WLNI-LPA: Detecting Overlapping Communities in Attributed Networks based on Label Propagation Process

Imen Ben El Kouni[1,2][a], Wafa Karoui[1,3][b] and Lotfi Ben Romdhane[1,2][c]

[1]*Université de Sousse, Laboratoire MARS LR17ES05, ISITCom, 4011, Sousse, Tunisia*

[2]*Université de Sousse, ISITCom, 4011, Sousse, Tunisia*

[3]*Université de Tunis El Manar, Institut Superieur d'Informatique, 2080, Tunis, Tunisia*

Keywords: Attributed Networks, Overlapping Community Detection, Node Similarity, Weighted Graph.

Abstract: Several networks are enriched by two types of information: the network topology and the attributes information about each node. Such graphs are typically called attributed networks, where the attributes are always as important as the topological structure. In these attributed networks, community detection is a critical task that aims to discover groups of similar users. However, the majority of the existing community detection methods in attributed networks were created to identify separated groups in attributed networks. Therefore, detecting overlapping communities using a combination of nodes attributes and topological structure is challenging. In this paper, we propose an algorithm, called WLNI-LPA, based on label propagation for detecting efficient community structure in the attributed network. WLNI-LPA is an extension of LPA that combines node importance, attributes information, and topology structure to improve the quality of graph partition. In the experiments, we validate the performance of our method on synthetic weighted networks. Also, a part of the experiment focuses on the impact of detecting significantly overlapping communities in the recommender system to improve the quality of recommendation.

1 INTRODUCTION

Network analysis has become a hot topic in recent years due to the rapid growth of real-world networks such as social networks (Leng and Jiang, 2016; Meo et al., 2014). It can found in a wide variety of contexts, for example, model friendships and acquaintances in a social context; in biology, networks capture metabolic processes in the organism (Garza and Schaeffer, 2019). A crucial task in network analysis is group identification, which is generally known as a group of nodes with large internal connections and minimal external connections. For example, in protein-protein interaction networks, communities refer to functional modules of interacting proteins. Thus far, a large number of community-detection algorithms have been proposed, and many of them have successfully addressed the various aspects of the community-detection issue. The attributed network is a common community detection framework scenario, in which nodes have attributes. In social networks, for example, the users can be characterized by several attributes such as gender, occupation, and hobbies. In attributed networks, community identification requires both network topology and nodes attributes examination (Huang et al., 2016). Many community detection studies, which only consider network structure details, are not adaptable for attributed networks. The label propagation algorithm is one of the most efficient community detection approaches where each node of the network is identified by a label. Exploring nodes attributes during the label propagation process is technically challenging because label propagation-based methods must assign one or more labels to each node of the network. In general, communities can be divided into two types: overlapping and non-overlapping. Many existing community detection algorithms in attributed network can solve the non-overlapping community detection problem (Zhou et al., 2009; Ruan et al., 2013). For that, studies start to develop overlapping community algorithms in the attributed networks to especially solve this problem (Sun et al., 2012; Xu et al., 2012). But, the problem is still open because these methods' applicability is

[a] https://orcid.org/0000-0003-3240-9647
[b] https://orcid.org/0000-0002-5311-0655
[c] https://orcid.org/0000-0003-2163-5809

El Kouni, I., Karoui, W. and Ben Romdhane, L.
WLNI-LPA: Detecting Overlapping Communities in Attributed Networks based on Label Propagation Process.
DOI: 10.5220/0010605904080416
In *Proceedings of the 16th International Conference on Software Technologies (ICSOFT 2021)*, pages 408-416
ISBN: 978-989-758-523-4

restricted which means that are not able to handle all types or size of networks.

In this paper, we propose an overlapping community detection algorithm based on label propagation process, called WLNI-LPA for attributed networks. This algorithm tackles the issue of analyzing attributed networks using label propagation. WLNI-LPA combines the network's topology information and node attributes to ameliorate the accuracy of detection. Our method is an extension of NI-LPA algorithm to detect overlapping communities for attributed networks. The original NI-LPA allows a node to contain more than one label and can conveniently leverage these labels in the label propagation step. In addition to the propagation of labels, the idea of this work is to improve the coefficient of each label by attributes information inherited from nodes. Moreover, WLNI-LPA requires neither a pre-defined objective function nor prior information about the size of communities.

