Hybrid Community Detection Approach in Multilayer Social Network: Scientific Collaboration Recommendation Case Study

Wala Rebhi
RIADI laboratory,
National School of Computer Sciences,
University of Manouba, Tunisia
Email: wala.rebhi@gmail.com

Nesrine Ben Yahia
RIADI laboratory,
National School of Computer Sciences,
University of Manouba, Tunisia
Email: Nesrine.benyahia@ensi.rnu.tn

Narjès Bellamine Ben Saoud
RIADI laboratory,
National School of Computer Sciences,
University of Manouba, Tunisia
Email: Narjes.Bellamine@ensi.rnu.tn

***Abstract*—Within real-world social networks people are linked with multiple types of relationships, which brings new challenges in community detection for multilayer social network where each layer represents one type of relationships. However, most of existing approaches consist on transforming the problem into a classical problem of community detection in monoplex network. In this work, we propose a new hybrid community detection approach in multilayer social networks. This approach considers simultaneously the network structure (different social connections) and the homophily of participants (similarities between users). To do so we propose a new multiplex information graph model to represent multilayer social network. Then, we adapt a combined community detection algorithm to the multiplex case. Furthermore, an example in the field of scientific collaboration recommendation is given to illustrate the practical usefulness of the proposed approach. Finally, a comparison with other community detection approaches evaluates its performance.**

I. Introduction

With the overwhelming explosion of online social networks in recent years such as Facebook[1], Twitter [2], ResearchGate[3], DBLP [4] and Linkedin[5], a huge number of users have become highly dynamic and they are continually seeking new collaborators to form communities. In this context, social networks analysis in particular community detection has been a main objective to identify and characterize relevant communities (groups of people) who participate in the creation of best practices and in the generation of innovative solutions through expertise sharing and socialization [1].

However, within real-world social networks, people are linked with different types of ties: friendship, family relationship, professional relationship, etc. Thus, the concept of multilayer social network has been introduced with the goal to account for this relationships multiplicity [2]. Then, multiplex graph has been proposed to provide an expressive model for representing a multilayer social network [3]. In this context, [4] defines a multiplex graph, named also multi-slice graph or multi-relational graph or multi-layer graph [5], as a multilayer graph composed by the same set of nodes but interconnected by different types of links. In fact, each layer contains exactly the same set of nodes. But the nature of links varies from one layer to another.

Therefore, great efforts have been made to solve the challenge of community detection in multilayer social networks [6]. Yet, most of the proposed approaches until recently transform this problem into a problem of community detection in monoplex networks where one relationship can be represented [7]. And even the few approaches, which consist on extending an existing algorithm to deal directly with multiplex graph, have considered only the structural aspect to detect communities [8].

Thus, in this work we propose a hybrid community detection approach to find and form dense communities in multilayer social networks. This approach considers simultaneously the network structure (different social connections) and the homophily of participants (similarities between users). To do so we propose a new multiplex information graph model to represent multilayer social network. Then, we adapt a combined community detection algorithm to the multiplex case.

Our paper is structured as follows. First, we introduce basic definitions and used notations related to community detection in social networks. In section 3, we present an overview of different community detection approaches in multilayer social networks. Then, we describe the proposed hybrid community detection approach and illustrate its practical usefulness via a real example in the field of scientific collaboration recommendation. In section 5, we compare the performance of the proposed solution with other community detection approaches. In the last section, we conclude and present an outlook on future work.

II. Definitions and notations

Community detection is presented by [9] as a technique for partitioning networks (in particular social networks) into communities with weak coupling (external relations among

[1]https://www.facebook.com/
[2]https://twitter.com/
[3]https://www.researchgate.net/
[4]dblp.uni-trier.de/
[5]https://www.linkedin.com/

978-1-5090-4320-0/16/$31.00 ©2016 IEEE

communities) and high cohesion (internal relations within communities).

Besides, as social network is a complex system [10], modeling it is a key issue for the process of community detection. Thus, most widespread community detection approaches consider the social network as a graph and then analyze its structure with graph properties and algorithms built around the graph structure [11]. Even more, in some very influential works in the literature such as [12], the terms "graph" and "network" are used interchangeably [13]. In fact, graph is a powerful mathematical abstraction for representing entities (i.e., actors in social network) and their relationships [14].

In this context, most community detection algorithms, such as [15], [16], [17] and [18], consider a social network as a monoplex graph [11]. A monoplex graph, named also single-layer graph [6], is defined as a tuple G = (V, E) where V is the set of vertices or nodes representing individuals and E is the set of edges that connect pairs of nodes [2]. In Figure 1, for example, we have represented a monoplex graph with four nodes and three edges (without specifying the weights). This may correspond to a portion of a research laboratory, where nodes represent researchers and edges represent the co-authoring relationships among these researchers. Weights can be used to represent the number of common publications.

