A Recommender System based on Discovering the Triadic Frequent Closed Patterns Using Hidden Markov Model in Folksonomy

Maryam Fahimi
Department of Computer Engineering
Mashhad Branch, Islamic Azad University
Mashhad, Iran
Fahimi.Maryam@gmail.com

Majid Vafaei Jahan
Department of Computer Engineering
Mashhad Branch, Islamic Azad University
Mashhad, Iran
VafaeiJahan@mshdiau.ac.ir

***Abstract*— With the expansion of web 2.0, social tagging systems became to the dominant form of web content classification. The core of these systems is folksonomy that allows users to upload their resources and introduce them to others using the optional tags. Tag recommendation and resource recommendation are two main problems in these systems, but the huge volume of data and assigning freely tags are challenging for knowledge discovery from folksonomies. To solve this problem, the present research aims at clustering the folksonomy based on the discovery of the triadic frequent closed patterns. Thus the common concepts among users, tags and resources are retrieved. To this end, Hidden Markov Model is used. Also to overcome the problem of personal tagging and to enhance the suggestions quality we use Wikipedia, WordNet, and Google "did you mean" mechanism. Testing the real data taken from "Del.icio.us" website and comparing the results with the same algorithm in the field of folksonomy show that the proposed method has a better average in tag recommendation and resource recommendation in terms of the F-Measure test.**

Keywords: recommender systems, tag recommendation, resource recommendation, triadic frequent closed patterns, Hidden Markov Model

I. INTRODUCTION

The advent of web 2.0 and its daily expansion has made internet users from customers to active producers of web content and leads to a significant increase in the volume of web-based information. So it is difficult to find a user's favorite content and users face with the challenge of information overload. Social tagging systems are one of the web 2.0 applications. Nowadays, common tagging is widely used as an important tool to classify web content for searching and sharing [1]. The result of these collaborative tagging activities on these systems leads to the user-generated classification called "Folksonomy". A folksonomy has three main parts including users, tags, and resources (triple u, t, r), in which user u assigns tag t to the resource r. In these systems, the tags reflect the views of the different users about unique resources such as blogs, news, photos, etc. So, folksonomy contains valuable information and knowledge. Finding common concepts between users is one of the significant tools to discover knowledge from folksonomies.

Tag recommendation and resource recommendation are two main problems in social tagging systems benefiting from user-generated contents. As far as tag suggestion issue is concerned, when the user provides a resource annotation, the system suggests suitable tags for the resource. Tag suggestion helps to users in assigning correct and unambiguous tags and enriches the tagging information by providing offers. Also the resource suggestion is a way to overcome the information overload problem and helps users find their favorite resource in massive amounts of data.

Discovery of the triadic frequent closed patterns is a way for extracting common concepts in a folksonomy. In this paper, a method for tag recommendation and resource recommendation in social bookmarking services is proposed that links similarity between resources, tags and users through discovering triadic frequent closed patterns in folksonomy and uses these relations for the recommendation goals. We use a Hidden Markov Model to extract these patterns and finally to cluster folksonomy. More recommender methods that have been presented, separately consider the similarity between user to resource and resource to tag and rarely benefit from dependency between all three folksonomy dimensions simultaneously [2, 3] while the proposed method jointly applies all three dimensions data including users, tags and resources. Another problem is that since each user can freely select a keyword as a tag, uncontrolled vocabulary labels may affect recommendation quality when the words consist of redundant, synonyms, acronyms, grammatical mistakes, or misspelling and other non-standard forms. [4]. To solve this problem, our method uses external resources, i.e., WordNet, Wikipedia and Google.

The remainder of the paper is organized as follows:

Section II reviews some related works. Section III recalls the key notions used throughout this paper. In section IV, a new method for tag and resource recommendation is introduced.

The empirical evidences about the performance of our approach are provided in section V. Finally, section VI contains conclusions and future works.