The following is a list of our major contributions:

- We propose an extension of the NI-LPA overlapping community detection algorithm, which considers both the topology of the network and attribute information, to uncover communities in attributed networks.
- We combine the node attributes information, the node importance, and the multi-label propagation by weighting the edges to add new knowledge to identify overlapping communities in attributed networks.
- The tests in synthetic networks show that our proposed WLNI-LPA ameliorates the detection process. Also, an application in the recommender system area shows the efficiency of our approach to detect overlapping communities and improve the quality of recommendation.

2 RELATED WORK

In this section, we introduce the community detection issue and give a classification of the different approaches. Then, we concentrate on the methods based on the label propagation process.

2.1 Community Detection

One of the most important concepts in complex network research is community detection. It helps to analyze networks and understand behavior grouping. Community detection is to group nodes into different groups, where nodes in the same group are more linked with each other than with nodes outside the group. This area consists of creating cohesive groups within networks, groups of friends in social networks, or groups of web pages with the same subject in knowledge networks. Other tasks are made possible by identifying such groups like marketing (Weng et al., 2013), recommendation (Zheng and Wang, 2018; Zheng et al., 2019) and document organization (Hachaj and Ogiela, 2017; Grineva et al., 2009). In recent years, researchers have proposed a variety of methods for identifying communities. Newman and Girvan proposed a quality function called modularity and optimized it for community detection (Newman and Girvan, 2004). Since then, modularity-based approaches that maximize the modularity value have been commonly used to detect groups. However, Lancichinetti and Fortunato (Lancichinetti and Fortunato, 2011) demonstrated that the modularity measure has a limited resolution and suffers extreme degradation. Also, they propose a local fitness maximization approach for community detection. Lancichinetti *et al.* (Lancichinetti et al., 2009) exploit the hierarchy of communities to apply a fitness maximization method for overlapping communities' detection. In random walk methods, the concept is based on the fact that walks appear to be trapped in densely areas of a network (Pons and Latapy, 2005). The most common random walk approach is Infomap, which is a flow-based group detection algorithm that combines information-theoretic techniques and random walks (Rosvall and Bergstrom, 2008). Label propagation is an intensively studied issue in the field of community detection proposed by Raghavan *et al.* (Raghavan et al., 2007). This approach may be appropriate for partitioning large networks in real time.

2.2 Label Propagation based Community Detection

Label propagation algorithm (LPA) is a popular and fast method for community detection. Initially, every node is identified with unique labels. The next step is for each node to update its label with the label that is most frequently used by its neighbors. When many neighbor labels are similarly frequent, it selects at random from the most frequent labels. The process of label propagation is repeated until all nodes with the same label are grouped into a single community. LPA is a simple, unsupervised near-linear algorithm with no parameters that require no previous information about the size and number of communities. Since there is a random factor in LPA when there is more than one most frequent label, different results can be obtained after several runs. Various improvements to LPA have been applied in recent years, in order to enhance its stability and robustness. COPRA (Gregory, 2010) algorithm with each vertex is assigned by pairs (c, b), where c is a community identifier and b is a belonging coefficient. BMLPA (Wu et al., 2012) requires that one vertex's community identifiers have balanced belonging coefficients, allowing nodes to belong to any number of communities without imposing a high number set by COPRA. SLPA (Xie et al., 2011) employs a dynamic speaker-listener interaction mechanism to maintain that each node can own several labels. An extension of SLPA called WLPA (Hu, 2013) which improves it by adding a similarity between any two nodes focused on the labels they got during label propagation. LPA-S (Li et al., 2017) is proposed using label propagation and similarity to develop a stepping community detection algorithm. LP-LPA (Berahmand and Bouyer, 2018) algorithm improves LPA by computing the link strength and node's label influence values and it processes label updating according to the highest label influence value among neighbors. AntLP (Hosseini and Rezvanian, 2020) is an improved version of LPA that assigns weights for edges based on several similarity indices, then, uses ant colony optimization to propagate labels and optimize modularity measure. NI-LPA (El Kouni et al., 2020) is an extension that adopts LPA strategy to allow a node to contain a set of labels and simulates a special propagation and filtering process using information deduced from the structural properties of nodes. On the other hand, these label propagation-based methods are unable to address attributed networks because they all ignore nodes attributes information during the propagation process. However, our work uses the structural and semantic attribute information of nodes, making it well adapted to attribute networks.

