A Survey on Hashtag Recommendations

Kirti Jain
Delhi Technological University
Delhi, India
kj.kirtijain25@gmail.com

Rajni Jindal
Delhi Technological University
Delhi, India
rajnijindal@dce.ac.in

***Abstract*—As the need for recommending most customized products is increasing in order to boost up sales and thus the profits, it has become necessary to make the recommendations more and more targeted towards the end user. While a lot of research has been done on video recommendations for Netflix/YouTube, product recommendations for Amazon/Flipkart, Friend/User Profile recommendations on Facebook/LinkedIn, not much research has been done on hashtag recommendations which have become an integral part of these platforms now. In this paper, we provide a glimpse of hashtag recommendations that have been implemented in various fields and also provide a scope of future work.**

1. Introduction

Recommendation systems help the users to get items to their likeability. They are everywhere, from Facebook to LinkedIn; from Youtube to Netflix; from Flipkart to Amazon. They have many applications such as item based recommendation systems [1],[2], comparing user profiles [3],[4], social networks [5],[6],[7],[8], technology enhanced learning [9]. They are also used in various other fields related to recommendation such as predicting users' behaviour, datasets to be analyzed and analyzing quality with which prediction will be done, identifying the attributes of both user and item to be analyzed [10],[11]. Also, recommending does not only mean to present products by mapping out visible attributes but it also means to map unobserved product attributes and customer characteristics[12].

They offer customers to buy products according to their taste in e-commerce sites, without taking much time and effort. They not only help in e-commerce sites to purchase items [13],[14] but also are a critical part in decision making systems like sales and marketing[15]. It helps in picking the best web services as per the requirements [16]. They are also used in the field of Software engineering to help developers in a variety of tasks from reusing code to writing bug reports[17],[18].

Social networking is a revolution in modern communication. Its emergence has taken the communication system to a next level. Text messages, photos and videos which are uploaded on the social networking sites, have become a way of communicating with others. So, there has risen a need for a label expressing the emotions or views on the current issues.
These days, a new trend of tagging is becoming popular. Hashtags are preceded by the symbol #. Instagram, Pinterest and Twitter are popularly followed for hashtags. Similarly, hashtags are getting space in microblogs and microvideos. Also, they are quite useful in tagging subjects/topics for the questions asked on StackOverflow or Quora platforms. Posting a picture and then getting hashtag recommendations is an emerging trend on social websites. Thus there is a need to study and improve hashtag recommendations.

The rest of the paper is organized as follows: Section 2 describes the importance of hashtags. Section 3 explains some of the key studies of hashtag recommendations. Section 4 lays out a comparative analysis of the studies. Conclusion and future work are described in Section 5 and 6 respectively.

II. Why Are Hashtags Important

There are so many images, text messages, microvideos, microblogs that are being floated over the internet daily. Hashtags are a way of finding out a sense of all that data. They not only help us to group the data of users' posts with similar context, but also sort the data into different categories and make the searching easier. They allow us to watch text of users other than one's friend circle. They help in capturing new trending topics that may not be found otherwise. Fig.1 depicts the various types of hashtags. Fig2. represents the various applications of hashtags.