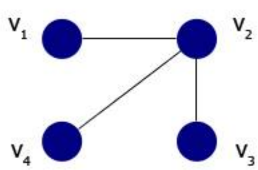

Fig. 1. An example of monoplex graph

With the emergence of Web 2.0 and digital networks, the concept of the social network has to be generalized to account for features describing the actors of the network and their relationships. This led to the definition of new concepts such as information graph by [19], Attribute augmented graph [20] or homogeneous or heterogeneous information network by [21]. Thus, [22] defines an information network as a graph where each node is described by data that can be structured or unstructured. For example, microblogging site can be considered as an information network where each user can bind friendships with others and he is described through a short biography (textual content), his age (digital vector) or his interest centers (a list of labels). Simultaneously, other community definition based on vertex similarity has been proposed. Indeed, [23] defines communities as groups of vertices that probably share common properties and/or play similar roles within the graph.

However, these graph-based models do not deal with the multiple aspects of relationships within social network. Thus, in some recent works of community detection, another issue arises around the multiplexe graph [6]. Formally, a multiplex graph, as illustrated by Figure 2, is defined as a tuple $G = (V, E)$ where V is a set of nodes, $E = \langle E_1, E_2, ..., E_K \rangle$ is a set of K types of edges between nodes in V; with $E_l = \{(V_i, V_j), i \neq j, V_i, V_j \in V\}; \forall l \in \{1..K\}$ [4].

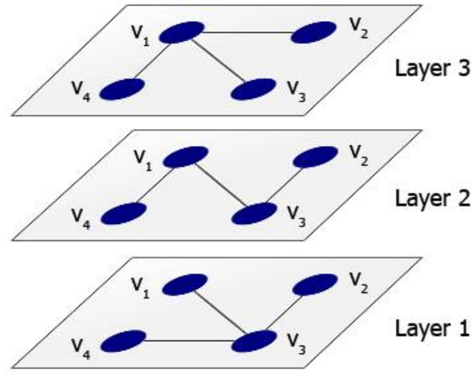

Fig. 2. An example of multiplex graph representing a 3-layer social network

In the next section we will present a brief survey of community detection approaches in multilayer social networks using multiplex graph.

III. Overview of different community detection approaches in multilayer social networks

Different approaches have been recently proposed to cope with community detection problem in the context of multilayer social networks [2]. We can classify these existing approaches into two broad classes:

- Network aggregation approach
- Multilayer network exploration approach

Next we detail both approaches.

A. Network aggregation approach

The idea is to transform a multiplex graph in a monoplex graph using an integration strategy of the different layers. Then, classical community detection algorithms can be applied to the resulting graph such as the Girvan-Newman algorithm [24], the Walktrap algorithm [25], the Louvain algorithm [16], etc. Differents ways to construct an aggregated network (which is also known as a superposition network [26], overlapping network [27], or overlay network [28]) have been proposed. For example, [29] propose to define edge weights between two nodes in the resulting monoplex network as a linear combination of the weights between those same nodes from each of the layers. As for [30], they propose to weight a link between two nodes by the average of weights of these two nodes' links in all layers in the multiplex graph. [31] and [32] propose to use a binary weights. In fact, two nodes are linked in the resulting monoplex graph if it exists at least one layer in the multiplex graph where these nodes are linked.

B. Multilayer network exploration approach

The idea is to integrate the social network multilayer aspect in the community detection process. Few studies have addressed the problem of simultaneous exploration of all multiplex network layers for the community detection. In this context, a generalized modularity function is proposed in [8]. With this new modularity, classical approaches for modularity maximization can be applied directly to multiplex networks. For example, an inspired version of the Louvain algorithm has been proposed using the generalized modularity [33]. Recently, the Infomap algorithm [34] has also been extended to the multiplex case [35]. In [36] an adaptation of the Walktrap community detection algorithm [25] is proposed. A seed-centric approach is also proposed in [7]. These adapted solutions rely on the network structure to detect communities in multilayer networks.

Inspired by these solutions, this paper aims to adapt a combined community detection algorithm to the multiplex case. However, this combined algorithm does not consider only the structural aspect like the previous presented algorithms; it also deals with the homophily of participants (similarities between users). This is why, we need a new graph-based model to represent multilayer social network which considers simultaneously the network structure (different social connections) and various members' profiles so as to calculate similarities between "nodes". This graph will be the input of the combined community detection algorithm.

In the next section, we present these different steps of the proposed hybrid approach and we illustrate its practical usefulness via a real example in the field of scientific collaboration recommendation.

IV. Hybrid community detection approach in multilayer social network

In this paper, our aim is to propose a new approach to deal with hybrid community detection in multilayer social networks. As a use case, we propose to apply this approach to support scientific collaboration recommendation in multilayer social networks.

For example, we can consider a research laboratory as a multilayer social network.