3 PROPOSED METHOD

In this section, the problem definition has been described. Then, we define the proposed weighted algorithm called WLNI-LPA to detect overlapping communities in weighted networks.

3.1 Problem Definition

Given a graph $G = (V, E, L, W)$ be an attributed network. $V = \{v_1, v_2, ..., v_n\}$ is the set of n nodes, $E \sqsubseteq V * V$ is the set of edges, $L = \{l_1, l_2, ..., l_s\}$ is the set of nodes attributes, and W is edge weight between two nodes. Overlapping community detection is to partition the network G into k communities $C = \{c_1, c_2, ..., c_k\}$ that satisfy the following criteria:

- Based on network structure, the nodes in the same group are closely connected, while the nodes in different communities are sparsely connected.
- Based on nodes attributes, nodes in the same group have similar attribute values, while nodes in different communities have different values.
- Based on overlap concept, node can allow to more than one community $c_i \bigcap c_j \neq \emptyset$

3.2 Contribution

In our work, we propose to improve the NI-LPA algorithm to be able to detect overlapping communities in attributed networks. On one hand, NI-LPA focuses on both topology information and the role or importance of the node in the network. On the other hand, this algorithm maintains the simplicity of the original LPA and obtains accurate results in large and complex networks. For that, we show that NILPA is an appropriate method to become adaptable for attributed networks with large size, diverse attribute nodes, and different typologies of the graph.
As for any algorithm of detection in attributed networks, the main topic is the use comprehensively of structural information and nodes attributes together. For example, in social networks, the profiles of users can be regarded as nodes attributes. As illustrated in Figure 1, each user is identified by his age, gender, and occupation.

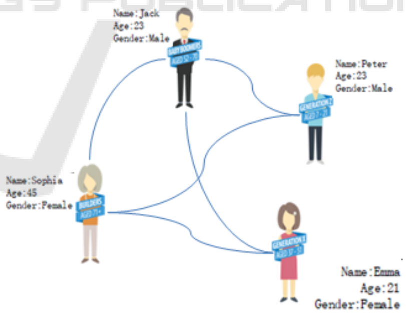

Figure 1: Social network with nodes attributes information.

Therefore, the proposed algorithm transforms the nodes attributes information to a weight of edge to compare the node in terms of semantic similarity. In fact, this method needs a step of weighted network construction. Then, the propagation phase will be improved to take into consideration both the node importance and the link weighting. The proposed algorithm is able to detect overlapping communities in attributed networks.

3.3 Weighted Graph Construction

To integrate the information of nodes attributes, we propose to transform the original unweighted network to weighted networks where the weights of links represent the similarity between nodes based on attributes information. Therefore, we use the attributes of each two node to calculate the similarity values between a pair of nodes. Then, affect this measure as a weight of the link which connects these nodes. In this work, we define the similarity between node u and node v as the following:

$$Sim(u,v) = \frac{\sum_{i=1}^{Nb} Sim(u,v)}{Nb} \tag{1}$$

where Nb indicates the number of attributes.

Figure 2 shows an explication of graph construction algorithm to transform attributed graph to a weighted network. In this example, the attributes information is the tags put by the film viewer. This example demonstrates that the weighted network is easier to use in the propagation process.

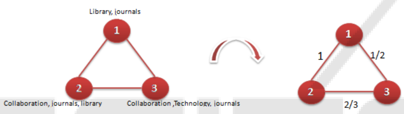

Figure 2: Transformation from attributed graph to weighted network.

3.4 Weighted NI-LPA

We propose an efficient algorithm which is an extension of NI-LPA (El Kouni et al., 2020) to detect overlapping communities in the attributed network. In fact, the proposed algorithm, called WLNI-LPA, is described in Algorithm 1. The step of weighted network construction takes an unweighted attributed network as input and creates a new weighted network using the nodes attributes information to give link weight as a similarity measure between two nodes. For That, our algorithm takes as input the weighted network and discovers a set of overlapping communities. It consists of three main components: initialization, propagation, and filtering. Initially, we consider each node as a community. The concept of overlap means that each vertex may belong to more than one community. For that, to find overlapping communities, we must allow a node to contain many community identifiers. In fact, a node is characterized by its label, its topological importance, and its connections with the others node. Similar to NI-LPA, we calculate the importance of each node based on its degree and coefficient of clustering. Besides, the connection of this node and its neighbors is assigned as a weight to the link that connects with another node. The higher the node importance and the link weight is, the higher this label will be dominant in the propagation phase. In the label propagation process of Lines 10–18, we iterate the nodes labels Nb times. For each iteration i, we update the labels of all nodes based on the labels at previous iteration and the labels of its neighbors updated in the current iteration. The node that is being processed receives a set of labels from its neighbors and aggregates the scores of the labels associated with the neighbors such as line 13 of the algorithm. At the end of the propagation phase, each node contains a set of labels with their scores. But, some of these labels are poor in contrast to the other labels' coefficients. For that, lines 19-23 describe the simple filtering process. In this stage, we compare the coefficient of each node with the threshold and remove those who are less than it. The threshold is fixed to 0.4 which means that any coefficient less than 0.4 is considered poor.