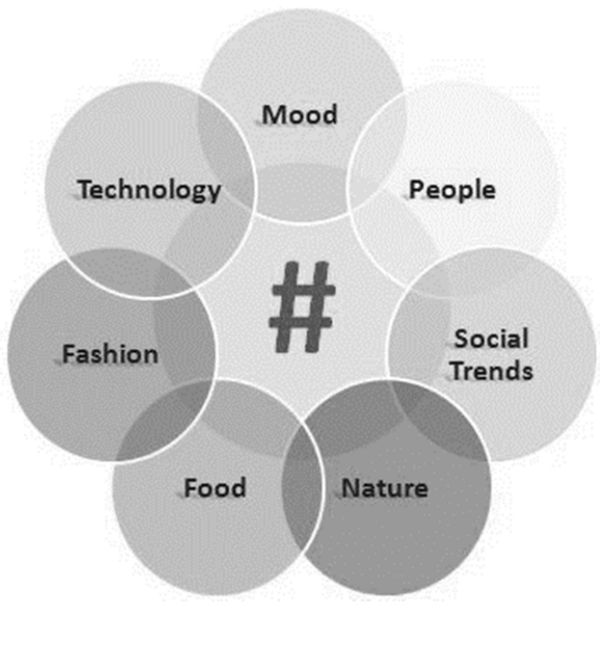

Fig. 1. Various types of Hashtags

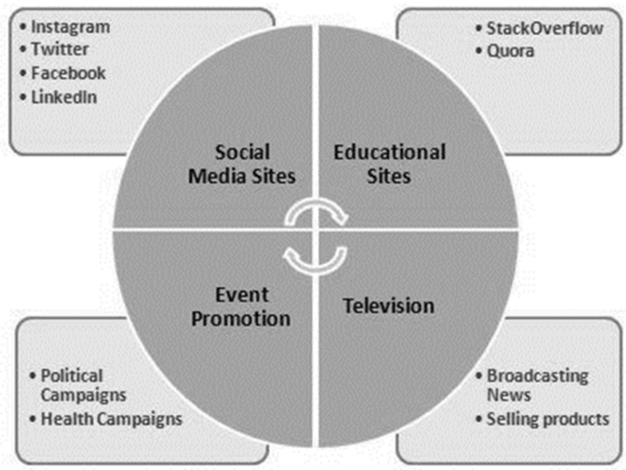

Fig. 2. Various Applications of Hashtags

II. Key Studies on Hashtag Recommendations

In this section, Tables I-XI represent summaries of a few key studies.

TABLE I. STUDY I

Paper Title	**Learning-to-Rank for Real-Time High-Precision Hashtag Recommendation for Streaming News, 2016 [19]**
Summary	Authors address the problem of hashtag recommendation to a stream of news articles as a Learning-To-rank problem so as to increase the interests of users for particular news on Twitter.
Data Set	RSS news feeds of 7 news organizations: Reuters, BBC, Irish Times, Irish Independent, Irish Examiner, RTE, and The Journal
Method/ Algorithm	Learning-to-Rank (L2R) Method is used to model the hashtag relevance. Here, news articles serve as query and hashtags serve as documents. Recommendation process is done in 2 steps: each article is connected to hashtag stream to produce candidate hashtags by a pre-ranking step and then L2R is used to measure the relevance of each candidate hashtag to recommend specific hashtags.
Future Work / Limitations	Analyzing the impact of hashtag recommendation on automatic story detection and tracking can be explored in future.

TABLE II. STUDY II

Paper Title	**Suggest what to tag: Recommending more precise hashtags based on users' dynamic interests and streaming tweet content, 2016 [20]**
Summary	In this paper, users' dynamic interests and streaming content on Twitter is combined to recommend hashtags.
Data Set	Twitter
Method/ Algorithm	Twitter-User LDA is used to learn the dynamic interests of users. Another model, Incremental Biterm Topic Model (IBTM) is used to discover the latent topics i.e. topics in current use of streaming Twitter content. By using these 2 models, hashtag recommendation is done by User-IBTM model.
Future Work / Limitations	Pre-processing can be included with extracting key phrases, as a hashtag may include several words instead of a single word.

TABLE III. STUDY III

Paper Title	**Hashtag Recommendation for Multimodal Microblog Using Co-Attention Network, 2017 [21]**
Summary	Authors take images along with textual information from microblogs for recommending hashtags. They propose a co-attention mechanism model which takes a combination of text and images.
Data Set	Twitter
Method/ Algorithm	VGG net is used to extract features from images. LSTM is used to extract text features. Both are combined using co-attention mechanism and then hashtags are recommended using single-layer Softmax classifier.
Limitations	The paper does not compare its results with the similar previous work.

TABLE IV. STUDY IV

Paper Title	**Modeling Chinese Microblogs with Five Ws for Topic Hashtags Extraction, 2017 [22]**
Summary	Authors first relate the microblogs to short-message-style news and then they relate words of microblogs to 5 Ws (Why, What, When, Where and hoW) and then recommend hashtags.
Data Set	Sina Weibo
Method/ Algorithm	This paper comprises of 4 phases: Microblogs spam filtering, handling popular Internet words and sentences, segmenting a microblogs into clauses and then finally recommending hashtag based on the 5Ws. Authors relate each word of microblog to one of 5Ws and then recommend hashtags using 5WTAG model.
Future Work / Limitations	This paper considers only Chinese microblogs. Other languages' microblogs can be explored in future.

TABLE V.: STUDY V

Paper Title	**Hashtagger+: Efficient High-Coverage Social Tagging of Streaming News, 2017 [23]**
Summary	In this paper, authors propose an approach Hashtagger+, a Learning-to-rank method for real-time hashtag recommendation of streaming news.
Data Set	Twitter
Method/ Algorithm	Hashtagger+, an efficient Learning-to-rank method is proposed to recommend hashtags by merging news in social streams in real-time. First, keywords are extracted from news articles and then Article-Hashtag Feature Vector is prepared by extracting candidate hashtags from relevant tweets. Hashtagger+ is then implemented in this feature vector to recommend hashtags.
Future Work	This approach can be applied in text mining and news story detection and tracking.