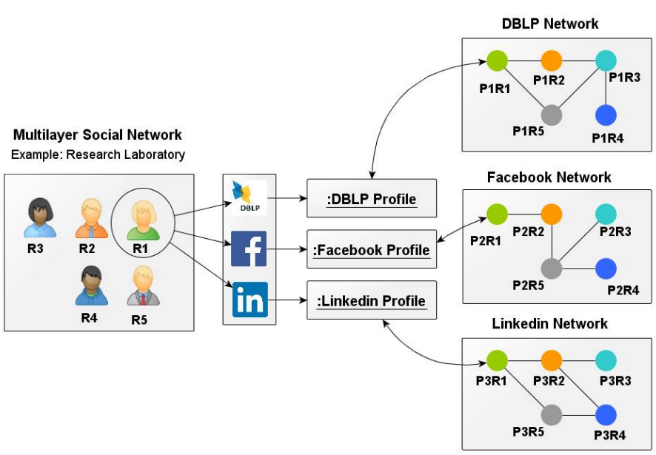

Fig. 3. Research laboratory as a multilayer social network

As illustrated by Figure 3, within this social network there are different types of relationships between researchers. Indeed, the same members can be connected by a co-publication relationship on DBLP, a friendship on Facebook, or a professional relationship on Linkedin. In addition, each member has a specific profile describing him in each social network: DBLP profile, Facebook profile and Linkedin profile.

When a researcher has a problem and he is seeking relevant collaborators to assist him, our aim is to recommend for him rapidly the most relevant group by detecting the community that includes this researcher. To do so, and as illustrated by Figure 4, the first phase of our approach is social network multilayer modeling. Then, when a member is looking for his community, a combined community detection algorithm will be applied to identify the relevant community.

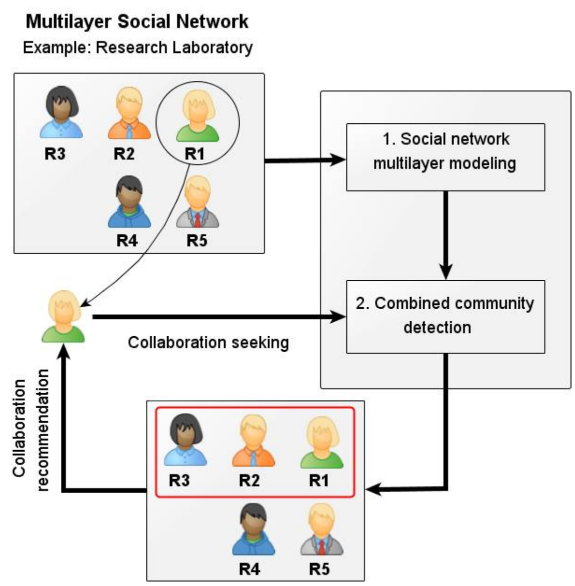

Fig. 4. Hybrid community detection approach for collaboration recommendation

These phases will be detailed in the following.

A. *Social network multilayer modeling*

To compute similarities between participants, we propose firstly to model conceptually the multilayer social network. In fact, after collecting and extracting information about social network users; these data have to be structured in order to facilitate their analysis. Thus, we propose to use ontology as a tool for structuring the different knowledge in a multilayer social network. We choose ontology because it is a standard modeling tool and it guarantees the interoperability [37].

As illustrated by Figure 5, we consider a multilayer social network as a set of participants connected through a variety of relationships. Thus, a multilayer social network is composed of a set of social networks. Each network is characterized by a particular social relationship type and whitin it each participant is described by a profile. Inspired by [38], each profile includes two facets of characteristics. The first facet captures explicit participant characteristics, such as demographic characteristics, skills, interests, profession, diplomas, etc. The second facet captures implicit (or behavioral) characteristics of the participant, such as his/her activities, his/her interactions in the social network, etc.

Within a multilayer social network, it is possible to detect communities. Inspired by community definitions proposed in [9] and [23], we propose to consider a community as a group of actors strongly connected and more weakly connected to the rest of the network and who share similar properties. Furthermore, each community is characterized by an objective. For example, in the case of scientific collaboration recommendation, the aim of the detected community is problem solving.

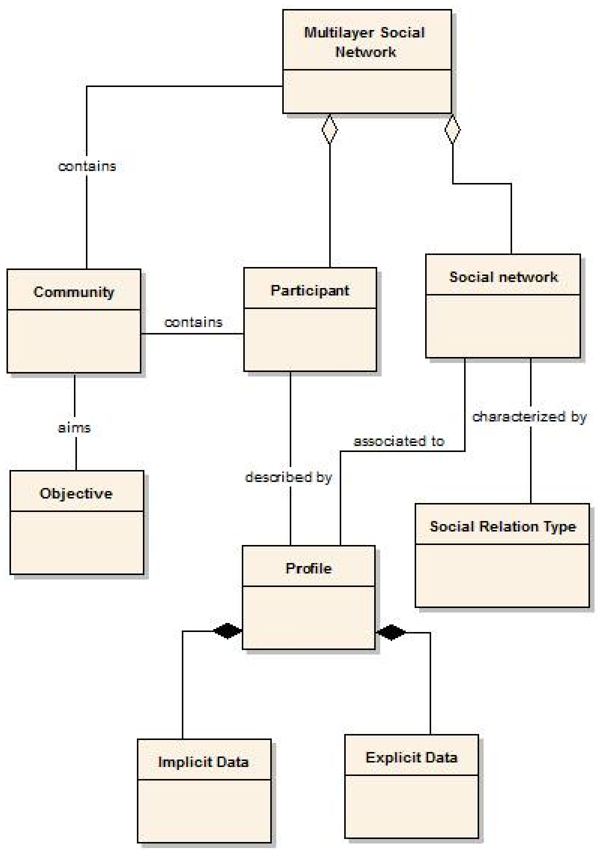

Fig. 5. Multilayer social network's ontological model

The second modeling step is graph construction. Thus, we propose a new graph-based model to represent multilayer social network. Indeed, we propose firstly to reuse a multiplex graph in order to represent different types of relationships. Then we propose to reuse information graph. In fact, due to this type of graph we can represent various members' profiles so as to calculate similarities between "nodes".