4 EXPERIMENTS

This section reports experimental settings and results. On the one hand, we evaluate the proposed algorithm in synthetic weighted networks with different settings. On the other hand, we apply this algorithm to evaluate its effectiveness to detect groups of similar users in order to improve the quality of recommendation.

4.1 Experiments on Synthetic Networks

To evaluate the efficacy of our proposed WLNI-LPA, we carried out experiments on artificial networks using LFR benchmarks (Lancichinetti and Fortunato, 2009) to generate weighted networks with overlapping communities. The LFR benchmark enables the generation of weighted networks with power-law node degree and community size distributions. For the accuracy evaluation, we generate a group of the dataset where N equal to 5000 and we varied the edge weight and mixing parameter for topology. The parameter settings are shown in Table 1.
In addition, we compare the performance of our community detection algorithm (denoted as WLNI-LPA) with three algorithms which can detect overlapping communities for weighted networks.

- Conductance (Lu et al., 2014): characterizes nodes in communities based on two metrics intra-centrality and inter-centrality.
- COPRA (Gregory, 2010): detects communities

Algorithm 1: Weighted NI-LPA.

```
Data: A Weighted network G, number of
      iterations Nb
Result: Overlapping communities
Map ← a new list with n empty dictionaries
 foreach e in E do
    u ← e.sourceNode
    v ← e.targetNode
    w ← e.weight
    Initialize each node with a unique label
      C_x = x
    Calculate importance of all the nodes
    Sort nodes according to their degree in
      descending order
    for i=1 To Nb do
        foreach label in List(u) do
            if label is in List(v) then
                List(u).label ←
                  List(v).label + w *
                  importance(v)
            else
                List(u).label ←
                  w * importance(v)
            end
        end
    end
end
Map ← List(u)
foreach Label with coefficient b in Map do
    if b < Threshold then
        delete label
    end
end
return Map
```

on weighted networks based on label propagation process.

- Strength (Chen et al., 2010): exploits belonging degree and node strength to detect the overlapping community structure.

Since it is hard to verify the detected communities for real networks, the evaluation of these methods is based mainly on synthetic networks.
As metric, we use the normalized mutual information NMI (Lancichinetti et al., 2009) defined by Equation 2 to compare communities with ground-truth partition. The higher NMI values mean good partition.

$$I_{norm}(X:Y) = \frac{H(X)+H(Y)-H(X,Y)}{(H(X)+H(Y))/2} \tag{2}$$

where H(X), (H(Y)), is the entropy of the random variable X, (Y), assigned to the partition $C\prime$, $(C\prime\prime)$.

Threshold Definition. In the filtering step, the coefficient of each node is compared with the threshold and remove those who are less than it. The choice of this value is clearly explained by the fact that any coefficient less than this threshold is considered poor. As a consequence, in this section, we will attempt to test distinct values experimentally. The parameters of networks are as follows: $mut = 0.2$; $muw = 0.2$; $k = 10$; $O_m = 2$; $O_n = 10\%$. Table 2 reports the results. We can conclude that our algorithm gives the best value of NMI when the threshold is 0.4, and the quality of detection decreases when we consider another value.

Evaluation of WLNI-LPA with Different Settings. Figure 3 presents the NMI values given by our algorithm with the different average degrees on the LFR networks. We test WLNI-LPA with three categories: when the mixing parameter *mut* equal to 0.2; 0.3; and 0.5. As the value of *mut* increases, the network becomes much more complex, and the boundaries between communities become more unclear. We show in this figure that all the values are between 0.96 and 1 which means that WLNI-LPA successfully provides a partition very close to the exact partition. This is mainly because our algorithm collects more and more useful information (degree measure, clustering coefficient, and node attributes) about the node to propagate labels.