II. RELATED WORK

With the advent of web 2.0 technologies, providing suggestions entered the new research area. In [5], Zhang et al. discussed on social tagging-based researches. In [6] Alper et al. proposed a Dirichlet-Multinomial distribution based model called Latent Interest Model (LIM) for resource and tag recommendation that accounts for users, tags and resources jointly. Alepidou et al. in [7], proposed a user-centric Bayesian model for tag recommendation that benefits from WordNet to overcoming tags. In [8] a resource recommendation is proposed that is dependent on the optimal stop techniques based on maximum reward. Based on the optimal stop, when the reward (recommendation quality) maximized, at this moment recommender system offers a list of suggestions to the user. FolkRank algorithm [9] used the PageRank algorithm to identify the important labels in order to tag recommendation. The proposed method in [10] uses a nonlinear optimization model for the computation of clustering parameters and maximizing the accuracy of clustering in tag recommender systems. Recommended algorithms in [11,12] Simultaneously models triple relations among users, tags and resources by using a three-dimensional matrix (known as three-order Tensor). Then, using analysis methods reveal hidden correlations in order to provide resource recommendation. The proposed method in [13] introduced a graph-based algorithm that models ternary relations between users, resources, and tags as a tripartite, weighted and undirected graph. Then applying the Katz measure to the tripartite graph provides personalized recommendations. The Katz score measures the proximity of two nodes based on weighted sums over collections of possible paths connecting those nodes. Recommender system in [14] is based on tripartite graph clustering. In this method, an improved k-means model for clustering tripartite graph is provided that used clusters as an indicator of similarity in the suggestion process. The recommended method in [15] is a graph-based method and Random walk algorithm. The proposed method in [16], first using Trias algorithm [17] discover triadic frequent closed patterns. Then a Hidden Markov Model based on these patterns is made and uses this model in tag recommendation process. Its difference with our proposed method is that our method uses external resources in order to improve the quality of suggestion. In addition, this method discovers triadic frequent closed patterns by using Viterbi algorithm based on Hidden Markov Model and then directly uses extracted patterns in tag and resource recommendation process. There are a lot of researches to improve the efficiency of the recommendation systems with the help of external resources. Lin et al. in [18] introduced a method to extract the ontological structures among tags that benefits from the power of association rules extraction supplemented semantic dictionary WordNet. Overell et al. in [19], compared the power of Wikipedia and WordNet for tags classification. They find important keywords that are not covered by WordNet but Wikipedia classifies them. Also in [20] a method for bookmark recommendation is proposed that uses Wikipedia and a small set of tags for finding the same users.

III. PRELIMINARIES

A. *Folksonomy*

As mentioned above, the purpose of this paper is to present a recommendation model in folksonomy systems. Folksonomies are the core structure of social bookmarking systems. A folksonomy has three main parts including users, tags, and resources (triple u, t, r), in which user u assigns tag t to the resource r. The following is an accurate and formal definition of folksonomy.

Definition (1): a Folksonomy is simply a tuple F:= (U,T,R,Y):

- U, T, and R are finite sets, whose elements are called users, tags, and resources, resp.
- Y is a ternary relation between them, i.e. $Y \subseteq U \times T \times R$ [21].

B. *Discovery of the triadic frequent closed patterns*

The frequent patterns are set of items that are repeated with a greater frequency than or equal to a threshold specified by the user. To recover the more important shared concepts we can additionally impose minimum support constraints on each of the three dimensions 'users', 'tags', and 'resources'. Also, closed in a triplex pattern means that none of these sets can be extended without shrinking one of the other two dimensions. Definition (2) describes triadic frequent closed patterns issue.

Definition (2): Suppose that F:= (U, T, R, Y) as defined in (1) is given folksonomy. A tri-set of F is a triple (A,B,C) with A⊆U, B⊆T ,C⊆R. Also, we consider u-minsupp and t-minsupp and r-minsupp the thresholds A, B, C in such a way that u-minsupp, t-minsupp, r-minsupp∈[0,1]. The task of mining all frequent tri-concepts consists in determining all tri-sets (A,B,C) of F with $\frac{|A|}{|U|} \geq u - minsupp,\ \frac{|B|}{|T|} \geq t - minsupp,\ \frac{|C|}{|R|} \geq r - minsupp$.

In addition, any set A, B, and C cannot be extended without shrinking one or two other sets. Also, sometimes it is more convenient to use absolute rather than relative thresholds. For this case we let τ_u:= |U|·u-minsupp, τ_t :=|T|·t-minsupp, and τ_r:= |R|·r-minsupp.

