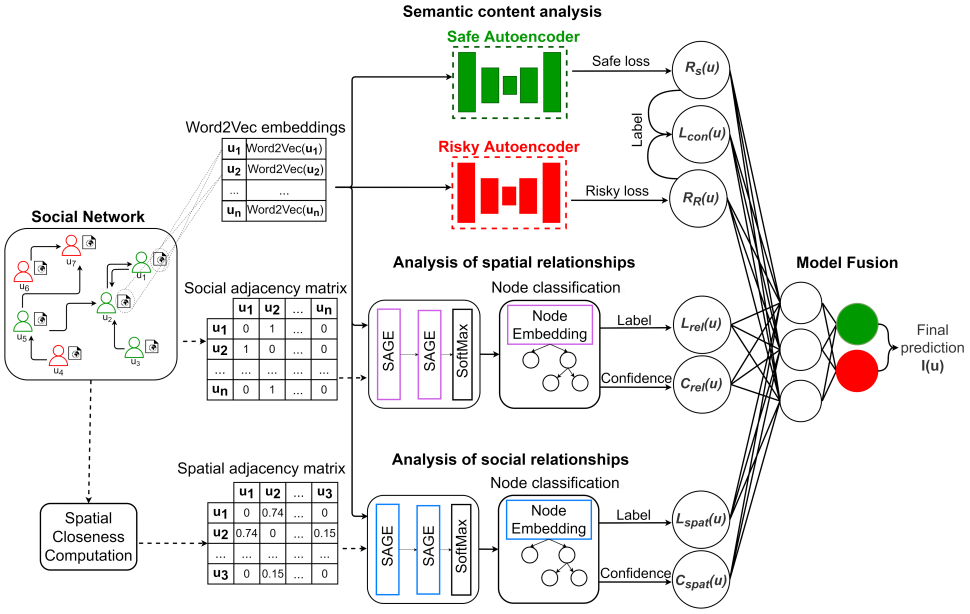

Figure 1: A graphical overview of the method IMMENSE.

user, the embeddings of the words appearing in their posts are aggregated through summation into a single vector representing the semantics of the published content. Although more recent approaches (e.g., BERT-based or LLM-based) could be used to identify a proper embedding of the textual content, Word2Vec proved to be more accurate in previous studies [11, 12].

Finally, a dedicated autoencoder is trained for safe users and for risky users, separately. Specifically, the autoencoder A_S is trained using the semantic vectors of users labeled as *safe*, while the autoencoder A_R is learned from the vector representations of *risky* users. Subsequently, each user embedding vector is processed through both autoencoders, leading to two reconstruction errors R_S and R_R, calculated as the mean squared error between the original and reconstructed vectors. Therefore, this module outputs the two reconstruction errors R_R and R_S, and a label $L_{con}(u)$ computed as $L_{con}(u) = 0$ if $R_S < R_R, 1$ otherwise, where 0 and 1 indicate the *safe* and *risky* labels, respectively.

We use two separate autoencoders because, as emphasized in [13], employing one distinct model per class rather than a single binary classifier provides greater stability in situations of label imbalance, that is the situation we expect in our scenario.

2.2. Analysis of social relationships

Nowadays, in social networks we can find a significant number of *passive* users, i.e., individuals who rarely share contents but primarily follow and consume content posted by other users. For these users, without complementary information, it would be very difficult to provide an estimate of the risk. Therefore, it is fundamental to integrate such information with interactions among users. Consequently, in this module, we analyze the social relationships among users as represented by E_T.

For this task, Graph Neural Networks (GNNs) have emerged as state-of-the-art approaches for processing graph-structured data and extracting latent representations of nodes [14]. An example of such GNNs is GraphSAGE [15], which generates the embedding of each node in the graph by sampling and aggregating features from neighboring nodes. In particular, GraphSAGE offers two significant advantages over other approaches (e.g., Node2Vec, adopted in [9]): *i)* it samples only a subset of each node's neighborhood, enabling efficient scaling to large networks, and *ii)* it provides inductive embedding capabilities, allowing generalization to previously unseen nodes or entirely new networks. Therefore, in IMMENSE, we perform the analysis of social relationships by stacking two GraphSAGE layers (see Figure 1). Notably, each layer allows expansion of the considered neighborhood by one

hop, resulting in a comprehensive two-hop neighborhood analysis when computing user embeddings. Furthermore, since GraphSAGE directly exploits the features of neighboring nodes, we associated each user in the graph with features corresponding to the user's semantic representation obtained in the semantic content analysis. In this way, although possibly introducing some redundancy with the information conveyed by the previous module, relationships are modeled while simultaneously considering the content posted by users.