TABLE VI. STUDY VI

Paper Title	**Temporal enhanced sentence-level attention model for hashtag recommendation, 2018 [24]**
Summary	Authors develop a novel approach using LSTM which incorporates selective sentence-level attention for reducing noise factor. It also considers temporal information for hashtag recommendation in microblogs.
Data Set	SINA Weibo
Method/ Algorithm	LSTM incorporating selective sentence-level attention model is used. Selective sentence-level attention is used to assign weights to each sentence and then remove the noisy data. Temporal information is then introduced to recommend hashtags using a Softmax pooling layer.
Future Work / Limitations	This work includes microblogs with a single hashtag. Microblogs with multiple hashtags can be considered in future.

TABLE VII. STUDY VII

Paper Title	**Hashtag Recommendation for Photo Sharing Services, 2019 [25]**
Summary	A novel approach MACON (Memory Augmented Co-attention model) is developed. Here, hashtags are recommended using both image and text using co-attention neural network (LSTM). It also uses a memory unit for learning users' past hashtag posts.
Data Set	Instagram
Method/ Algorithm	Co-attention neural network (LSTM) is used for learning the features of both image and text. It also uses a memory unit for learning users' past hashtag posts. This is done in 2 steps: first step includes user-based random sampling of users' historical posts and their corresponding hashtags and the second step includes learning the tagging habits from users' historical posts and then connecting with the current post to tag.
Future Work / Limitations	Sampling method used is user-based random sampling, but user based temporal sampling or community based random sampling (in which posts of users' friends can be considered) can be used.

TABLE VIII. STUDY VIII

Paper Title	**DeepTagRec: A Content-cum-User based Tag Recommendation Framework for Stack Overflow, 2019 [26]**
Summary	The authors develop a deep learning model: DeepTagRec which uses content of questions (title and body) in StackOverflow and also heterogeneous network of user-tag relationships for tag recommendation.
Data Set	StackOverflow
Method/ Algorithm	Gated Recurrent Unit (GRU) model is used to represent content (title and body) of questions and encode the content to sequence of words (word2vec). node2vec model is used to capture the representation of user-tag relationships in heterogeneous networks. word2vec and node2vec are added and concatenated to perform tag prediction and recommendation.
Limitations	The source of data collection is not mentioned.

TABLE IX. STUDY IX

Paper Title	**Personalized Hashtag Recommendation for Micro-videos, 2019 [27]**
Summary	Authors involve interactions among users, hashtags and micro-videos for hashtag recommendation. This model considers users' preference on post contents and their personal understandings on hashtags. It also removes noise from micro-videos by using an attention mechanism.
Data Set	Public dataset YFCC100M and a self-collected Instagram dataset.
Method/ Algorithm	Graph Convolutional Network based Personalized Hashtag Recommendation (GCN-PHR) is used to analyze the graph involving interactions among three nodes: user, hashtags and micro-videos. It also uses an attention mechanism to remove redundant messages passed by micro-videos to users and hashtags.
Limitations	Even after filtering the noise, there is no significant improvement in terms of accuracy as compared with previous works.

TABLE X. STUDY X

Paper Title	**Sentiment Enhanced Multi-Modal Hashtag Recommendation for Micro-Videos, 2020 [28]**
Summary	Authors propose a novel approach TOAST for hashtag recommendation of micro-videos by considering 2 features, sentiment features and content features of 3 multi-modalities i.e. visual, audio and text.
Data Set	Vine
Method/ Algorithm	Multi-Layer Perceptron (MLP) is used for sentiment feature extraction and Bi-directional LSTM (Bi-LSTM) is used for content feature extraction of each modality (visual, audio and text). Then the proposed model, senTiment enhanced multi-mOdal Attentive hashtag recommendation (TOAST) is used to recommend hashtags.
Future Work	Emojis can be considered in text modality for recommending sentiment hashtags.

TABLE XI. STUDY XI

Paper Title	**User Conditional Hashtag Recommendation for Micro-Videos, 2020 [29]**
Summary	In this paper, authors include user profiles and historical hashtags for hashtag recommendation of micro-videos.
Data Set	Musical.ly
Method/ Algorithm	User Guided Hierarchical Multi-Head Attention Network (UHMAN) is proposed which combines user profiles and historical hashtags for micro-video hashtag recommendation. It is used to attend both image-level and video-level representations of micro-videos including user profiles.
Future Work / Limitation	This paper does not address the cold start problem that if a user is posting a micro-video for the first time, then there are no previous hashtags.