To present the proposed graph-model in a formal way, we will use the following notations:

- MIG : Multiplex Information Graph.
- N : The set of Nodes of MIG.
- E_i : The set of edges of MIG which represents edges between nodes of MIG linked by the same type of relationship i.
- G_i : Monoplex Information Graph which represents the layer i of MIG.
- P_i : The set of Nodes of G_i.

The Multiplex information graph, as illustrated by Figure 6, is characterized by a number Nbr of nodes (which represents the number of social network' participants) and a number K of layers (which represents the number of relationships types). Each layer i is represented by a monoplex information graph G_i.

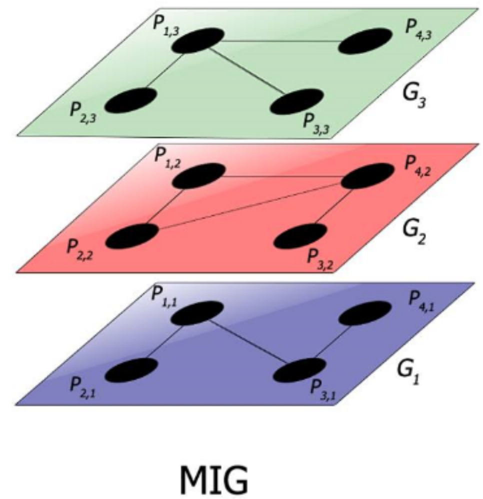

Fig. 6. Example of multiplex information graph (without specifying edges and nodes weights)

$$MIG = \langle N, E_1, E_2, ..., E_K \rangle \ . \ (1)$$
$$N = \langle N_1, N_2, N_3, ..., N_{Nbr} \rangle \ . \ (2)$$

Each node N_i (which represents a person/human being, for example a researcher) is associated at the most K profiles. Each profile is represented conceptually in the ontological model and graphically by the node called $P_{i,j}$ in the monoplex information graph G_j. $P_{i,j}$, $j \in \{1..K\}$ is characterized by K weights $p_{i,j}$, $j \in \{1..K\}$. Each weight represents similarities between profiles. If the node isn't associated at a profile (for example, if a member of the laboratory hasn't a Facebook account) its weight will be zero.

$$N_i = \{(P_{i,j}, p_{i,j}); j \in \{1..K\}\}; \forall i \in \{1..Nbr\} \ . \ (3)$$

Each set of edges $E_i; \forall i \in \{1..K\}$ can be written as follows:

$$E_i = \left\{ (P_{x,i}, P_{y,i}, v_{x,y}^{(i)}); x, y \in \{1..Nbr\} \right\} \ . \ (4)$$

$v_{x,y}^{(i)}$ is the weight of the edge between $P_{x,i}$ and $P_{y,i}$. This weight represents the structural similarity between the two nodes. For examples, for a co-publishing relationship, structural similarity may be the number of published papers; or the number of friends in common in the case of a friendship.

Each monoplex information graph G_i can be written as follows:

$$G_i = \langle P_i, E_i \rangle \ . \ (5)$$

With P_i is:

$$P_i = \{P_{j,i}; j \in \{1..Nbr\}\}; \forall i \in \{1..K\} \ . (6)$$

B. Combined community detection

Once the modeling ended and as a member who is looking for his/her community, the second step is combined community detection. This step begins by calculating the similarities between participants and the participant (represented in the graph by his/her three profiles $P_{1,1}$, $P_{1,2}$ and $P_{1,3}$) who is seeking collaborators. To do so, we propose to use the measurement of [39]. This measure has the advantage of being simple to calculate, more powerful and more expressive than others [40]. Thus, we obtain weighted multiplex information graph as illustrated by Figure 7.

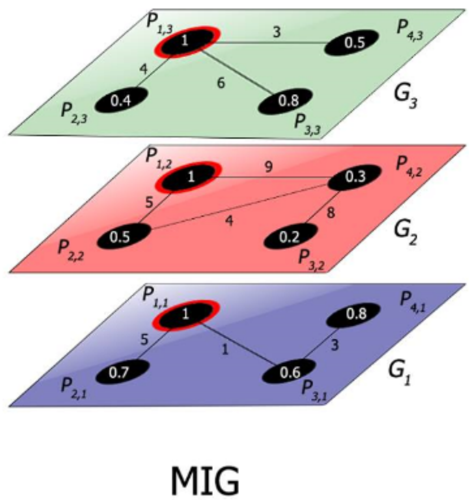

Fig. 7. Example of weighted multiplex information graph for collaboration recommendation

