Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 13, No. 2, 2025, 241-249.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Exploring the Relationship between User Posts and List Subscription Behaviors on Twitter/X

Havva Alizadeh Noughabi*

Computer and Electrical Engineering Department, University of Gonabad, Gonabad, Iran.

Article Info

Article History:
Received 04 February 2025
Revised 23 February 2025
Accepted 06 March 2025

DOI:10.22044/jadm.2025.15715.2688

Keywords:
Social Network Analysis, Twitter Lists, Behavioral Patterns in List Subscriptions.

**Corresponding authors: h.alizadeh@gonabad.ac.ir (H. Alizadeh Noughabi).*

Abstract

Social media platforms have transformed information consumption, offering personalized features that enhance engagement and streamline content discovery. Among these, the Twitter Lists functionality allows users to curate content by grouping accounts based on shared themes, fostering focused interactions and diverse perspectives. Despite their widespread use, the relationship between user-generated content and List subscription behaviors remains insufficiently explored. This study examines the alignment between users' post topics and their subscribed Lists, along with the influence of activity levels on this alignment. The role of content diversity in shaping subscription patterns to Lists covering a range of topics is also investigated. Additionally, the extent to which the affective characteristics—sentiment and emotion—of user posts correspond with the emotional tone of subscribed List content is analyzed. Utilizing a comprehensive Twitter dataset, advanced techniques for topic modeling, sentiment analysis, and emotion extraction were applied, and profiles for both users and Lists were developed to facilitate the exploration of their interrelationship. These insights advance the understanding of user interactions with Lists, informing the development of personalized recommendation systems and improved content curation strategies, with broad implications for social media research and platform functionality.

1. Introduction

Social media platforms have revolutionized the way users access and engage with information, offering tools to customize and streamline their online experiences. Among these tools, the *Twitter Lists* feature stands out as a unique functionality that enables users to curate and organize content by grouping accounts based on shared themes, interests, or industries. Through subscription to these Lists, users can streamline their feeds, prioritize relevant content, and minimize distractions from their broader timelines. As a result, they contribute to a more structured and manageable feed, empowering users to interact with content that aligns closely with their interests and preferences. This makes Lists an essential feature for users seeking to balance content variety with relevance in an increasingly dynamic digital space [1].

Over the past decade, Twitter Lists have gained significant popularity and have been utilized in various empirical studies for diverse purposes. Prior research has leveraged Lists to infer users' latent characteristics [2], identify relevant topics for individual users [3], track well-connected and topic-sensitive followers [4], differentiate between elite and ordinary users [5], assess the degree of homophily among users [6], capture emergent semantics [7], and explore the relationships between following, membership, and subscription patterns [8]. Recently, Twitter Lists have been employed to analyze terrorism-related activities [9], identify local user communities and their shared interests [10], and examine their role in branding strategies, particularly in facilitating celebrity endorsement decisions [11].

In comparison to prior research, this study examines the relationship between users' posting behaviors—such as the topics and frequency of their posts—and their List subscription patterns, with the aim of uncovering potential correlations. By examining this relationship from multiple dimensions, the study aims to bridge a gap in the existing literature concerning the interplay between subscription Lists and user-generated content. The findings aim to advance personalization strategies by refining recommendation systems, improve platform functionality by better aligning content recommendations with user interests and engagement patterns, and deepen the understanding of user behavior on social media platforms. To achieve this, we formulate the following research questions to address:

RQ1) What is the relationship between the topics of users' posts and their subscribed Lists, and how does this alignment differ across varying levels of user activity?

RQ2) How does the diversity of users' posts impact their subscription to Lists with varying topics?

RQ3) To what extent is user activity level correlated with List subscription behaviors, such as the number of Lists subscribed to?

RQ4) How do the affective characteristics (sentiment and emotion) of user posts align with the affective attributes of the content within Lists they subscribe to?

To address these research questions, we begin by employing BERTopic [12,13], a state-of-the-art semantic topic modeling technique, to extract topics from user posts and List content. Using these outputs, we construct topic profiles for users and Lists, enabling the analysis of topic alignment between users and their subscribed Lists, as well as its variation across different levels of user activity. Furthermore, we examine the relationship between the diversity of users' posts and the diversity of the Lists they subscribe to, uncovering patterns in content consumption and engagement. Beyond topic-based analysis, we incorporate affective features, such as emotion and sentiment, from both user posts and Lists to explore the role of emotional alignment in subscription behaviors. This comprehensive approach offers valuable insights into how posting behaviors, activity levels, topic diversity, and emotional tone influence users' engagement with Lists, shedding light on broader patterns of content interaction on Twitter/X. Therefore, this paper makes the following key contributions:

- We propose a methodology to examine the relationship between users' posts and their List subscription behaviors on Twitter/X. By employing BERTopic, we uncover patterns that highlight the alignment between the topics of users' content and the Lists they subscribe to.
- We assess how varying user activity levels influence List subscription behaviors, including the number and diversity of Lists subscribed to, and examine the emotional alignment between users' posts and their subscribed List content.
- Through experiments on Twitter data, we present findings elucidating the relationship between user posts and List subscription behaviors, enhancing understanding of content interaction patterns.