IV. DISCUSSION

Based on the above summaries, we divide the papers broadly in 3 categories: Streaming Content, Microblogs and Microvideos.

TABLE XII. HASHTAG RECOMMENDATION IN STREAMING CONTENT

Year	Parameters	Dataset	Results
2016 [19]	Streaming News	Twitter	Precision - 0.899
2016 [20]	Streaming Tweets	Twitter	HitRate - 78.50%
2017 [23]	Streaming News	Twitter	Precision - 0.9, Recall - 0.93, F1 Score - 0.5

From the Table XII, we can infer that not much work has been done on Instagram to inculcate current stories. As the stories on Instagram disappear after 24 hours, hashtag recommendation on streaming content will attract the users as they will get new hashtags after every 24 hours. Another application is recommending hashtags in YouTube comments boxes for live songs or streaming sessions. User’s interests can also be included so as to recommend hashtags according to their interests during the streaming contents.

TABLE XIII. HASHTAG RECOMMENDATION IN MICROBLOGS

Year	Parameters	Dataset	Results
2017 [21]	Text and images	Twitter	Precision - 0.311, Recall - 0.286, F1 Score - 0.298
2017 [22]	Relating words of text with 5 Ws (Why, What, When, Where and hoW)	Sina Weibo	Accuracy - 64.80%
2018 [24]	Temporal Information	Sina Weibo	Precision - 0.673, Recall - 0.665, F1 Score - 0.669
2019 [25]	Text, images and users' historical hashtags	Instagram	Precision - 0.62, Recall - 0.58, F1 Score - 0.34

As shown in the Table XIII, the authors of [25] yield a better result in terms of Precision, Recall and F1-Score as compared to [21] because they consider users' historical posts and their previous hashtags which has resulted in a better hashtag recommendation. Also, noise filtration and including temporal data [24] yields a better result. So, we conclude that extracting content features from text and images does not give a good result but including user profiles and temporal information gives a much better result.

Sentiment features can also be extracted from texts and images to recommend hashtags according to the swinging moods.

TABLE XIV. HASHTAG RECOMMENDATION IN MICROVIDEOS

Year	Parameters	Dataset	Results
2019 [27]	Interactions among users, hashtags and micro-videos	YFCC100, Instagram	Precision - 0.6847, Recall - 0.4075, Accuracy - 91.34%
2020 [28]	Sentiment features and content features of visual, audio and text.	Vine	Recall - 0.736
2020 [29]	User profiles and historical hashtags	Musical.ly	Precision - 0.673, Recall - 0.665, F1 Score - 0.669

As depicted in the Table XIV, the authors of [27] and [29] both consider user profiles but [27] removes noise or redundant messages by using an attention mechanism which improves the precision factor as compared to [29]. Also, Sentiment features in [28] are extracted from texts, audio and video clips which improve the recall factor. Along with user interactions/ user historical hashtags, user demographics can be included as a parameter for hashtag recommendation in the future.

V. CONCLUSION

In today's era, where social media plays an important role in a person's life, hashtag is an important feature within this social media. This is now widely used in order to gather the different messages from different users having similar and relative interests. It is particularly used in order to understand the relative trending topic within a particular social media. In this paper, we review some studies where hashtags are recommended. We categorized our literature survey in 3 fields: Streaming Content, Micro-Videos and Microblogs. The studies reveal that removing noise plays an important role in improving accuracy of the algorithm. Also, including user profiles, their interests and demographics can help improve the recommendation systems.

VI. FUTURE WORK

User based temporal sampling or community based random sampling for data collection can be done as future work. Also, microblogs with a single hashtag are used but microblogs with many hashtags can be considered in future. Analyzing the impact of hashtag recommendation on streaming content for automatic story detection and tracking can be explored in future.

REFERENCES

[1] Sarwar, B. M., Karypis, G., Konstan, J. A., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. Www, 1, 285-295.

[2] Linden, G., Smith, B., & York, J. (2003). Amazon. com recommendations: Item-to-item collaborative filtering. IEEE Internet computing, (1), 76-80.

[3] Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In The adaptive web (pp. 325-341). Springer, Berlin, Heidelberg.

[4] Lops, P., De Gemmis, M., & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. In Recommender systems handbook (pp. 73-105). Springer, Boston, MA.

[5] Liu, F., & Lee, H. J. (2010). Use of social network information to enhance collaborative filtering performance. Expert systems with applications, 37(7), 4772-4778.

[6] Lam, C. P., & Goeksel, M. (2010). U.S. Patent No. 7,689,452. Washington, DC: U.S. Patent and Trademark Office.