IV. THE PROPOSED METHOD

In this paper, in order to clustering folksonomy, we discover triadic frequent closed patterns. Thus, simultaneously three types of data will cluster and provide valuable information for recommender systems. Since similar users are interested in implementing similar behavior, resulting clusters can be offered as well. In the field of discovery of the triadic frequent closed patterns in folksonomy, we already in [22] proposed a method based on Hidden Markov Model [23]. This method is used in this article for clustering, and we use the extracted patterns in order to generate recommendations.

Our proposed model is able to offer tag and resource recommendation. This model consists of three main steps. The first step is preprocessing which includes tag integration and non-useful data deletion. In the second step with the discovery of the triadic frequent closed patterns, folksonomy will be clustered. The final step is suggestion which includes two sections for tag recommendation and resource recommendation. In the following, each of these steps is described.

A. Preprocessing

This step includes preprocessing on the folksonomy data and consisting of two phases, data integration and removing non-useful data.

Data integration: Although folksonomies are rich resources for recommender systems, since there is no controlled mechanism to assign tags in these systems, people often make grammatical mistakes (e.g."webdesgn" instead of "webdesign"), tag concepts indistinctly in singular, plural or derived forms ("start", "starts", "starting"). Sometimes adjectives, adverbs, prepositions or pronouns are added to the main concept of the tag ("beautiful car", "to read"), or use synonyms and acronyms that could be converted into a single tag ("Web" and "WWW", "hk" and "hong kong"). This phenomenon directly affects the performance of the recommended systems. [24,25]

The first step in preprocessing is a sequential implementation that its inputs are folksonomy tags and the output of one step is used as an input to the next step. The output of preprocessing step is a set of integrated tags which belong to an agreed representation. Hence, in order to perform this step correctly, we have considered three external resources:

1) WordNet: A lexical database and thesaurus that groups English words into sets of cognitive synonyms, providing definitions of terms, and modeling various semantic relations between concepts.

2) Wikipedia: A multilingual web-based encyclopedia that classifies inputs and also supports ambiguous words and acronyms.

3) Google "did you mean" mechanism: When a searched term is entered, the Google engine checks whether more relevant search results are found with an alternative spelling. Since Google's spell check is based on occurrences of all words on the Internet, it is able to suggest common spellings for proper nouns that would not appear in a standard dictionary.

The goal of this step is cleansing and enriching the user-generated tags. Hence inspired from the work of [26], for each tag different refining steps are executed.

1) Tag transformation: The selected tag t is set to lower case and special characters, such as accents and caret symbol, and are converted to their base form. For example, the tag "*Naïve*" is converted to "*Naive*". Additionally, common stop-words, such as pronouns, prepositions and conjunctions are discarded.

2) Syntactic filtering: It aims to check whether the transformed tag t is a standard one, i.e., the tag has an exact matching with an entry in WordNet. In this case, we return back to the stage (1) and select the next tag.

3) Wikipedia correlation: In order to provide an agreed representation of t, we check whether this tag is a jargon one. In such case, we correlate it to its appropriate Wikipedia entry. For example, when searching the tag "hk" in Wikipedia, the entry with the title "Hong Kong" is retrieved. The advantage of using Wikipedia to agree on tags from folksonomies is that Wikipedia is a community-driven knowledge base, much like folksonomies are, so that it rapidly adapts to accommodate new terminology.

4) Tag correction: If t is not a standard tag or a jargon one, i.e., nonsense tag, e.g., "webdesgn", we consider possible misspellings, e.g., "webdesign" and/or compound nouns, e.g., "web design". To solve these problems, we make use of Google which suggests a common spelling for t, e.g., "web design".

Removing non-useful data: In order to avoid additional processing and increase speed of folksonomy clustering, non-useful data will be deleted. This step consists of two phases: in the first phase only those rows of dataset where in each set of U,T,R the number of occurrences of each element is equal to multiply of other sets thresholds will be extracted. This means that, this step will extract the data that for users as much as $\tau_r \times \tau_t$ or more, to resources as much as $\tau_u \times \tau_t$ or more and for tags as much as $\tau_u \times \tau_r$ or more are repeated. In the second phase, from the output of the previous phase, we extract a part of the dataset where for each element in the first set, its Corresponding element in the second set is repeated as much as the threshold that is considered for the third set. e.g. for each t, its related u is repeated as much as τ_r.