2. Data and Experimental Design

This section outlines the dataset used in this study, along with the methods applied to structure and analyze it. Furthermore, it provides an overview of the experimental design employed to explore the relationship between user posting behaviors and List subscription patterns.

2.1. Dataset

This study employs the dataset previously introduced and described in [14], which encompasses users and their subscribed Twitter Lists. The dataset collection process began by identifying the Lists associated with Ashton Kutcher, a widely followed Twitter user, following a similar approach to [2]. Using these Lists as an initial seed, all users subscribed to these Lists were retrieved. Subsequently, the dataset was expanded by collecting all Lists subscribed to by these newly identified users. This iterative expansion process was repeated four times to ensure comprehensive coverage and a diverse set of users and Lists. Finally, all user interactions, including tweets, retweets, likes, and replies, were gathered, and users with engagement in more than 200 English tweets were designated as golden users. Additionally, up to 500 of the most recent tweets from each List were collected. This resulted in a final preprocessed dataset comprising 1,284 users, 1,535 Lists, 6,059 subscription relations, and 1,581,321 tweets.

Table 1. Some frequent topics extracted by BERTopic.

Topic Theme	Topic Words
Cryptocurrency	'bitcoin', 'btc', 'cryptocurrency', 'blockchain', 'defi', 'dip', 'crypto', 'ethereum', 'market', 'buy'
Music	'music', 'album', 'song', 'kanye', 'songs', 'spotify', 'wizkid', 'tickets', 'dance', 'dj'
AI	'artificial intelligence', 'data science', 'ai', 'robot', 'big data', 'learning', 'machine learning', 'deep learning', 'artificial', 'ml'
Covid	'covid', 'cases', 'coronavirus', 'virus', 'pandemic', 'covid cases', 'deaths', 'new cases', 'covid deaths', 'hospitalizations'
Climate	'carbon', 'fossil', 'climate change', 'emissions', 'cop', 'climate', 'change', 'oil', 'energy', 'gas'

2.2. Topic Profile Creation using BERTopic.

Traditional topic modeling methods, such as Latent Dirichlet Allocation (LDA) [15], have been widely used for topic extraction in longer and more structured documents. However, when applied to short, unstructured, and noisy texts like tweets, LDA often struggles to generate meaningful results due to the brevity and lack of contextual depth in these texts [16,17]. Comparative evaluations of various topic modeling methods, including LDA, non-negative matrix factorization (NMF), Top2Vec, and BERTopic, have demonstrated that BERTopic and NMF are particularly effective for analyzing Twitter data [13]. Therefore, we employ BERTopic [12], a semantic topic modeling technique specifically designed to work effectively with short texts. BERTopic leverages pre-trained transformer models to generate high-quality document embeddings that capture semantic relationships within the text. These embeddings are then processed through dimensionality reduction techniques to enable efficient clustering, with the resulting clusters used to generate interpretable topic representations using a class-based TF-IDF approach. By employing BERTopic, we extract topics from user posts and List content, enabling the construction of detailed topic profiles for further analysis.

Formally, let M represent the set of posts published by users U and in the Lists L, from which active topics Z can be extracted using BERTopic. Each post $m \in M$ is treated as an individual document, with a topic assigned to each. Given $Z = \{z_1, z_2, ..., z_K\}$ be K active topics, we define *List Topic Profile* and *User Topic Profile* as follows:

List Topic Profile. Let $M_l = \{m_1, m_2, ..., m_N\}$ represent the set of posts published in List $l \in L$, and Z be the set of K topics. The List topic profile $TP(l)$ is represented as a vector of weights over these K topics, i.e., $(f_l(z_1), f_l(z_2), ..., f_l(z_K))$, where $f_l(z_k)$ denotes the number of posts in List l associated with topic z_k. To ensure the profile accurately reflects the proportional distribution of topics, each $f_l(z_k)$ is normalized by dividing it by the total number of posts in the List $(|M_l|)$.

User Topic Profile. Let $M_u = \{m_1, m_2, ..., m_N\}$ represent the set of posts user u has interacted with, and Z be a set of K topics. A user topic profile $TP(u)$ is represented by a vector of weights over K topics, i.e., $(f_u(z_1), f_u(z_2), ..., f_u(z_K))$, where $f_u(z_K)$ is the number of posts by user u associated with topic z_k. To ensure the profile reflects the proportional distribution of topics, each $f_u(z_k)$ is normalized by dividing it by the total number of u's posts $(|M_u|)$.

Prior to applying BERTopic, we performed light preprocessing on the tweets, following the recommendations in [12]. This included lowercasing the text and removing URLs, mentions, punctuation, and special characters. For embedding tweets into dense vector representations, we utilized BERTopic's default model, "*all-MiniLM-L6-v2*". Additionally, we ran BERTopic in its auto-configuration mode, which generated 539 topics for our dataset. Table 1 provides an example of the 5 most frequent topics identified and their corresponding top 10 words.