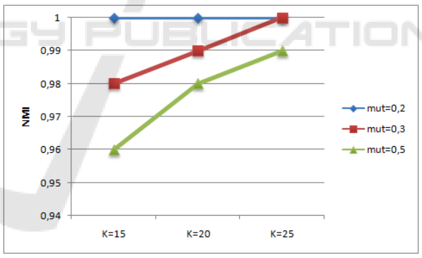

Figure 3: The NMI measures as a function of the average degree in a weighted network with 5000 nodes.

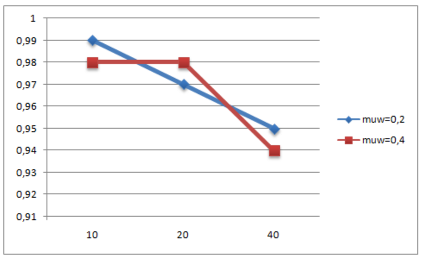

Figure 4: The NMI measures as a function of the percentage of the overlapping node in a weighted network with 5000 nodes.

As illustrated in Figure 4, we test our method with two constraints which are the percentage of overlapping nodes from 10% to 40% and the variation of

Table 1: Parameter settings of benchmarks.

Param	value	Description
N	5000	Number of nodes
k	15 - 25	Average node degree
maxk	50	Max node degree
mut	0.2 - 0.5	Mixing parameter for topology
muw	0.2 - 0.5	Mixing parameter for edge weight
minc	20	Minimum for community sizes
maxc	50	Maximum for community sizes
On	10-40	Number of overlapping nodes
Om	2	Number of communities of overlapping nodes

Table 2: NMI results for test with different threshold.

Network	0.2	0,3	0,4	0,5	0,6
5000	0,81	0,9	1	0,96	0,72

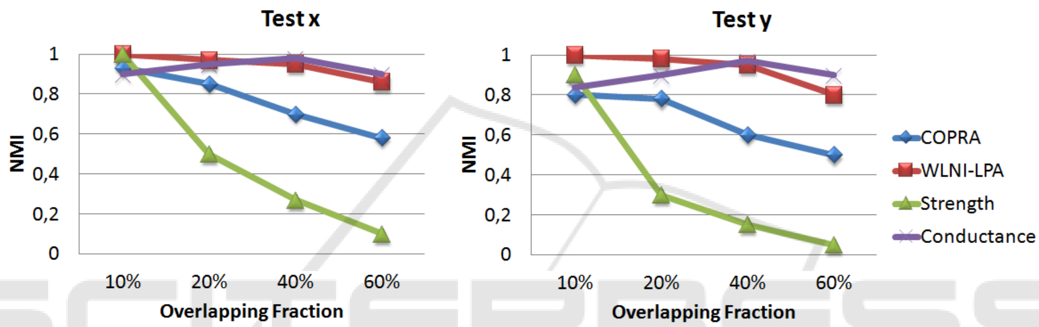

Figure 5: The quality in terms of NMI of the community structure detected by Conductance (Lu et al., 2014), COPRA (Gregory, 2010), Strength (Chen et al., 2010) and WLNI-LPA for different parameter settings.
Test x: N=5000, $\mu_w = 0.1$, $\mu_t = 0.1$
Test y: N=5000, $\mu_w = 0.3$, $\mu_t = 0.3$.

the mixing parameter for edge weight. The results demonstrate that WLNI-LPA detect overlapping communities in all the cases. This indicates that our algorithm can propagates labels by considering topological and attributes information and then, filter useless labels to find more accurate communities with different numbers of overlapping nodes.

Figure 5 shows the results of our algorithm compared to three other algorithms from the literature for different parameter settings. We can see the performance all these approaches degrade dramatically when the overlapping fraction increase from 10% to 60%. The NMI values given by WLNI-LPA indicate that it is able to detect good partition in artificial weighted networks compared to other methods. WLNI-LPA has a very stable performance compared to strength and COPRA even with the increase of overlapping nodes fraction. However, WLNI-LPA and conductance give closer partitions to exact ones.