B. Clustering

At this step the proposed system performs clustering task. The folksonomy clustering by the discovery of the triadic frequent closed patterns leads to extract triplex groups from similar users who assigned same tags to common sources. How to produce these clusters is described in detail below.

Step 1: Learning Hidden Markov Model

A real folksonomy dataset in addition to user number, information resource address and dedicated tags may include time of tag registration, source format, tag weight, and other additional information. In this article, only user numbers, information resource address, and assigned tags, among others, are important and other information will be ignored.

HMM[1] is a powerful mathematical model that obeys from the concepts of Markov models. We only use the resources and tags data to produce HMM, in order to consider resources as states of the HMM and tags as observations of HMM.

As mentioned in the previous sections, in the problem of mining all frequent tri-concepts, there are three sets as <U,T,R> where any user in U, allocated all tags in T to each resource in R. In this problem, Q is a set of tags and is considered as observations of HMM, $Q = \{q_1, \dots, q_{nq}\}$. Also, S is a set of resources and is considered as states of the HMM, $S = \{s_1, \dots, s_{ns}\}$. $nu_{q,j}$ is the number of users who have devoted tag q to the resource j, and $ns_{i,j}$ is the number of times that resource j is tagged after resource i. Our HMM is as follow:

[1] Hidden Markov Model

$$\lambda = (A, B, \pi) \quad (1)$$

Equation (1) is a probabilistic model which its parameters using folksonomy are computed as follows:

$\pi = [..\pi_i..] = P(s_i) = \frac{ns_i}{|Y|}$: In this problem, the probability of being in each state as the initial factor is equal to result of dividing the number of times a tag is assigned to resource i, divided by number of rows in the dataset.

$B = [\ldots b_j(q) \ldots] = P(q|s_j) = \frac{nu_{q,j}}{ns_j}$: The probability of observing tag q in state s_j is equal to the number of times that users assigned tag q to the resource j, divided by the total times resource j is tagged.

$A = [\ldots a_{ij} \ldots] = P(s_j|s_i) = \frac{ns_{i,j}}{ns_j}$: The probability of transition from state s_i to s_j is equal to the number of times that resource j is tagged after resource i, divided by the total times resource j is tagged.

Figure 1 shows an example of the proposed method HMM.

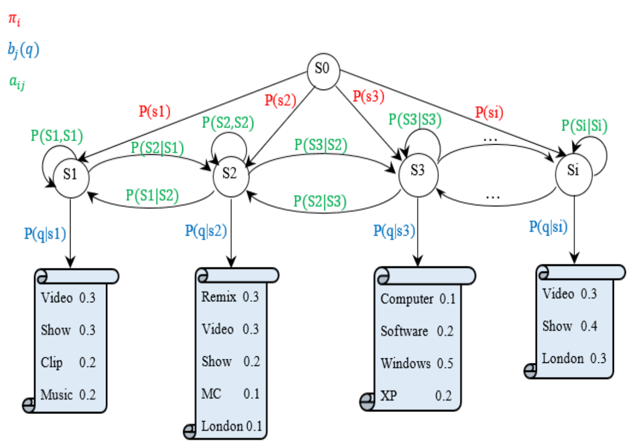

Figure 1. Example of the proposed method HMM

Step 2: Inference of Hidden Markov Model

To infer from created hidden model, we use a Viterbi algorithm. This function receives an observation sequence and the HMM, then returns the probability that the observation sequence observes in each state of the HMM.

In order to support the τ_t, and mining groups of the same resources that have received similar tags, we make non repeated combinations of set t with length τ_t and consider them as an observation sequence for Viterbi algorithm. Then among generated outputs, only those results for later processing will be saved that at least for τ_r resources the probability of observing input string exists. The output of this step is shared tags and resource groups that satisfy the threshold values. In fact, this step will lead to the dyadic clustering of similar tags and resources.

C. Final processing

In fact, this step includes applying third dimension means that users extracted on each cluster from the previous step, which may break down or join groups according to the corresponding users.