For this module, contrary to the previous one, we adopt a tree-based classifier, due to their proven performances with network data [16]. Specifically, the node embeddings resulting from GraphSAGE serve as input to a random forest classifier that produces two outputs for each user: the classification label L_{rel} and an associated confidence value c_{rel}. The confidence is computed by averaging the purity of the leaf nodes where each instance falls across all trees in the forest.

2.3. Analysis of spatial relationships

Users who live in close geographical proximity often share similar real-life experiences, cultural contexts, and have increased opportunities for in-person interactions. These shared environmental factors can significantly influence opinions, behaviors, and vulnerability to certain types of content. In order to capture this precious information, we construct a network of spatial relationships that models the geographical connections among users. This task is performed through two key steps. First, we estimate each user's physical location by identifying the most frequent coordinates (latitude and longitude) associated with published posts[1]. Such coordinates, when available, are metadata associated with the posts. Second, we compute a weight to associate with the connections among users, that is inversely proportional to their geographical distance. Specifically, for any pair of users u_1 and u_2 with respective latitudes ϕ_1, ϕ_2 and longitudes λ_1, λ_2, we compute their geodetic distance as:

$$d(u_1, u_2) = 2r \cdot \arctan\left(\sqrt{\frac{a(u_1, u_2)}{1 - a(u_1, u_2)}}\right) \tag{1}$$

where r is the Earth radius (≈ 6371 km) and $a(u_1, u_2) = \sin^2\left(\frac{\phi_1 - \phi_2}{2}\right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2\left(\frac{\lambda_1 - \lambda_2}{2}\right)$ is the Haversine Formula. Then, we compute the mean μ_d and standard deviation σ_d among all distances, that are used to compute the z-score normalization of the distance $z(u_1, u_2)$ in $\mathcal{N}(0,1)$. Finally, the weight of the edge linking the two users in the graph is defined as:

$$closeness(u_1, u_2) = \begin{cases} \frac{z(u_1,u_2)}{min_z} & \text{if } z(u_1, u_2) < 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where min_z is the minimum of the normalized distances among users. In other words, if two users are closer than the average, the closeness score will range in the interval $(0, 1]$, otherwise, their closeness score will be set to zero (meaning that they are not connected in the spatial network).

Once the spatial network is built, it is processed in the same way as the social relationships network: a GraphSAGE-based model is trained, and the obtained embeddings are fed to a random forest classifier. This module outputs two values for a user: the label L_{spat} and the confidence value c_{spat}.

2.4. Model Fusion

After examining each individual perspective within the social network, the final step consists in integrating their contribution to make a final prediction. For this purpose, we adopt a Multi-Layer Perceptron (MLP) architecture following the stacked generalization approach. As illustrated in the right side of Figure 1, the employed MLP processes seven inputs, derived from the preceding single-perspectives analyses. More specifically:

[1]Note that if a user never shares the position of its post, he/she will be represented as an isolated node in the graph, with its embedding corresponding to its original feature vector.

- Semantic content analysis: the reconstruction errors R_R, R_S, and the classification label L_{con};
- Social relationships analysis: the classification label L_{rel} and its confidence value c_{rel};
- Spatial relationships analysis: the classification label L_{spat} and its confidence value c_{spat}.

The MLP-based model fusion component processes these diverse features through one hidden layer with ReLU activation function and produces a final binary classification for each user, exploiting the softmax function. Notably, this integration approach potentially allows IMMENSE to make more robust predictions than any single analysis module could achieve independently, in particular when certain features might be ambiguous or misleading when considered in isolation.