4.2 Application to Social Recommender System

Recently, the socialized recommendation has become one of the most common methods of recommendation in a variety of recommender systems used in areas such as e-commerce, social media sites, and web search engines (Ikeda et al., 2013). Integrating detection community algorithm in social recommender system becomes an important challenge studied by many authors (Gasparetti et al., 2020; Boussaadi et al., 2020; Ai et al., 2019). In social recommender systems, it's crucial to find users that have similar interests in order to provide recommendations. To address these problems, user profile techniques are used to reflect users' interests and detect similar users. Therefore, integrating the community detection process in the recommender system can enhance recommendations. In this paper, since WLNI-LPA is a community detection algorithm specific to attributed networks, we use it as a step in the recommender system algorithm to detect similar users. To evaluate

Table 3: Datasets used in recommender system.

Datasets	Users	Items	R-scale	demographic information
Flickr	1000	500	1-5	age, location, gender
Book crossing	2000	1000	1-10	age, location

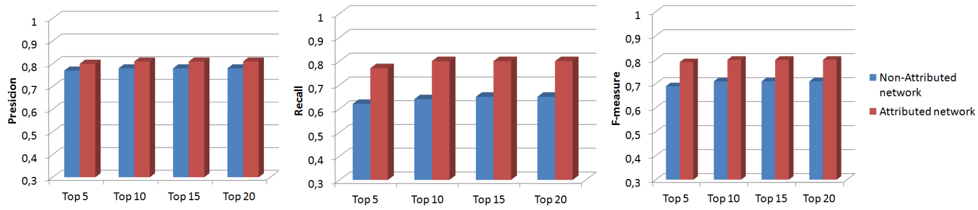

Figure 6: Performance comparison (Precision, Recall, and F-measure) on attributed and non-attributed networks using Flickr datasets.

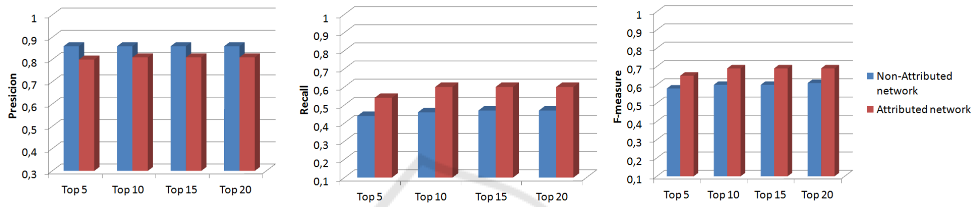

Figure 7: Performance comparison (Precision, Recall, and F-measure) on attributed and non-attributed networks using Book crossing datasets.

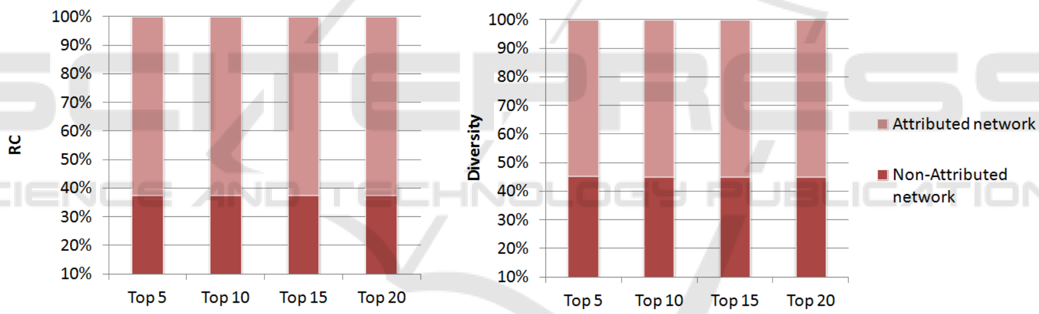

Figure 8: Performance comparison (Diversity, rate Coverage) on attributed and non-attributed networks using Book crossing datasets.

the performance of this method using WLNI-LPA, we use two known datasets in the recommender system field and described in Table 3. Flickr is an image-sharing network in which nodes represent users and links represent relationships between users. Book-crossing contains a set of users (with demographic information) providing ratings about a set of books.

We used precision, recall, F1 measure, diversity, and rate coverage (RC) to evaluate the performance of recommendation.
Let TP is the recommendations generated by the algorithm that users like defined as True Positive, and the others those not like by the users are defined as False Positive (FP). The items that are not recommended and the uses not like are defined as True Negative (TN), and those not recommended by the algorithm but liked by users are defined as False Negative (FN).