This graph will be the input of a combined community detection algorithm. To do so, we propose to define a new generalized combined metric Q_{MIG} for multiplex graph. Inspired from [41], the new proposed metric Q given by (7) is based on a weighted combination of two components that must be maximized simultaneously:

$$Q_{MIG} = \alpha M_{MIG} + (1-\alpha) I_{MIG} \ . (7)$$

The first component concerns the frequency of social interactions between individuals based on the assumption that people who frequently socialize (have interactions between them) are more likely to collaborate together. It relies on the structural quality, thus we propose to define a new multiplex modularity noted M_{MIG} and given by (8). This modularity is based on the modularity M of Newman for weighted graphs.

$$M_{MIG} = \sum_{i=1}^{K} \beta_i M_{G_i} \ . (8)$$

With M_{G_i} is the modularity of Newman for the graph G_i.

The second component concerns the attribute similarity. Thus, we propose to reuse the notion of inertia I_{MIG} [42]. Inertia is a metric that permits to measure the dispersion of a weighted cloud (a set of nodes where each node has a weight). In order to calculate this inertia, we propose to define for each node N_i a global weight p_i as:

$$p_i = \sum_{j=1}^{K} \gamma_j p_{i,j} \ . (9)$$

α is a dynamic weighting factor where $0 < \alpha < 1$ that can be changed. This factor is related to the community detection context . If we want for example to obtain equitability between modularity proportion and inertia proportion, we can set α to 0.5.

The coefficients β_i and γ_i are related to the community detection context too. And they must be chosen so that:

$$\sum_{i=1}^{K} \beta_i = 1 \ . (10)$$
$$\sum_{i=1}^{K} \gamma_i = 1 \ . (11)$$

At this stage of our research, and to do simply we propose to choose:

$$\beta_i = \gamma_i = \frac{1}{K}; \forall i \in \{1..K\} \ . (12)$$

Finally, in order to maximize this generalized combined quality, which is a NP-hard (non-deterministic polynomial-time hard) optimization, we use a computational optimization technique (i.e. Particle Swarm Optimization) as proposed in the community detection approach of [41]. In fact, a comparative study, done by [43], confirms that particle swarm optimization (PSO) is faster than other used techniques such as simulated annealing [44], the genetic algorithm [45], or the ant colony algorithm [46], etc. It gives best results for NP-hard optimization and it is equally characterized by its simple implementation and its fast convergence [47].

V. Experimentation: Scientific collaboration recommendation case study

To evaluate the proposed hybrid community detection approach, we propose to compare its results with results of basic community detection approaches for multilayer social networks using network aggregation and multilayer network exploration. Thus, we choose to apply Louvain algorithm [16] and the Generalized-modularity optimization approach [8].

To experiment these approaches, we choose to deal with a real network. Thus, we firstly choose the sciences research laboratory RIADI[6] as a multilayer social network that contains 155 members and three layers (DBLP, Facebook, and Linkedin). Then for the same network we consider five layers (DBLP, ResearchGate, Facebook, Twitter and Linkedin). All these data are collected manually. Then, the research laboratory RIADI is conceptually modelled by the ontolgical model and graphically modelled by a multiplex information graph. We consider a researcher (the problem holder), represented by a node, who has a problem (for example a problem in the Java development) and he is seeking relevant collaborators to assist him. For each layer, the weight of each node represents the similarity between this node profile and the profile of the problem holder. For the three approaches, our aim is to recommend for the problem holder the most relevant community.

[6]www.riadi.rnu.tn

Next, we rely on the Barabási-Albert model [48] which is an algorithm for generating random scale-free networks using a preferential attachment mechanism in order to increase the number of nodes. Thus, in each time as shown in Table 1 we increase the size of the graph then the number of layer and we compare the three approaches.

TABLE I
DATASETS DESCRIPTION

Data Notation	Nodes number	Layers number	Edges number in each layer L_i
D1	155	3	L1(3582), L2(1253), L3(261)
D2	155	5	L1(3582), L2(2645), L3(1253), L4(193), L5(261)
D3	2000	3	L1(53439), L2(53092), L3(54537)
D4	2000	5	L1(54441), L2(54045), L3(53564), L4(53664), L5(54417)
D5	3000	3	L1(93297), L2(95475), L3(94603)
D6	3000	5	L1(93944), L2(95746), L3(94023), L4(93377), L5(93893)

The comparison between these approaches, given in Table 2, is based on time execution (in minute), the modularity, the inertia and the redundancy [31] of the detected community.