2.3. Measuring user-List similarity.

To assess the similarity between each user-List pair $(u.l)$ where user u subscribes to List l, we examine their respective topic profiles. As these profiles are represented as vectors, we calculate the similarity using the cosine similarity measure, as formulated below:

$$S(u,l) = cosine\_similarity(TP(u), TP(l)) \quad (1)$$

where $TP(u)$ and $TP(l)$ are the profile vectors of user u and List l, respectively, and $S(u,l)$ represents the similarity score between the user and the List.

2.4. User Categorization based on Activity Level

To examine the impact of user activity levels on addressing specific research questions, we categorized users based on their activeness, similar to [18]. The activity level of a user was computed as the total number of posts they interacted with on Twitter, including their own posts and other interactions such as likes, retweets, and replies. Users were ranked by their activity level and divided into four equal quartiles to ensure balanced group sizes. The first quartile represents *low-active* users, while the fourth quartile includes *high-active* users, who exhibit the highest levels of activity. This categorization allows us to investigate how activity levels influence user behavior and their relationships with Twitter Lists.

2.5. Affective Features.

To deepen our understanding of the relationship between users and their subscribed Lists, we integrated affective features into our analysis. Affective features capture the emotional and sentiment-driven aspects of textual content, offering a nuanced perspective on the tone and emotional resonance of interactions. We extracted two primary affective attributes: *emotion* and *sentiment*.

For emotion analysis, we utilized the DistilBERT pre-trained language model [19] to classify textual content into seven emotional dimensions: *anger, fear, joy, surprise, sadness, disgust,* and *neutral*, similar to [20]. We quantified the frequency of these emotions in tweets and aggregated them to create emotional profiles for users and Lists. This approach allows us to measure the distribution of emotional tones in the content posted by or associated with users and Lists. The emotional profile of a user is defined as follows, with a similar approach applied to Lists.

Emotional Profile. Let $M_u = \{m_1, m_2, ..., m_N\}$ represent the set of posts that user u has interacted with. An emotional profile $EP(u)$ is represented by a vector of weights over 7 emotion dimensions, i.e., $(f_u(e_1), f_u(e_2), ..., f_u(e_7))$, where $f_u(e_i)$ is the number of posts by user u associated with emotion e_i. To ensure the profile reflects the proportional distribution of emotions, each $f_u(e_i)$ is normalized by dividing it by the total number of posts of user u $(|M_u|)$.

For sentiment analysis, we map each post to its sentiment—positive or negative—using the RoBERTa pre-trained language model [21], which has been successfully employed in recent studies on social media analysis [22][23]. Then, we count the number of posts associated with each sentiment label for every user/List and then normalize so that the sum of weights of two sentiment labels equals 1. The sentiment profile of a user is defined as follows, with a similar approach applied to Lists.

Sentiment Profile. Let $M_u = \{m_1, m_2, ..., m_N\}$ represent the set of posts that user u has interacted with. A sentiment profile $SP(u)$ is defined as a vector of weights over two sentiment categories, i.e., $(f_u(positive), f_u(negative))$, where $f_u(positive)$ is the number of posts associated with positive sentiment and $f_u(negative)$ represents the number of posts associated with negative sentiment for user u.

2.6. Diversity Measurement

To assess the diversity of a user's engagement with topics, we focus on the significant topics within the user's topic profile $(TP(u))$, which are defined as those topics with which the user interacts most frequently. The diversity of a user's topic interactions is then calculated by considering both the *number of significant topics* and the *relationships* between these topics. If a user engages with a large number of distinct topics, the diversity score is higher, indicating a broad range of interests. However, if the topics are highly related, the diversity score is adjusted downward, reflecting a narrower range of interactions.

The similarity between topics t_i and t_j is represented by $S(t_i, t_j)$, where a higher similarity indicates that the topics are more related. This score is derived from the output of a topic modeling approach, such as BERTopic. The diversity score $DIV(u)$ for user u is then calculated as follows:

$$DIV(u) = |\{z_k \mid f_u(z_k) \in TP(u) > 0\}| + \frac{1}{\sum_{i=1}^{|TP(u)|}\sum_{j=i+1}^{|TP(u)|} S(t_i, t_j)} \quad (2)$$

To quantify the diversity of the Lists subscribed to by a user u, we calculate the average pairwise similarity across all the Lists the user subscribes. This metric reflects how closely related the topics of these Lists are to each other. A higher average similarity indicates less diversity, whereas a lower average similarity suggests a more diverse selection of Lists. Formally, the *Subscription Diversity Index* ($SDI(u)$) is defined as:

$$SDI(u) = \frac{1}{n(n-1)/2} \sum_{i=1}^{|L_u|}\sum_{j=i+1}^{|L_u|} (1 - S(l_i, l_j)) \quad (3)$$

where n represents the number of Lists the user u subscribes to $(|L_u|)$, and $S(l_i, l_j)$ denotes the similarity between the topic profiles of l_i and l_j, calculated using cosine similarity of their corresponding vectors.

3. Results and Analysis

In this section, we present and analyze the findings of our study by addressing each RQ. We provide a detailed examination of the results, supported by quantitative metrics, visualizations, and qualitative interpretations to derive meaningful insights.