These metrics are calculated as follows:

$$Prec = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{TP}{TP+FP}\right) \tag{3}$$

$$Rec = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{TP}{TP+FN}\right) \tag{4}$$

$$F-measure = 2*\frac{Prec*Rec}{Prec+Rec} \tag{5}$$

$$Diversity = \left| \bigcup_{i=1,..,n} L_N(i) \right| \tag{6}$$

where $L_N(i)$ is the Top-N items in the recommendation list for the i^{th} user, and n is the number of users.

$$RC = \frac{number\ of\ predicted\ ratings}{number\ of\ all\ ratings} \tag{7}$$

The objective of this recommender system is to recommend items to users based on the similarity between them. For that, we apply our algorithm WLNI-LPA to detect a group of similar users with overlap. Thus, improving the algorithm of detection means enhancing the quality of recommendation.

In fact, the demographic information of users will be considered as nodes attributes. To demonstrate how the nodes attributes affect the quality of recommendations, we use our proposed algorithm WLNI-LPA in the social recommendation process with and without attributes also to see how well it utilizes the attribute information in the label propagation.

Figure 6 and 7 recapitulate the experimental results in both datasets. The results show that running WLNI-LPA on an attributed network achieves great promotion in the community detection task in all cases (Top 5 to 20) and even in two datasets with different densities. The scores given when using WLNI-LPA in the attributed network is always better than in non attributed network in terms of quality of results which means that the attributed information improves the quality of detecting similar users and can bring satisfaction about recommended items. In the other part, the test on Book-crossing datasets with considering attributed network achieves maximum scores of diversity and RC than non attributed network. The good partition of the network in communities affects the result of predicted rating produced by a recommender system.

5 CONCLUSION

In this paper, we examined the community detection analyses in attributed graphs. We propose WLNI-LPA, an accurate method for detecting overlapping communities in attributed networks as only a few works are considering the overlap concept. Our method combines the nodes attributes with the network topology based on the label propagation process to improve the quality of graph partition. WLNI-LPA transforms the attributed network to a weighted graph where the attributes information will be represented as link weight between nodes. The experimental tests are divided into two types: firstly, in an artificial weighted network and secondly by integrating WLNI-LPA in a recommender system to detect similar users. The results show that our algorithm can effectively discover the overlapping communities on one hand and improve the quality of recommendation on the other hand.

Future work can be conducted in various directions. First, to improve the accuracy by optimizing the weights of node attributes. Second, to propose temporal similarity measurements which consider the time factor to detect communities in dynamic attributed networks. Third, since real-life problems models use dynamic real-world networks graphs to represent entities and the relations between them, we must study these dynamic networks evolution. As a solution, on one hand, we can add some information into group detection problems that are not intrinsic to the graph structure; on the other hand, we can consider the ability of the members to communicate among them. Indeed, the community would be more homogeneous if we take into consideration different exchanges between members like public or private messages, videos, photos, hypertext links, or games.

REFERENCES

Ai, J., Liu, Y., Su, Z., Zhang, H., and Zhao, F. (2019). Link prediction in recommender systems based on multi-factor network modeling and community detection. *EPL (Europhysics Letters)*, 126(3):38003.

Berahmand, K. and Bouyer, A. (2018). Lp-lpa: A link influence-based label propagation algorithm for discovering community structures in networks. *International Journal of Modern Physics B*, 32(06):1850062.

Boussaadi, S., Aliane, H., Abdeldjalil, O., Houari, D., and Djoumagh, M. (2020). Recommender systems based on detection community in academic social network. In *2020 International Multi-Conference on:"Organization of Knowledge and Advanced Technologies"(OCTA)*, pages 1–7. IEEE.

Chen, D., Shang, M., Lv, Z., and Fu, Y. (2010). Detecting overlapping communities of weighted networks via a local algorithm. *Physica A: Statistical Mechanics and its Applications*, 389(19):4177–4187.

El Kouni, I. B., Karoui, W., and Romdhane, L. B. (2020). Node importance based label propagation algorithm for overlapping community detection in networks. *Expert Systems with Applications*, 162:113020.

Garza, S. E. and Schaeffer, S. E. (2019). Community detection with the label propagation algorithm: A survey. *Physica A: Statistical Mechanics and its Applications*, 534:122058.

Gasparetti, F., Sansonetti, G., and Micarelli, A. (2020). Community detection in social recommender systems: a survey. *Applied Intelligence*, pages 1–21.

Gregory, S. (2010). Finding overlapping communities in networks by label propagation. *New journal of Physics*, 12(10):103018.