TABLE II
COMPARISON BETWEEN QUALITIES OF COMMUNITY DETECTION

Community Detection approaches for multilayer network	Multilayer Social Network	Times T(min)	Detected community modularity M	Detected community Inertia I	Detected community Redundancy R
Network aggregation: Louvain algorithm [16]	D1	0.001	0.103	0.710	0.690
	D2	0.001	0.100	0.703	0.452
	D3	11.690	0.040	0.020	0.190
	D4	11.830	0.015	0.100	0.150
	D5	14.600	0.008	0.017	0.089
	D6	13.930	0.009	0.047	0.022
Multilayer network exploration: Generalized-modularity optimization [8]	D1	0.002	0.191	0.680	0.870
	D2	0.004	0.153	0.712	0.891
	D3	31.420	0.039	0.261	0.468
	D4	52.210	0.024	0.092	0.533
	D5	91.250	0.014	0.214	0.658
	D6	195.110	0.009	0.174	0.499
Proposed hybrid community detection	D1	0.008	0.210	0.963	0.890
	D2	0.009	0.192	0.923	0.744
	D3	1.650	0.008	0.720	0.666
	D4	2.550	0.020	0.590	0.498
	D5	3.910	0.004	0.703	0.575
	D6	4.670	0.007	0.679	0.401

To better visualize these measures, we exploited these values to plot the following four curves.

Concerning the convergence of each approach, Figure 8 illustrates comparison between execution time of community detection. We remark that for small networks (D1 and D2), the Louvain approach is the fastest. Yet, for large-scale networks (D3, D4, D5 and D6), the proposed approach is slightly faster than the other and needs less time to converge. In addition, the number of layers does not have much influence on the convergence time of our approach as in the case of the Generalized-modularity optimization approach.

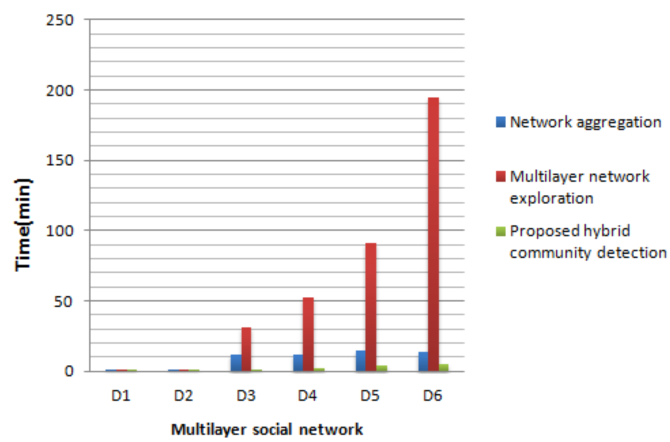

Fig. 8. Comparison between execution times of community detection

Considering only the inertia proportion, illustrated by Figure 9, results show that the proposed approach provides better inertia values for almost all networks.

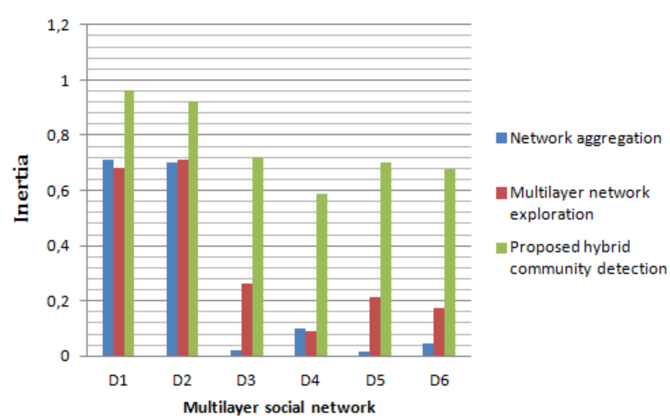

Fig. 9. Comparison between Inertia of the detected community

However, for modularity, the Generalized-modularity optimization approach gives better results for large-scale networks but the difference with the proposed approach is not very high as shown in Figure 10.

For the redundancy, which is a new measure defined in [31] to characterize communities in multilayer network by capturing the phenomenon for which a set of nodes that constitute a community in a dimension tend to constitute a community also in other dimensions, we remark, as shown

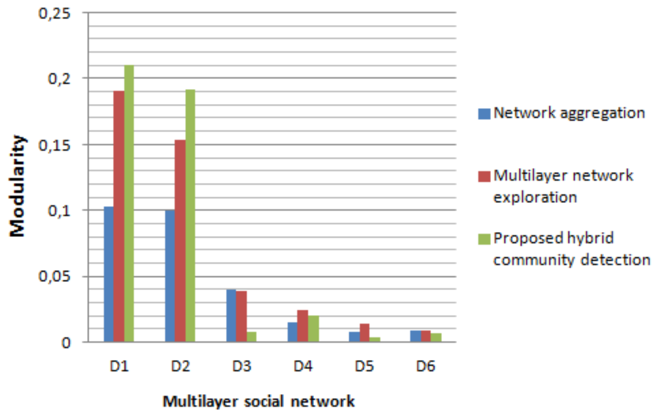

Fig. 10. Comparison between Modularity of the detected community

in Figure 11, that the Generalized-modularity optimization approach and the proposed one gives better results then the Louvain approach.