3.1. Analyzing Topic Alignment Between Users' Posts and Subscribed Lists (RQ1)

Figure 1 illustrates the distribution of similarity scores between the topics of users' posts and the content of their subscribed Lists, computed based on their topic profiles (Eq. 1). A notable proportion of user-List pairs exhibit low similarity scores, indicating that many users subscribe to Lists that are not strongly aligned with their own posting behavior. This suggests that users may follow Lists for purposes beyond content alignment, such as exploratory interests, or passive consumption of information outside their primary focus. Furthermore, a smaller yet notable peak near a similarity score of 1 reflects a subset of user-List pairs with strong alignment, where users' posts closely match the topics of their subscribed Lists. These users may seek Lists that reflect their specific interests or expertise.

Further investigation into the relationship between topic alignment in users' posts and subscribed Lists and their *activity levels* can provide valuable insights for designing recommendation and content categorization systems. A t-test analysis reveals that the user-List similarity scores for low-active users are significantly lower than those of high-active users (p-value ~ $0 < 0.05$), indicating a statistically significant difference between the two groups. To visualize this relationship, Figure 2 compares the alignment between the topics of users' posts and their subscribed Lists across low-active and high-active users. Specifically, for each user category, Figure 2 depicts the percentage of subscribed Lists at varying similarity levels. The results reveal distinct trends based on the similarity score:

1) For similarity scores up to 0.5: Low-active users are more likely to subscribe to Lists with lower content alignment than high-active users, as reflected by a higher proportion of Lists with lower similarity scores. This is evident from the blue line (low-active users) consistently appearing above the red line (high-active users).

2) For similarity scores above 0.5: High-active users demonstrate a stronger preference for Lists with higher content alignment, subscribing more frequently to Lists that closely match their posting behavior.

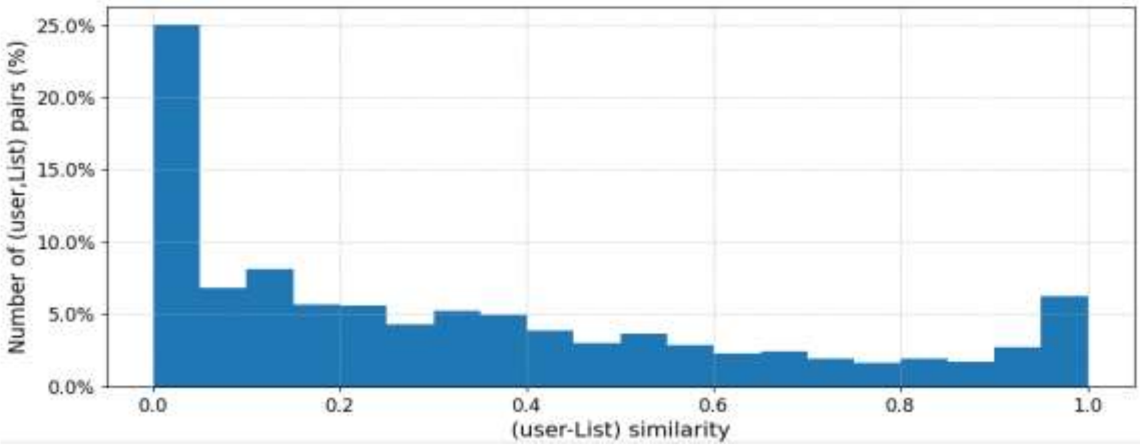

Figure 1. Distribution of similarity scores between the topics of users' posts and the content of their subscribed Lists.

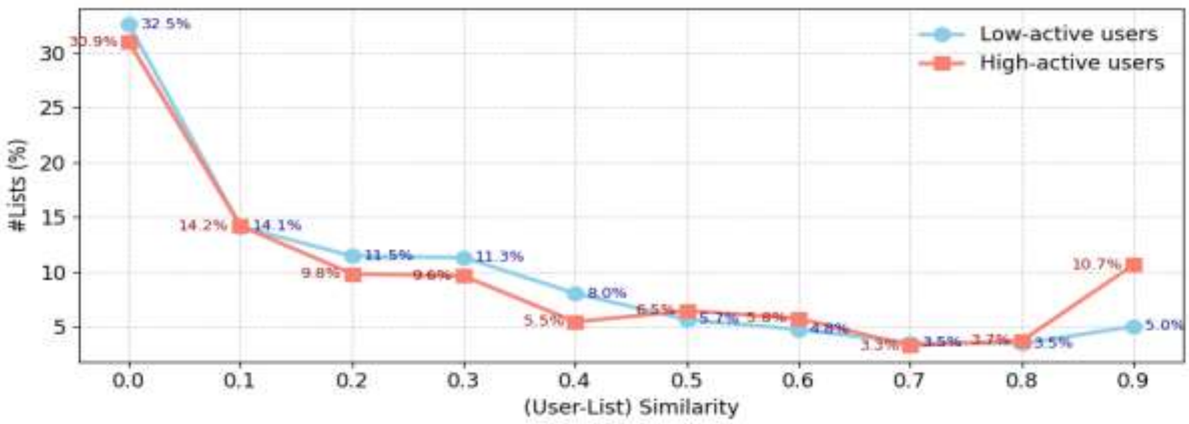

Figure 2. Distribution of user-List similarity scores for low-active and high-active users, showing the percentage of subscribed Lists at each similarity level.