3. Experiments

To evaluate IMMENSE, we conducted comparative experiments against its closest counterpart, namely SAIRUS [9], using two real-world datasets derived from X (formerly Twitter). The first dataset, denoted as D_1, is the same used in [9]. The ground truth for D_1 was established using a keyword-based approach: tweets containing terms from two manually curated lists related to terrorism, extremism, hate speech, and discrimination against minorities[2,3] were flagged as risky. Subsequently, users were classified as risky if the majority of their published tweets received this flag. The second dataset, denoted as D_2, is a novel dataset constructed using the X API and keywords related to radicalism identified through the Horizon 2020 project CounteR[4]. The data collection followed a multi-step process: *i)* we began by retrieving up to 1500 tweets containing each radicalism-related keyword; *ii)* for each author of these tweets, we collected up to 1000 of their followers to establish a connected social network; *iii)* to ensure sufficient network connectivity, we kept users who follow more than 5 other users within our dataset; *iv)* we gathered the 20 most recent tweets from each user. Then, to create the ground truth labels, we implemented a semantic similarity approach using Google's pre-trained Word2Vec model[5]. Specifically, we generated semantic vector representations for each user, along with a reference vector derived from a corpus of known malicious content provided by CounteR project partners. Then, users whose semantic similarity to the malicious vector exceeded a defined threshold $\Delta = 0.88$ received the *risky* label, while the remaining users were labeled as *safe*. Such a threshold was determined through a preliminary analysis, and led to around 7% of users being labeled as risky, that can be considered reasonable. Furthermore, to incorporate the possible influence of social network relationships, we enhanced this labeling strategy by switching *safe* users to the *risky* class if at least 10% of their followed accounts were already identified as *risky*. We report key statistics about both datasets in Table 1.

3.1. Experimental setup

In order to show the advantages derived from the multi-perspective approach, we performed experiments with both IMMENSE and SAIRUS, considering several combinations of the three considered dimensions, namely content (C), social relationships (R), and spatial relationships (S). When a given dimension is not considered, the values outputted by its corresponding module are set to zero for all the users.

Both datasets D_1 and D_2 were split using 80% for training and the remaining 20% for testing. It is important to note that SAIRUS, being transductive, requires access to the complete network of relationships and geographical information during training, and cannot provide predictions for users absent in those networks. IMMENSE, on the other hand, exhibits a key advantage since it enables generalization to completely new networks in the testing phase. This aspect makes the comparison inherently unfair in favor of SAIRUS, since the competitor SAIRUS is aware of the users in the testing set, while we purposely make IMMENSE unaware of them. However, such a comparison allows us to assess

[2] https://www.dailymail.co.uk/news/article-2150281/
[3] https://github.com/msang/hateval
[4] https://counter-project.eu/
[5] https://code.google.com/archive/p/word2vec/

Table 1
Datasets information

Dataset	D_1	D_2
Number of Users	3834	37945
Number of Risky users	1559	2807
Number of Safe users	2275	35138
Number of Users with spatial information	3834	1043
Avg social following per user	5	16
Avg followers per user	6	2

the performance of IMMENSE in such a more challenging scenario, compared with its transductive counterpart that could not be applied at all to unseen users.

As evaluation measures, we considered precision, recall, F1 score, and accuracy, computed on the whole test set and for each class. Given that both datasets exhibit class imbalance, with safe users being the majority, our primary interest lies in how effectively the systems detect risky users.

3.2. Results and discussion

In Tables 2 and 3, we present the experimental results achieved for datasets D_1 and D_2, respectively. The best values achieved for accuracy and micro/macro F1 scores are highlighted in bold.

Examining the results for dataset D_1, we observe a clear benefit of considering multiple perspectives instead of only the semantics of the posted content. Indeed, when relying solely on the content, both IMMENSE and SAIRUS demonstrated suboptimal performances, with a similar F1 score and accuracy of 0.76 and 0.71, respectively. However, incorporating additional dimensions consistently improved the overall effectiveness. Notably, the inclusion of the spatial dimension generally yielded the most significant gains, highlighting the importance of geographical relationships in identifying risky users. Indeed, when the spatial dimension was considered, SAIRUS achieved the best F1 score of 0.88, while IMMENSE reached 0.86. Generally, IMMENSE is slightly outperformed by SAIRUS, but we remind that SAIRUS is aware of the testing nodes during the training. These results are further confirmed by looking at class-specific performances, especially for the risky class. Indeed, in their best configuration (which is $C + S$ for both systems), SAIRUS again achieves slightly better metrics compared to IMMENSE.