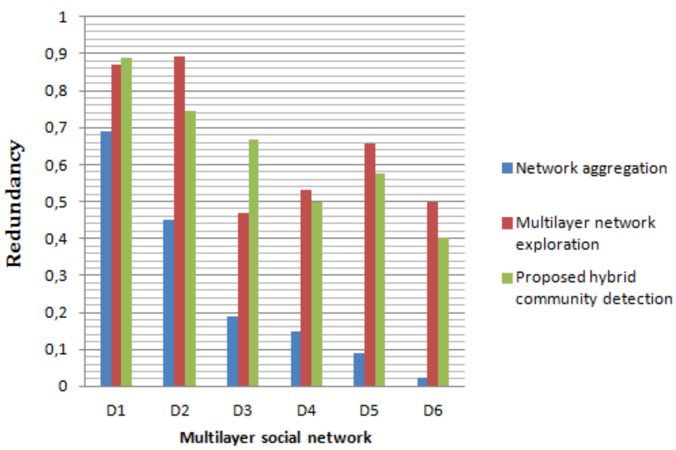

Fig. 11. Comparison between Redundancy of the detected community

These results may suggest that the Louvain approach based on layer aggregation is not the adequate for multiplex community detection. Indeed, layer aggregation approach leads to a great loss of information about the heterogeneous nature of links in multilayer social network. For the Generalized-modularity optimization approach, it gives important results but it does not deal with similarities between users.

VI. Conclusion

In this paper we proposed a new hybrid community detection approach in multilayer social network. As a use case, we proposed to apply this approach to form communities in a research laboratory network in order to support scientific collaboration recommendation. The proposed approach is based on a multiplex information graph which considers simultaneously the network structure (different social connections) and various members' profiles so as to calculate similarities between "nodes". Then, to detect the relevant community, we propose to define a new generalized metric which combines modularity based on network structure (edges weights) and inertia based on the homophily of participants (nodes weights). Finally, we test the proposed approach with other classical approaches. The experimentation showed that our approach gives better results for community detection in multilayer social networks.

In future works, we aim to test the performance and the scalability of the proposed community detection approach in other contexts using large-scale real social networks. In addition, we propose to add a temporal aspect to study community resilience using its productivity and its maturity during the evolution of the social network.

References

[1] M. S. Ackerman, J. Dachtera, V. Pipek, and V. Wulf, "Sharing knowledge and expertise: The cscw view of knowledge management," *Computer Supported Cooperative Work (CSCW)*, vol. 22, no. 4-6, pp. 531–573, 2013.

[2] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, "Multilayer networks," *Journal of Complex Networks*, vol. 2, no. 3, pp. 203–271, 2014.

[3] M. Berlingerio, M. Coscia, F. Giannotti, A. Monreale, and D. Pedreschi, "Multidimensional networks: foundations of structural analysis," *World Wide Web*, vol. 16, no. 5-6, pp. 567–593, 2013.

[4] R. Kanawati, "Multiplex network mining: A brief survey."

[5] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, "The structure and dynamics of multilayer networks," *Physics Reports*, vol. 544, no. 1, pp. 1–122, 2014.

[6] J. Kim and J.-G. Lee, "Community detection in multi-layer graphs: A survey," *ACM SIGMOD Record*, vol. 44, no. 3, pp. 37–48, 2015.

[7] M. Hmimida and R. Kanawati, "Community detection in multiplex networks: A seed-centric approach." *NHM*, vol. 10, no. 1, pp. 71–85, 2015.

[8] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, "Community structure in time-dependent, multiscale, and multiplex networks," *science*, vol. 328, no. 5980, pp. 876–878, 2010.

[9] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical review E*, vol. 69, no. 2, p. 026113, 2004.

[10] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek, "Evolution of the social network of scientific collaborations," *Physica A: Statistical mechanics and its applications*, vol. 311, no. 3, pp. 590–614, 2002.

[11] M. Plantié and M. Crampes, "Survey on social community detection," in *Social Media Retrieval*. Springer, 2013, pp. 65–85.

[12] M. Newman, *Networks: an introduction*. OUP Oxford, 2010.

[13] I. Psorakis, S. J. Roberts, I. Rezek, and B. C. Sheldon, "Inferring social network structure in ecological systems from spatio-temporal data streams," *Journal of the Royal Society Interface*, vol. 9, no. 76, pp. 3055–3066, 2012.

[14] D. A. Ford, J. H. Kaufman, and Y. Mesika, "Modeling in space and time," in *Infectious Disease Informatics and Biosurveillance*. Springer, 2011, pp. 191–206.

[15] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[16] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, p. P10008, 2008.

[17] L. Wan, J. Liao, and X. Zhu, "Cdpm: Finding and evaluating community structure in social networks," in *Advanced Data Mining and Applications*. Springer, 2008, pp. 620–627.

[18] U. Brandes, D. Delling, M. Gaertler, R. Görke, M. Hoefer, Z. Nikoloski, and D. Wagner, "On modularity clustering," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 20, no. 2, pp. 172–188, 2008.

[19] F. Moser, R. Ge, and M. Ester, "Joint cluster analysis of attribute and relationship data withouta-priori specification of the number of clusters," in *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2007, pp. 510–519.

[20] Y. Zhou, H. Cheng, and J. X. Yu, "Graph clustering based on structural/attribute similarities," *Proceedings of the VLDB Endowment*, vol. 2, no. 1, pp. 718–729, 2009.