3.2. Effect of User Post Diversity on Subscriptions to Diverse Lists (RQ2)

This section explores the relationship between users' post diversity and the diversity of their subscribed Lists. Users are first categorized into four bins based on their post diversity (Eq. 2): low, low-medium, high-medium, and high. Similarly, the Subscription Diversity Index ($SDI(u)$) (Eq. 3) is computed for each user and grouped into four bins. Figure 3 presents the distribution of users across these bins, highlighting how post diversity influences the diversity of subscribed Lists.

Across all user post diversity bins, there is a notable preference for subscribing to Lists with high and high-medium diversity, relative to other List diversity categories. Additionally, users with high post diversity show a significantly greater proportion of subscriptions to high-diversity Lists compared to those with lower post diversity, with a corresponding decrease in subscriptions to low-diversity Lists. In other words, users with high post diversity demonstrate a stronger preference for high-diversity Lists, highlighting a clear alignment between their broader content interests and subscription patterns. Additionally, the t-test analysis reveals a statistically significant difference in the SDI between high-active and low-active users (p-value=0.046), with high-active users exhibiting greater diversity in their subscribed lists.

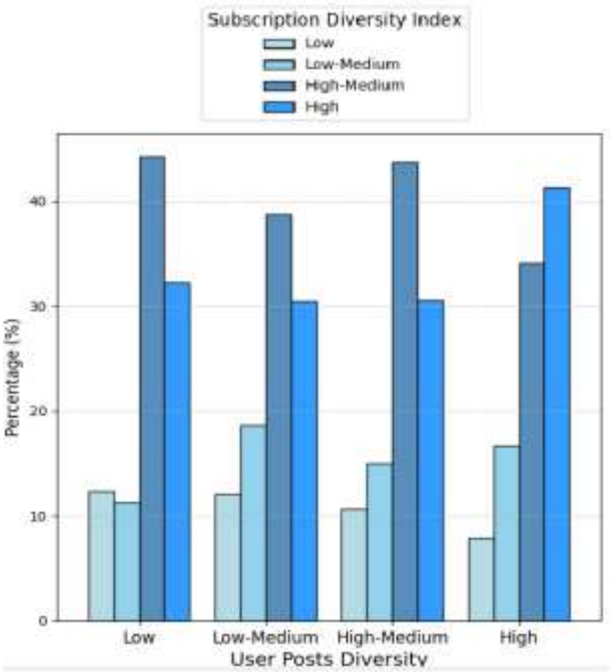

Figure 3. User distribution across post diversity levels and their corresponding subscriptions to Lists by diversity categories.

3.3. Impact of User Activity Levels on List Subscription Behavior (RQ3)

To examine the impact of user activity level on List subscription behavior, we first analyze the number of Lists users subscribe to. Figure 4 depicts the distribution of List subscriptions across four user activity levels: Low, Low-Medium, High-Medium, and High. The figure shows that users with higher activity levels tend to subscribe to a greater number of Lists, as reflected in their higher median number of subscriptions and larger interquartile range. This finding suggests that highly active users engage more extensively with Lists compared to their less active counterparts.

It is important to note that while our primary analysis focuses on the number of Lists to which users are subscribed, we also examine the number of Lists in which they are members across different user activity levels. As illustrated in Figure 5, highly active users are significantly more likely to be members of a larger number of Lists compared to low-active users. Notably, the disparity in List membership is even more pronounced than the differences observed in subscription behavior, indicating that the most active users not only subscribe to more Lists but are also more frequently included as members. Specifically, A one-way analysis of variance (ANOVA) was conducted to examine whether significant variations exist in the membership patterns across four user groups with different activity levels. The results revealed a statistically significant difference between the groups (p-value = $6.9e^{-7}$). However, no significant difference was observed in the subscription patterns (p-value = 0.19).

3.4. Affective Alignment Between User Posts and Subscribed Lists (RQ4)

As outlined in Section 2, we constructed sentiment and emotional profiles for each user and List by analyzing their respective posts. Figure 6 visualizes the alignment between the sentiment expressed in user posts and the sentiment of the Lists to which they subscribe, highlighting discrepancies in sentiment scores. The kernel density estimate (KDE) plot reveals that the observed differences are relatively minor, with most values concentrated within a narrow range of 0.2. This suggests a strong alignment between the sentiment profiles of users and the Lists they follow.

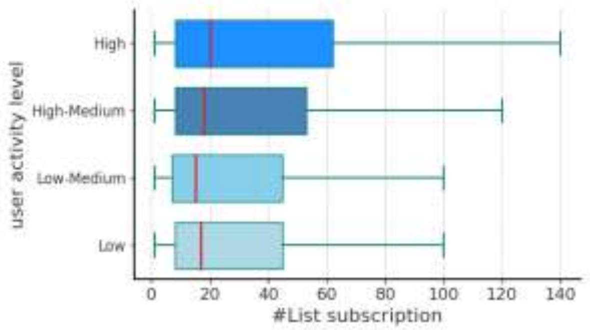

Figure 4. Relationship between user activity levels and the number of Lists to which users subscribe.