Grineva, M., Grinev, M., and Lizorkin, D. (2009). Extracting key terms from noisy and multitheme documents. In *Proceedings of the 18th international conference on World wide web*, pages 661–670.

Hachaj, T. and Ogiela, M. R. (2017). Clustering of trending topics in microblogging posts: A graphbased approach. *Future Generation Computer Systems*, 67:297–304.

Hosseini, R. and Rezvanian, A. (2020). Antlp: ant-based label propagation algorithm for community detection in social networks. *CAAI Transactions on Intelligence Technology*, 5(1):34–41.

Hu, W. (2013). Finding statistically significant communities in networks with weighted label propagation.

Huang, X., Cheng, H., and Yu, J. X. (2016). Attributed community analysis: Global and ego-centric views. *IEEE Data Eng. Bull.*, 39(3):29–40.

Ikeda, K., Hattori, G., Ono, C., Asoh, H., and Higashino, T. (2013). Twitter user profiling based on text and community mining for market analysis. *Knowledge-Based Systems*, 51:35–47.

Lancichinetti, A. and Fortunato, S. (2009). Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Physical Review E*, 80(1):016118.

Lancichinetti, A. and Fortunato, S. (2011). Limits of modularity maximization in community detection. *Physical review E*, 84(6):066122.

Lancichinetti, A., Fortunato, S., and Kertész, J. (2009). Detecting the overlapping and hierarchical community structure in complex networks. *New journal of physics*, 11(3):033015.

Leng, J. and Jiang, P. (2016). Mining and matching relationships from interaction contexts in a social manufacturing paradigm. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47(2):276–288.

Li, W., Huang, C., Wang, M., and Chen, X. (2017). Stepping community detection algorithm based on label propagation and similarity. *Physica A: Statistical Mechanics and its Applications*, 472:145–155.

Lu, Z., Sun, X., Wen, Y., Cao, G., and La Porta, T. (2014). Algorithms and applications for community detection in weighted networks. *IEEE Transactions on Parallel and Distributed Systems*, 26(11):2916–2926.

Meo, P. d., Ferrara, E., Abel, F., Aroyo, L., and Houben, G.-J. (2014). Analyzing user behavior across social sharing environments. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1):1–31.

Newman, M. E. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113.

Pons, P. and Latapy, M. (2005). Computing communities in large networks using random walks. In *International symposium on computer and information sciences*, pages 284–293. Springer.

Raghavan, U. N., Albert, R., and Kumara, S. (2007). Near linear time algorithm to detect community structures in large-scale networks. *Physical review E*, 76(3):036106.

Rosvall, M. and Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123.

Ruan, Y., Fuhry, D., and Parthasarathy, S. (2013). Efficient community detection in large networks using content and links. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1089–1098.

Sun, Y., Aggarwal, C. C., and Han, J. (2012). Relation strength-aware clustering of heterogeneous information networks with incomplete attributes. *arXiv preprint arXiv:1201.6563*.

Weng, L., Menczer, F., and Ahn, Y.-Y. (2013). Virality prediction and community structure in social networks. *Scientific reports*, 3(1):1–6.

Wu, Z.-H., Lin, Y.-F., Gregory, S., Wan, H.-Y., and Tian, S.-F. (2012). Balanced multi-label propagation for overlapping community detection in social networks. *Journal of Computer Science and Technology*, 27(3):468–479.

Xie, J., Szymanski, B. K., and Liu, X. (2011). Slpa: Uncovering overlapping communities in social networks via a speaker-listener interaction dynamic process. In *2011 ieee 11th international conference on data mining workshops*, pages 344–349. IEEE.

Xu, Z., Ke, Y., Wang, Y., Cheng, H., and Cheng, J. (2012). A model-based approach to attributed graph clustering. In *Proceedings of the 2012 ACM SIGMOD international conference on management of data*, pages 505–516.

Zheng, J., Wang, S., Li, D., and Zhang, B. (2019). Personalized recommendation based on hierarchical interest overlapping community. *Information Sciences*, 479:55–75.

Zheng, J. and Wang, Y. (2018). Personalized recommendations based on sentimental interest community detection. *Scientific Programming*, 2018.

Zhou, Y., Cheng, H., and Yu, J. X. (2009). Graph clustering based on structural/attribute similarities. *Proceedings of the VLDB Endowment*, 2(1):718–729.