Breaking clusters: In this process, based on each extracted dyadic frequent pattern in inference of HMM step, the clusters are broken into several clusters so that each cluster includes triple users, resources and tags and each user assigns all existing tags to each resource.

Joining clusters: After applying the third dimension and breaking clusters, clusters containing same users and resources having different sets of tags may be created. This means a group of similar users are assigned two sets of tags to the same resource. In this case, in order to mine all frequent tri-concepts, we integrate two clusters on tags.

Finally, this section provides all tri-concepts. How to determine the non-useful data in phase C, also the clustering in phase D and final processing in phase E are exactly modeled from our presented method in [22].

D. The recommender system

In social bookmarking systems, tag recommendation supports user tagging process and provides the appropriate tags for a resource when a user wants to annotate it. In addition, the resource recommendation helps users retrieve their favorite and similar resources. When multiple users are assigned similar tags to a group of resources the following can be concluded:

- These groups of users have the same interests and are similar, because they assigned same tags to the same resources.
- The resources that the similar users assigned same tags to them are similar.
- The tags which are assigned to similar resources by same users are similar.

So in this section extracted triadic frequent closed patterns from the previous step are used to offer recommendations and include tag and resource recommendation to the users. Suppose that the proposed sample size is equal to k. In this case the recommendation process will consist of the following steps:

1. First, the system creates a candidate set of resources that have been used more frequently.
2. Searches for target user in the extracted clusters from previous step. If the user is known in a cluster, the resources which are in this cluster, or in clusters that have similar users with this user will be selected.
3. Among these resources, k resources that are visited more by these users are recommended. Otherwise, the created candidate set for the resources will be offered to the user.

In tag recommendation same to resource recommendation the following steps are performed:

1. First, the system creates a candidate set of tags that have been used more frequently.
2. In the extracted clusters from previous step searches for target user. If the user is known in a cluster, the tags which are in this cluster, or in clusters that have similar users with this user, will be selected.

3. Among these tags, k tags that are visited more by these users are recommended. Otherwise, the created candidate set for the tags will be offered to the user.

In fact, the process of candidate set creation is performed to overcome the cold start.

V. Experiments

The proposed method was tested on a dataset taken from the Delicious website which is ordered based on the time of tag assignment. This dataset consists of |U| different users, |T| different tags and |R| different resources, which are linked by |Y| triples (u, t, r) and is easily downloadable[2]. After applying preprocessing steps, the resulting dataset is divided into training and test dataset with 70% and 30% ratio which is shown in detail in Table 1.

TABLE I. DETAILS OF DATASETS

DataSet	The size of each set			
	|*Y*|	**|*U*|**	**|*T*|**	**|*R*|**
Main	131968	941	8290	65994
Train set	92018	572	6020	46087
Test set	39950	703	4137	22427

The proposed method for each above-mentioned problem compared with the similar algorithm presented in [14] which is based on the clustering triple graph and is provided in folksonomy. In this method, each node in the network can be denoted by its link vector to the other type of nodes. For example, a resource can be represented by two link vectors, the first one includes the weight vector of users, and the second one includes weight vector of tags. Tripartite Clustering Algorithm has a similar approach with k-means. Following the random initialization of cluster numbers of nodes, nodes are iteratively replaced into the clusters depending on their distance to the centers of clusters. Different from k-means, the distance between the node to be replaced and the center of the cluster is calculated by cosine similarity of the two types of node vectors. The other difference of this approach is the calculation of the centroids of the clusters. To take the interactions among the cluster structures of different types of nodes into account, a centroid of a cluster is calculated by not only the current cluster nodes the but also by the other two types of nodes. In this approach, in order to reduce the negative effect of these characteristics of folksonomies and to get a more connected and semantically related samples, the Porter stemming algorithm is applied on tags before clustering for removing the commoner morphological and in flexional endings from words. It also uses WordNet in the recommendation step to compare the semantic similarity of tags. This method is able to offer tag and resource recommendation and named TGC in the charts.

The important thing is that choosing small values as mentioned above is not consistent with the logic of discovering frequent closed patterns, since the goal is to find overlapping sets. This can be achieved by choosing larger values for the parameters. In addition, if these parameters choose very large, causes a lot of useful relations remain hidden. Thus, in all experiments we have taken threshold values $\tau_{\boldsymbol{u}} = \tau_{\boldsymbol{t}} = \tau_{\boldsymbol{r}} = 2$. These values are used in more researches of discovery of the triadic frequent closed patterns as the standard values.