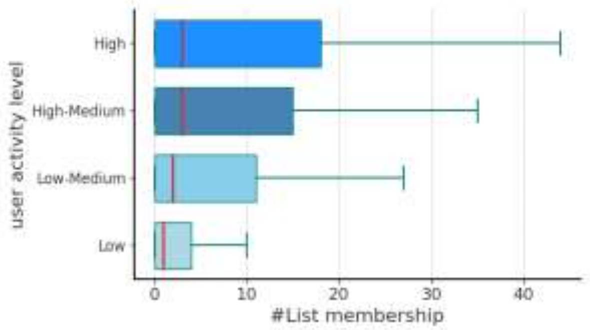

Figure 5. Relationship between user activity levels and the number of Lists in which users are members.

To analyze emotional differences across seven emotion categories—*anger, fear, joy, surprise, sadness, disgust,* and *neutral*—we compare the emotional profiles of users with those of the Lists to which they subscribe. As an illustrative example, Figure 7 depicts the emotional profile of a user who subscribes to two different Lists. This visualization demonstrates that the emotional profile of *List1* is more closely aligned with the user's profile compared to *List2*, highlighting variations in emotional compatibility between the user and the subscribed Lists.

Figure 8 illustrates the distribution of difference scores for each emotion dimension, revealing varying degrees of alignment between users' emotional responses and the emotional tone of the Lists they engage with. Overall, the observed differences in emotional profiles are relatively small, suggesting a general alignment between users' emotional profiles and the emotional tone of the Lists they follow. This indicates that, in most cases, users' emotional responses are consistent with the content they engage with.

Furthermore, the identification of outliers in certain emotion categories highlights users whose emotional profiles significantly deviate from the typical distribution of their subscribed Lists. These outliers may represent individuals with distinct engagement patterns, preferences, or emotional responses, providing valuable insights for enhancing personalized recommendation systems and developing more effective emotion-based content filtering strategies.

4) Conclusion

In this study, we explored the intricate relationships between user behaviors and their subscription to Lists on social media platforms like Twitter/X. By examining the alignment between users' post topics and the thematic focus of their subscribed Lists, the influence of activity levels and posting diversity on List subscriptions, and the affective characteristics shared between user posts and List content, we provided valuable insights into how users interact with and curate content. These insights hold significant implications for enhancing personalized recommendation systems, improving platform functionality, and fostering better user experiences through tailored content curation strategies.

As future work, we plan to explore dynamic modeling of user behavior to better capture the evolving nature of List subscription patterns over time. Additionally, we will investigate various forms of bias, such as demographic bias (e.g., age, gender), to gain a deeper understanding of how these factors influence user engagement and Twitter List subscription choices.

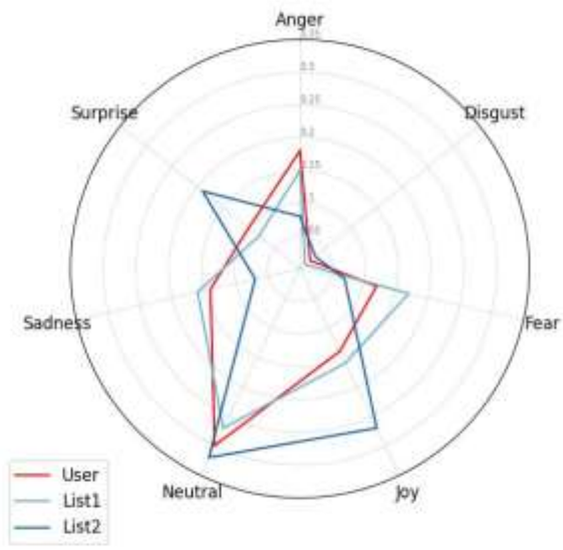

Figure 7. Example illustrating the alignment between the emotion expressed in user posts and the emotion of their subscribed Lists.

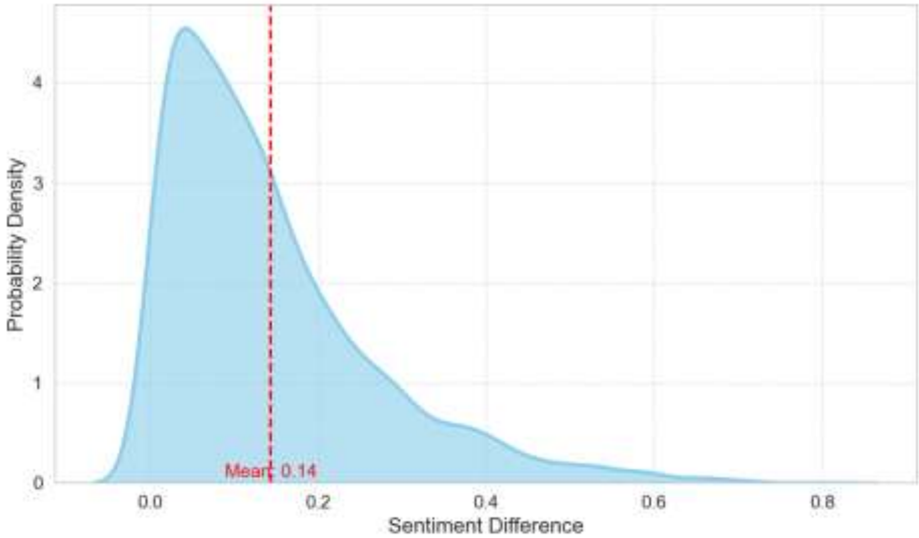

Figure 6. The alignment of sentiment between user posts and their subscribed Lists.