[7] Basu, C., Hirsh, H., & Cohen, W. (1998, July). Recommendation as classification: Using social and content-based information in recommendation. In Aaai/iaai (pp. 714-720).

[8] Cantador, I., Bellogín, A., & Vallet, D. (2010, September). Content-based recommendation in social tagging systems. In Proceedings of the fourth ACM conference on Recommender systems (pp. 237-240). ACM.

[9] Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender systems in technology enhanced learning. In Recommender systems handbook (pp. 387-415). Springer, Boston, MA.

[10] Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems (TOIS), 22(1), 5-53.

[11] Ekstrand, M. D., Riedl, J. T., & Konstan, J. A. (2011). Collaborative filtering recommender systems. Foundations and Trends® in Human–Computer Interaction, 4(2), 81-173.

[12] Ansari, A., Essegaier, S., & Kohli, R. (2000). Internet recommendation systems

[13] Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. Data mining and knowledge discovery, 5(1-2), 115-153.

[14] Schafer, J. B., Konstan, J., & Riedl, J. (1999, November). Recommender systems in e-commerce. In Proceedings of the 1st ACM conference on Electronic commerce (pp. 158-166). ACM.

[15] Pathak, B., Garfinkel, R., Gopal, R. D., Venkatesan, R., & Yin, F. (2010). Empirical analysis of the impact of recommender systems on sales. Journal of Management Information Systems, 27(2), 159-188.

[16] Manikrao, U. S., & Prabhakar, T. V. (2005, August). Dynamic selection of web services with recommendation system. In International Conference on Next Generation Web Services Practices (NWESP'05) (pp. 5-pp). IEEE.

[17] Robillard, M., Walker, R., & Zimmermann, T. (2009). Recommendation systems for software engineering. IEEE software, 27(4), 80-86.

[18] Pakdeetrakulwong, U., Wongthongtham, P., & Siricharoen, W. V. (2014, December). Recommendation systems for software engineering: A survey from software development life cycle phase perspective. In *The 9th International Conference for Internet Technology and Secured Transactions (ICITST-2014)* (pp. 137-142). IEEE.

[19] Shi, B., Ifrim, G., & Hurley, N. (2016, April). Learning-to-rank for real-time high-precision hashtag recommendation for streaming news. In Proceedings of the 25th International Conference on World Wide Web (pp. 1191-1202). International World Wide Web Conferences Steering Committee.

[20] Li, J., & Xu, H. (2016). Suggest what to tag: Recommending more precise hashtags based on users' dynamic interests and streaming tweet content. Knowledge-Based Systems, 106, 196-205.

[21] Zhang, Q., Wang, J., Huang, H., Huang, X., & Gong, Y. (2017, August). Hashtag Recommendation for Multimodal Microblog Using Co-Attention Network. In IJCAI (pp. 3420-3426).

[22] Zhao, Z., Sun, J., Yao, L., Wang, X., Chu, J., Liu, H., & Yu, G. (2017). Modeling Chinese microblogs with five Ws for topic hashtags extraction. Tsinghua Science and Technology, 22(2), 135-148.

[23] Shi, B., Poghosyan, G., Ifrim, G., & Hurley, N. (2017). Hashtagger+: Efficient high-coverage social tagging of streaming news. IEEE Transactions on Knowledge and Data Engineering, 30(1), 43-58.

[24] Ma, J., Feng, C., Shi, G., Shi, X., & Huang, H. (2018). Temporal enhanced sentence-level attention model for hashtag recommendation. CAAI Transactions on Intelligence Technology, 3(2), 95-100.

[25] Zhang, S., Yao, Y., Xu, F., Tong, H., Yan, X., & Lu, J. (2019). Hashtag Recommendation for Photo Sharing Services.

[26] Maity, S. K., Panigrahi, A., Ghosh, S., Banerjee, A., Goyal, P., & Mukherjee, A. (2019, April). DeepTagRec: A Content-cum-User Based Tag Recommendation Framework for Stack Overflow. In European Conference on Information Retrieval (pp. 125-131). Springer, Cham.

[27] Wei, Y., Cheng, Z., Yu, X., Zhao, Z., Zhu, L., & Nie, L. (2019). Personalized Hashtag Recommendation for Micro-videos. arXiv preprint arXiv:1908.09987.

[28] Yang, C., Wang, X., & Jiang, B. (2020). Sentiment Enhanced Multi-Modal Hashtag Recommendation for Micro-Videos. IEEE Access, 8, 78252-78264.

[29] Liu, S., Xie, J., Zou, C., & Chen, Z. (2020, July). User Conditional Hashtag Recommendation for Micro-Videos. In 2020 IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). IEEE.