A. Evaluation metrics

We use the precision, recall and F-Measure as accuracy metrics to evaluate the proposed recommendation algorithm. When referring to recommender systems, these metrics can be defined as follows:

The precision is the percentage of correctly recommended tags over the recommended tags:

$$Precision = \frac{|I_{Rec} \cap I_U|}{|I_{Rec}|}. \qquad (2)$$

Where I_{Rec} is the number of recommended items and I_U is the real samples used by the user. Thus $I_{Rec} \cap I_U$ Contains the correct samples.

The recall of the tag recommendation is the percentage of correctly recommended tags over the total tags:

$$Recall = \frac{|I_{Rec} \cap I_U|}{|I_U|} \qquad (3)$$

F-Measure is defined as the harmonic mean of precision and recall as follows:

$$\text{F-Measure} = \frac{2 \times Precision \times Recall}{Precision + Recall} \qquad (4)$$

B. Results

The results of the experiments are presented in charts. In each chart, a point on the curve is the number of samples used in the recommendation process. The experiments are repeated for 5 to 10 samples, and the results are marked on the curves from left to right.

1) The results of resource recommendation

In folksonomy-based recommender systems when there are many resources, it is unlikely that the user will visit exact recommendation of the item. Thus, we compare the resource with the same resources that are in a cluster with it.

[2] http://www.uni-koblenz-landau.de/koblenz/fb4/ AGStaab/ Research/DataSets/PINTSExperimentsDataSets/index_html

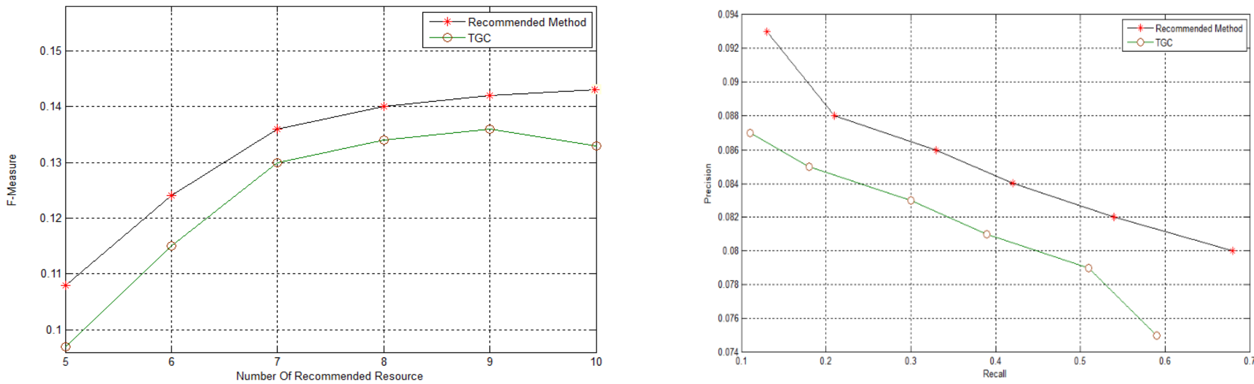

Figure 2. Precision/Recall and F-Measur for resource recommendation

The right chart in Figure 2 shows the Precision/Recall curve of recommended method and TGC for resource recommendation. In addition, the left chart in Figure 2 shows the F-Measure values when suggest the different number of resources. The results of these experiments show that our proposed method compared to the TGS on average and approximate increase 12% accuracy of resource recommendation.

2) The results of tag recommendation

Folksonomy clustering can also be used for tag recommendation, but uncontrolled vocabulary makes the evaluation difficult. For overcoming this problem, we also run data integration on the tags which are compared with the proposed collection. The test results in Figure 3 indicate that the proposed method in average and approximate increases 25% accuracy of tags recommendation compared to the TGS. So we can find from these two charts that the proposed method offers better accuracy in tag recommendation in comparison with TGS.