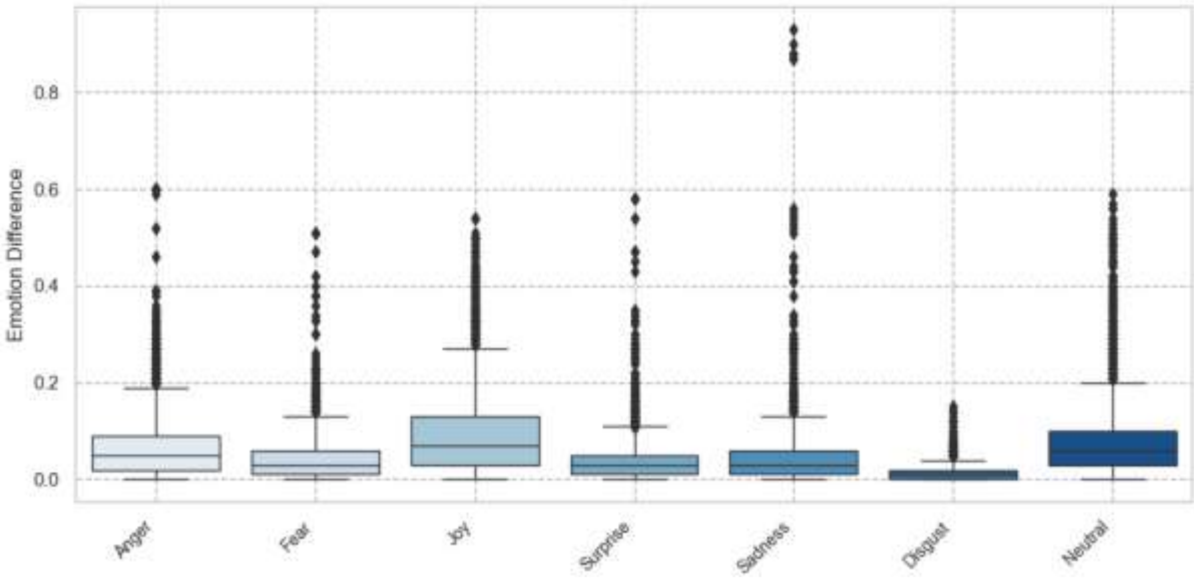

Figure 8. The alignment of emotions between user posts and their subscribed Lists.

References

[1] S. de la Rouviere and K. Ehlers, "Lists as coping strategy for information overload on Twitter," in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 199-200.

[2] D. Kim, Y. Jo, I.-C. Moon, and A. Oh, "Analysis of twitter lists as a potential source for discovering latent characteristics of users," in *ACM CHI Workshop on Microblogging*, 2010.

[3] Y. Yamaguchi, T. Amagasa, and H. Kitagawa, "Tag-based user topic discovery using Twitter lists," in *2011 International Conference on Advances in Social Networks Analysis and Mining*, 2011, pp. 13-20.

[4] P. Nasirifard and C. Hayes, "Tadvise: A Twitter assistant based on Twitter lists," in *Social Informatics: Third International Conference*, 2011, pp. 153-160.

[5] S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts, "Who says what to whom on Twitter," in *Proceedings of the 20th International Conference on World Wide Web*, 2011, pp. 705-714.

[6] J. H. Kang and K. Lerman, "Using lists to measure homophily on Twitter," in *AAAI Workshop on Intelligent Techniques for Web Personalization and Recommendation*, 2012.

[7] A. García-Silva, J.-H. Kang, K. Lerman, and O. Corcho, "Characterising emergent semantics in Twitter lists," in *The Semantic Web: Research and Applications – 9th Extended Semantic Web Conference*, 2012, pp. 530-544.

[8] S. Velichety and S. Ram, "Examining lists on Twitter to uncover relationships between following, membership, and subscription," in *Proceedings of the 22nd International Conference on World Wide Web*, 2013, pp. 673-676.

[9] K. Kaur, "Development of a framework for analyzing terrorism actions via Twitter lists," in *2016 International Conference on Computational Techniques*

in Information and Communication Technologies, 2016, pp. 19-24.

[10] M. Benabdelkrim, J. Savinien, and C. Robardet, "Finding interest groups from Twitter lists," in *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2020, pp. 1885-1887.

[11] C. Saridakis, C. S. Katsikeas, S. Angelidou, M. Oikonomidou, and P. Pratikakis, "Mining Twitter lists to extract brand-related associative information for celebrity endorsement," *Eur. J. Oper. Res.*, vol. 311, no. 1, pp. 316-332, 2023.

[12] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," *arXiv preprint arXiv:2203.05794*, 2022. [Online]. Available: https://arxiv.org/abs/2203.05794.