Shifting our focus to dataset D_2, we can see that the results are quite different. Indeed, it is evident that generally, relying solely on the content is enough to achieve strong performances, with both systems showing identical F1 score (0.92) and accuracy (0.98). Examining the impact of multiple dimensions, SAIRUS shows modest improvements when adding spatial relationships (see $C + S$ configuration), reaching its best F1 score of 0.94 for all users and 0.88 for risky users. On the other hand, IMMENSE shows a diverse behavior across the different configurations. In particular, the use of spatial relationships (see $C + S$ configuration) leads to a significant drop in terms of F1 score, with 0.72 for all users and 0.29

Table 2
Results on dataset D_1

	Configuration			**All users**				**Safe**			**Risky**		
	C	**R**	**S**	**Prec**	**Rec**	**F1**	**Acc**	**Prec**	**Rec**	**F1**	**Prec**	**Rec**	**F1**
SAIRUS	✓			0.79	0.74	0.76	0.71	1.00	0.49	0.66	0.59	1.00	0.74
	✓	✓		0.87	0.87	0.87	0.87	0.92	0.84	0.88	0.81	0.91	0.85
	✓		✓	0.88	0.89	**0.88**	**0.88**	0.96	0.83	0.89	0.81	0.95	**0.87**
	✓	✓	✓	0.88	0.88	**0.88**	**0.88**	0.92	0.87	**0.90**	0.84	0.90	**0.87**
IMMENSE	✓			0.79	0.74	0.76	0.71	1.00	0.49	0.66	0.59	1.00	0.74
	✓	✓		0.71	0.60	0.65	0.65	0.63	0.96	0.76	0.80	0.24	0.37
	✓		✓	0.86	0.87	0.86	0.86	0.94	0.80	0.87	0.78	0.93	**0.85**
	✓	✓	✓	0.83	0.79	0.81	0.81	0.79	0.93	0.85	0.87	0.66	0.75

Table 3
Results on dataset D_2

	Configuration			All users				Safe			Risky		
	C	R	S	Prec	Rec	F1	Acc	Prec	Rec	F1	Prec	Rec	F1
SAIRUS	✓			0.90	0.95	0.92	**0.98**	0.99	0.98	**0.99**	0.81	0.92	0.86
	✓	✓		0.90	0.97	0.93	**0.98**	1.00	0.98	**0.99**	0.80	0.95	0.87
	✓		✓	0.91	0.97	**0.94**	**0.98**	1.00	0.98	**0.99**	0.82	0.96	0.88
	✓	✓	✓	0.89	0.97	0.93	**0.98**	1.00	0.98	**0.99**	0.78	0.96	0.86
IMMENSE	✓			0.90	0.95	0.92	**0.98**	0.99	0.98	**0.99**	0.81	0.92	0.86
	✓	✓		0.95	0.89	0.92	**0.98**	0.98	1.00	**0.99**	0.92	0.79	0.85
	✓		✓	0.91	0.59	0.72	0.94	0.94	1.00	0.97	0.88	0.18	0.29
	✓	✓	✓	0.93	0.95	**0.94**	**0.98**	0.99	0.99	**0.99**	0.87	0.92	**0.89**

for risky users. This different behavior can be due to the fact that, unlike D_1, only a small portion of users in D_2 have associated location data (see Table 1). As a result, many users in the spatial graph remain isolated, causing their node embeddings to merely reflect their initial features without neighborhood aggregation. SAIRUS does not appear to suffer from this limitation possibly because of its transductive nature. Indeed, the locations of nodes in the testing set is known during the training, and, therefore, during the embedding. Therefore, it can exploit such a perspective more comprehensively, even when it is available for a limited number of users. In any case, IMMENSE achieves its optimal performance when all three dimensions are exploited $(C + R + S)$, attaining an F1 score of 0.94 that matches SAIRUS's best result. More importantly, when examining the risky class specifically, IMMENSE's full configuration outperforms all SAIRUS configurations with an F1 score of 0.89 versus SAIRUS's best of 0.88, even though it works in the more challenging inductive setting.

Overall, we can conclude that IMMENSE achieves interesting performances, that are comparable with its transductive closest competitor SAIRUS, in most of the configurations. This is an important result, considering the additional information from the test set available during the training for SAIRUS. Moreover, these results confirm the practical applicability of IMMENSE in real-world environments, since it can be adopted to estimate the risk of new users, without a full re-training of the model.

4. Conclusions

In this paper, we presented IMMENSE, an inductive, multi-perspective method for detecting malicious accounts in social networks. Unlike previous approaches that work in the transductive setting, thus requiring complete retraining to analyze new users, the primary contribution of IMMENSE is its ability to generalize to previously unseen users through an inductive learning framework.

Our experimental evaluation on two real-world datasets from X demonstrates that IMMENSE achieves competitive performance compared to the state-of-the-art transductive approach SAIRUS [9], despite the additional challenges introduced by the inductive learning. Furthermore, the experiments confirm that integrating multiple perspectives provides benefits to the classification performance. Particularly, the spatial dimension, when available, proved to be valuable in identifying risky users, highlighting the importance of geographical information in understanding user's behavior.