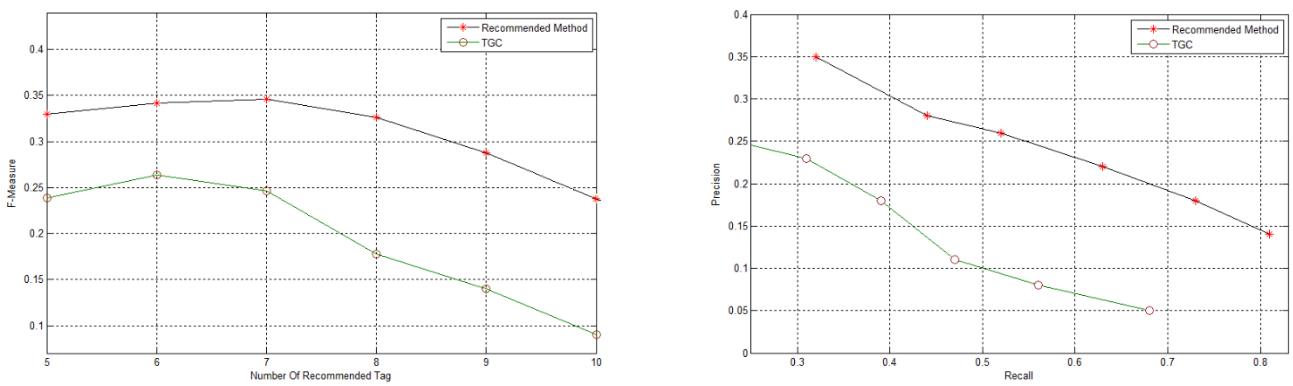

Figure 3. Precision/Recall and F-Measur for tag recommendation

CONCLUSIONS AND FUTURE WORK

In this paper, a method for overcoming the tag and resource recommendation problem in social tagging systems is presented. The proposed method benefits from advantage of all three dimensions of folksonomy simultaneously, and using HMM discovers all triadic frequent closed patterns. Thus, it extracts semantic dependency and similarity among users, tags and resources and leads to clustering folksonomy. It then uses the created clusters for recommendation purposes. To overcome the problem of uncontrolled vocabulary such as synonyms, acronyms, misspellings, etc., it uses external resources, including WordNet, Wikipedia and Google “did you mean” mechanism. Accuracy is the main problem in the field of recommendation systems, especially when we are facing with large volumes of data. Evaluation results show that the proposed method in resource recommendation is able to increase the average of F-Measure up to 10% and in tag recommendation up to 22% compared to the same tested method. Testing the proposed method with other methods and on other social resource sharing system databases or larger datasets with more evaluation criteria is one of the future goals. To enhance recommendations accuracy, this method can also be used on datasets with more dimensions, such as when the user location or time is intended.

REFERENCES

[1] H.N. Kim, A.T. Ji, I. Ha, G.S. Jo, “Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation”, Journal of Electronic Commerce Research and Applications, Vol. 9, pp. 73–83, 2010.

[2] J. Gemmell, T. Schimoler, B. Mobasher, R. Burke, “Resource recommendation in social annotation systems: A linear-weighted hybrid approach”, Journal of Computer and System Sciences, Vol. 78, pp. 1160–1174, 2012.

[3] D. Xu, Z. Wang, L. He, W. Huang, “A Simple and Heuristic Model of Tag Recommendation”, Fourth International Symposium on Computational Intelligence and Design (ISCID), pp. 162–165, 2011.

[4] J. Gemmell, M. Ramezani, T. Schimoler, L. Christiansen, B. Mobasher, “The impact of ambiguity and redundancy on tag recommendation in

folksonomies", in: Proceedings of the Third ACM Conference on Recommender Systems, pp. 45–52, 2009.

[5] ZK. Zhang, T. Zhou, and YC. Zhang, "Tag-aware recommender systems: a state of the-art survey", Journal of Computer Science and Technology - Special issue on Community Analysis and Information Recommendation, Vol. 26, pp. 767-777, 2011.

[6] M. E. Alper, S. G. Oguducu, "Personalized Recommendation in Folksonomies using a Joint Probabilistic Model of Users, Resources and Tags", 11th International Conference on Machine Learning and Applications, pp. 368-373, 2012.