[13] R. Egger and J. Yu, "A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts," *Front. Sociol.*, vol. 7, 2022.

[14] H. Alizadeh Noughabi, B. Behkamal, and M. Kahani, "Description-based post-hoc explanation for Twitter list recommendations," *Comput. Knowl. Eng.*, vol. 7, no. 2, pp. 43-50, 2024.

[15] Blei, D.M., Ng, A.Y. and Jordan, M.I., 2003. Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), pp.993-1022.

[16] X. Cheng, X. Yan, Y. Lan, and J. Guo, "BTM: Topic modeling over short texts," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 12, pp. 2928-2941, 2014.

[17] W. Gao, M. Peng, H. Wang, Y. Zhang, Q. Xie, and G. Tian, "Incorporating word embeddings into topic modeling of short text," *Knowl. Inf. Syst.*, vol. 61, pp. 1123-1145, 2019.

[18] F. Zarrinkalam, H. A. Noughabi, Z. Noorian, H. Fani, and E. Bagheri, "Predicting users' future interests on social networks: A reference framework," *Inf. Process. Manag.*, vol. 61, no. 5, 2024.

[19] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter," *arXiv preprint arXiv:1910.01108*, 2019. [Online]. Available: https://arxiv.org/abs/1910.01108.

[20] F. Amiri, S. Abbasi, and M. Babaie Mohamadeh, "Clustering methods to analyze social media posts during coronavirus pandemic in Iran," *J. AI Data Mining*, vol. 10, no. 2, pp. 159–169, 2022.

[21] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "RoBETa: A robustly optimized BERT pretraining approach," *arXiv preprint arXiv:1907.11692*, 2019. [Online]. Available: https://arxiv.org/abs/1907.11692.

[22] Z Noorian, A. Ghenai, H. Moradisani, F. Zarrinkaam, and S. Z. Alavijeh, "User-centric modeling of online hate through the lens of psycholinguistic patterns and behaviors in social media," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 3, pp. 4354-4366, 2024.

[23] S. Z. Alavijeh, F. Zarrinkalam, Z. Noorian, A. Mehrpour, and K. Etminani, "What users' musical preference on Twitter reveals about psychological disorders," *Inf. Process. Manag.*, vol. 60, no. 3, pp. 103269, 2023.

تحلیل رابطه بین پست‌های کاربران و رفتارهای اشتراک‌گذاری لیست‌ها در توییتر

حوا علیزاده نوقابی*

مهندسی برق و کامپیوتر، مجتمع آموزش عالی گناباد، گناباد، ایران.

ارسال ۲۰۲۵/۰۲/۰۴؛ بازنگری ۲۰۲۵/۰۲/۲۳؛ پذیرش ۲۰۲۵/۰۳/۰۶

چکیده:

پلتفرم‌های رسانه‌های اجتماعی نحوه مصرف اطلاعات را تغییر داده‌اند و ویژگی‌های شخصی‌سازی‌شده‌ای را ارائه دادند که تعامل را افزایش می‌دهد و کشف محتوا را ساده می‌کند. در این میان، قابلیت لیست‌های توییتر به کاربران امکان می‌دهد با گروه‌بندی حساب‌های کاربری بر اساس موضوعات مشترک، محتوای خود را سازماندهی کرده و تعاملات متمرکزتر و دیدگاه‌های متنوع‌تری را تجربه کنند. با وجود استفاده گسترده از آن‌ها، رابطه بین محتوای تولیدشده توسط کاربر و رفتارهای اشتراک‌گذاری لیست‌ها به اندازه کافی بررسی نشده است. این مطالعه میزان تطابق بین موضوع‌های پست‌های کاربران و لیست‌های آن‌ها را به همراه تأثیر سطوح فعالیت بر روی این تطابق تحلیل می‌کند. همچنین، نقش تنوع محتوایی در شکل‌گیری الگوهای اشتراک‌گذاری لیست‌ها با موضوعات گوناگون مورد بررسی قرار می‌گیرد. علاوه بر این، میزان مطابقت ویژگی‌های عاطفی- احساس و هیجان - پست‌های کاربران با محتوای لیست‌های آن‌ها تجزیه و تحلیل می‌شود. با بهره‌گیری از یک مجموعه داده جامع از توییتر، تکنیک‌های پیشرفته مدل‌سازی موضوعی، تحلیل احساسات و استخراج عواطف، به کار گرفته شد و پروفایل‌هایی برای کاربران و لیست‌ها توسعه یافت تا ارتباط میان آن‌ها به‌دقت مورد بررسی قرار گیرد. یافته‌های این پژوهش به درک بهتر تعامل کاربران با لیست‌ها کمک کرده و زمینه را برای توسعه سیستم‌های توصیه‌گر شخصی‌سازی‌شده و استراتژی‌های بهینه‌سازی مدیریت محتوا فراهم می‌کند که تأثیرات گسترده‌ای در حوزه پژوهش‌های رسانه‌های اجتماعی و بهبود عملکرد پلتفرم‌ها خواهد داشت.

کلمات کلیدی: تحلیل شبکه‌های اجتماعی، لیست‌های توئیتر، الگوهای رفتاری در اشتراک‌گذاری لیست‌ها.