[21] Y. Sun and J. Han, "Mining heterogeneous information networks: principles and methodologies," *Synthesis Lectures on Data Mining and Knowledge Discovery*, vol. 3, no. 2, pp. 1–159, 2012.

[22] D. Combe, *Détection de communautés dans les réseaux dinformation utilisant liens et attributs*. David Combe, 2013.

[23] S. Fortunato, "Community detection in graphs," *Physics reports*, vol. 486, no. 3, pp. 75–174, 2010.

[24] M. E. Newman, "Coauthorship networks and patterns of scientific collaboration," *Proceedings of the national academy of sciences*, vol. 101, no. suppl 1, pp. 5200–5205, 2004.

[25] P. Pons and M. Latapy, "Computing communities in large networks using random walks." *J. Graph Algorithms Appl.*, vol. 10, no. 2, pp. 191–218, 2006.

[26] S. Gomez, A. Diaz-Guilera, J. Gomez-Gardeñes, C. J. Perez-Vicente, Y. Moreno, and A. Arenas, "Diffusion dynamics on multiplex networks," *Physical review letters*, vol. 110, no. 2, p. 028701, 2013.

[27] F. Battiston, V. Nicosia, and V. Latora, "Metrics for the analysis of multiplex networks," *arXiv preprint arXiv:1308.3182*, 2013.

[28] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, "Mathematical formulation of multilayer networks," *Physical Review X*, vol. 3, no. 4, p. 041022, 2013.

[29] D. Cai, Z. Shao, X. He, X. Yan, and J. Han, "Mining hidden community in heterogeneous social networks," in *Proceedings of the 3rd international workshop on Link discovery*. ACM, 2005, pp. 58–65.

[30] L. Tang and H. Liu, "Community detection and mining in social media," *Synthesis Lectures on Data Mining and Knowledge Discovery*, vol. 2, no. 1, pp. 1–137, 2010.

[31] M. Berlingerio, M. Coscia, and F. Giannotti, "Finding and characterizing communities in multidimensional networks," in *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*. IEEE, 2011, pp. 490–494.

[32] D. Suthers, J. Fusco, P. Schank, K.-H. Chu, and M. Schlager, "Discovery of community structures in a heterogeneous professional online network," in *System Sciences (HICSS), 2013 46th Hawaii International Conference on*. IEEE, 2013, pp. 3262–3271.

[33] V. Carchiolo, A. Longheu, M. Malgeri, and G. Mangioni, "Communities unfolding in multislice networks," in *Complex Networks*. Springer, 2011, pp. 187–195.

[34] M. Rosvall, D. Axelsson, and C. T. Bergstrom, "The map equation," *The European Physical Journal Special Topics*, vol. 178, no. 1, pp. 13–23, 2009.

[35] M. De Domenico, A. Lancichinetti, A. Arenas, and M. Rosvall, "Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems," *Physical Review X*, vol. 5, no. 1, p. 011027, 2015.

[36] Z. Kuncheva and G. Montana, "Community detection in multiplex networks using locally adaptive random walks," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. ACM, 2015, pp. 1308–1315.

[37] J. F. Sowa, "Knowledge representation: logical, philosophical, and computational foundations," *Pacific Grove, CA, Brooks/Cole*, vol. 594, 2000.

[38] A. Abdel-Hafez and Y. Xu, "A survey of user modelling in social media websites," *Computer and Information Science*, vol. 6, no. 4, p. 59, 2013.

[39] Z. Wu and M. Palmer, "Verbs semantics and lexical selection," in *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 1994, pp. 133–138.

[40] D. Lin, "An information-theoretic definition of similarity." in *ICML*, vol. 98, 1998, pp. 296–304.

[41] N. Ben Yahia, N. Bellamine Ben Saoud, and H. Ben Ghezala, "Community-based collaboration recommendation to support mixed decision-making support," *Journal of Decision Systems*, vol. 23, no. 3, pp. 350–371, 2014.

[42] L. Lebart, A. Morineau, and J.-P. Fénelon, "Traitement des données statistiques(méthodes et programmes)," 1979.

[43] A. El Dor, "Perfectionnement des algorithmes d'optimisation par essaim particulaire: applications en segmentation d'images et en électronique," Ph.D. dissertation, Université Paris-Est, 2012.

[44] R. Guimera, M. Sales-Pardo, and L. A. N. Amaral, "Modularity from fluctuations in random graphs and complex networks," *Physical Review E*, vol. 70, no. 2, p. 025101, 2004.

[45] C. Pizzuti, "Community detection in social networks with genetic algorithms," in *Proceedings of the 10th annual conference on Genetic and evolutionary computation*. ACM, 2008, pp. 1137–1138.

[46] D. He, J. Liu, D. Liu, D. Jin, and Z. Jia, "Ant colony optimization for community detection in large-scale complex networks," in *Natural Computation (ICNC), 2011 Seventh International Conference on*, vol. 2. IEEE, 2011, pp. 1151–1155.

[47] K. E. Parsopoulos, *Particle Swarm Optimization and Intelligence: Advances and Applications: Advances and Applications*. IGI Global, 2010.

[48] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.