[7] Z. I. Alepidou, K. N. Vavliakis, P. A. Mitkas, "A Semantic Tag Recommendation Framework for Collaborative Tagging Systems", IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing, pp. 633 – 636, 2011.

[8] K. Kolomvatsos, C. Anagnostopoulos, S. Hadjiefthymiades, "An efficient Recommendation System based on the Optimal Stopping Theory", Journal of Expert Systems with Applications, Vol. 41, pp. 6796–6806, 2014.

[9] R. Jaschke, A. Hotho, G. Stumme, "Tag Recommendations in Social Bookmarking Systems", Journal of AI Communications, Vol. 21, pp. 231-247, 2008.

[10] M. T. Hassan, A. Karim, F. Javed, N. Arshad, "Self-Optimizing a Clustering-based Tag Recommender for Social Bookmarking Systems", Ninth International Conference on Machine Learning and Applications, pp. 601-606, 2010.

[11] A. Nanopoulos, "Item Recommendation in Collaborative Tagging Systems", IEEE transactions on systems, man, and cybernetics—part a: systems and humans, Vol. 41, pp. 760-771.

[12] R. Wetzker, C. Zimmermann, C. Bauckhage and S. Albayrak, "I tag, you tag: Translating tags for advanced user models", Web Search and Data Mining conference. pp. 71-80, 2010.

[13] M. Rawashdeh et., "graph-based personalized recommedation in social tagging systems", IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp. 1-6, 2014.

[14] Y. Ustunbas, S. oguducu, "A Recommendation Model For Social Resource Sharing Systems Based on Tripartite Graph Clustering", European Intelligence and Security Informatics Conference, pp. 378-381, 2011.

[15] Z. Wang, Y. Tan, M. Zhang, "Graph-based Recommendationon Social Networks", 12th International Asia-Pacific Web Conference, pp. 116-122, 2010.

[16] M. Fan, Q. Zhou, T. F. Zheng, "Content-based Semantic Tag Ranking for Recommendation", IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, pp. 292-296, 2012.

[17] C. Trabelsi, B. Moulahi, S. B. Yahia, "HMM-CARE: Hidden Markov Models for context-aware tag recommendation in Folksonomies", Proceedings of the 27th Annual ACM Symposium on Applied Computing, pp. 957-961, 2012.

[18] H. Lin, J. Davis, and J. Zhou, "*An integrated approach to extracting ontological structures from folksonomies*", The Semantic Web: Research and Applications, pp. 654-668, 2009.

[19] S. Overell, B. Sigurbjornsson, and R. Zwol, "Classifying tags using open content resources", Proc. of the 2nd ACM Int. Conf. on Web Search and Data Mining (WSDM '09), pp. 64-73, 2009.

[20] T. Yoshida, U. Inoue, "A Bookmark Recommender System based on Social Bookmarking Services and Wikipedia Categories", 14th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, pp. 409-413, 2013.

[21] R. J¨aschke, A. Hotho, C. Schmitz, B. Ganter, G. Stumme, "Discovering shared conceptualizations in folksonomies", Journal of Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 6, pp. 38-53, 2008.

[22] M. Fahimi, M. Vafaei Jahan, M. Niazi Torshiz, "Discovery Of The Triadic Frequent Closed Patterns Based On Hidden Markov Model In Folksonomy",First International Congress on Technology, Communication and Knowledge (ICTCK 2014), pp. ,2014.

[23] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", Proceedings of the IEEE, Vol 77, pp. 257–286, 2002.

[24] J. Gemmell, M. Ramezani, T. Schimoler, L. Christiansen, and B. Mobasher, "The impact of ambiguity and redundancy on tag recommendation in folksonomies", in Proceedings of the third ACM conference on Recommender systems, pp. 45–52, 2009.

[25] R. Wetzker, C. Zimmermann, C. Bauckhage, S. Albayrak, "I tag, you tag: translating tags for advanced user models", in: Proceedings of the Third ACM International Conference on Web Search and Data Mining, pp. 71–80, 2010.

[26] C. Trablesi, A. B. Jerad, S. B. Yahia, "Bridging folksonomies and domain ontologies : Getting out non-taxonomic relations", IEEE International Conference on Data Mining Workshops, pp. 369-379, 2010.