For future work, we plan to consider an additional critical aspect, that is the temporal dimension. The personalities of users evolve over time, and so do their behaviors on social media platforms. This can lead safe users to gradually move towards risky views, or vice versa.

Acknowledgments

The authors acknowledge the support of the EU Commission through the H2020 Project "CounteR - Privacy-First Situational Awareness Platform for Violent Terrorism and Crime Prediction, Counter Radicalisation and Citizen Protection" (Grant N. 101021607), and of the project FAIR - Future AI Research

(PE00000013), spoke 6 - Symbiotic AI, under the NRRP MUR program funded by the NextGenerationEU.

Declaration on Generative AI

The authors used Grammarly for grammar and spelling check. The authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] R. A. Igawa, S. Barbon Jr, K. C. S. Paulo, G. S. Kido, R. C. Guido, M. L. P. Júnior, I. N. da Silva, Account classification in online social networks with lbca and wavelets, Information Sciences 332 (2016) 72–83.

[2] A. Bhattacharyya, A. Kulkarni, Machine learning-based detection and categorization of malicious accounts on social media, in: International Conference on Human-Computer Interaction, Springer, 2024, pp. 328–337.

[3] Q. Cao, M. Sirivianos, X. Yang, T. Pregueiro, Aiding the detection of fake accounts in large scale social online services, in: Proceedings of the 9th USENIX Symposium on NSDI 2012, San Jose, CA, USA, April 25-27, 2012, USENIX Association, 2012, pp. 197–210.

[4] W. Li, Z. Liu, X. Li, S. Nie, Detecting malicious accounts in web3 through transaction graph, in: Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering, 2024, pp. 2482–2483.

[5] M. Ribeiro, P. Calais, Y. Santos, V. Almeida, W. Meira Jr, Characterizing and detecting hateful users on twitter, in: Proceedings of the international AAAI conference on web and social media, volume 12, 2018.

[6] J. Wang, X. He, Q. Gong, Y. Chen, T. Wang, X. Wang, Deep learning-based malicious account detection in the momo social network, in: 2018 27th International Conference on Computer Communication and Networks (ICCCN), IEEE, 2018, pp. 1–2.

[7] Q. Gong, Y. Liu, J. Zhang, Y. Chen, Q. Li, Y. Xiao, X. Wang, P. Hui, Detecting malicious accounts in online developer communities using deep learning, IEEE Transactions on Knowledge and Data Engineering 35 (2023) 10633–10649. doi:10.1109/TKDE.2023.3237838.

[8] Y. Tang, D. Zhang, W. Liang, K.-C. Li, K. Li, Uncovering malicious accounts in open mobile social networks using a graph and text-based attention fusion algorithm, IEEE Internet of Things (2024).

[9] A. Pellicani, G. Pio, D. Redavid, M. Ceci, SAIRUS: spatially-aware identification of risky users in social networks, Inf. Fusion 92 (2023) 435–449. doi:10.1016/J.INFFUS.2022.11.029.

[10] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, in: Y. Bengio, Y. LeCun (Eds.), 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings, 2013.

[11] G. D. Martino, G. Pio, M. Ceci, PRILJ: an efficient two-step method based on embedding and clustering for the identification of regularities in legal case judgments, Artif. Intell. Law 30 (2022) 359–390. doi:10.1007/S10506-021-09297-1.

[12] G. De Martino, G. Pio, M. Ceci, Multi-view overlapping clustering for the identification of the subject matter of legal judgments, Information Sciences 638 (2023) 118956.

[13] C. Bellinger, S. Sharma, N. Japkowicz, One-class versus binary classification: Which and when?, in: 2012 11th International Conference on Machine Learning and Applications, volume 2, 2012, pp. 102–106. doi:10.1109/ICMLA.2012.212.

[14] T. R. Murgod, P. S. Reddy, S. Gaddam, S. M. Sundaram, C. Anitha, A survey on graph neural networks and its applications in various domains, SN Computer Science 6 (2025) 1–12.

[15] W. L. Hamilton, Z. Ying, J. Leskovec, Inductive representation learning on large graphs, in: Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA, 2017, pp. 1024–1034.

[16] D. Stojanova, M. Ceci, A. Appice, S. Dzeroski, Network regression with predictive clustering

trees, Data Min. Knowl. Discov. 25 (2012) 378–413. URL: https://doi.org/10.1007/s10618-012-0278-6. doi:10.1007/S10618-012-0278-6.