

Incremental Community Mining in Location-based Social Network

Loubna Boujlaleb, Ali Idarrou, Driss Mammass

IRF-SIC Laboratory, Ibn Zohr University, Agadir, Morocco

Email: loubna84@gmail.com

Abstract— A social network can be defined as a set of social entities connected by a set of social relations. These relations often change and differ in time. Thus, the fundamental structure of these networks is dynamic and increasingly developing. Investigating how the structure of these networks evolves over the observation time affords visions into their evolution structure, elements that initiate the changes, and finally foresee the future structure of these networks. One of the most relevant properties of networks is their community structure – set of vertices highly connected between each other and loosely connected with the rest of the network. Subsequently networks are dynamic, their underlying community structure changes over time as well, i.e they have social entities that appear and disappear which make their communities shrinking and growing over time. The goal of this paper is to study community detection in dynamic social network in the context of location-based social network. In this respect, we extend the static Louvain method to incrementally detect communities in a dynamic scenario following the direct method and considering both overlapping and non-overlapping setting. Finally, extensive experiments on real datasets and comparison with two previous methods demonstrate the effectiveness and potential of our suggested method.

Keywords— *Data Mining; Dynamic Social Network, Community Detection, Location-based social network; Dynamic Network Analysis.*

I. INTRODUCTION

A social network is an information network frequently represented by a graph where vertices represent social entities and edges represent social relationships between these entities. Actually, many types of data can be naturally observed as networks; for example, citation networks where vertices are documents and citation links between them. The web is seen as html documents and hyperlinks between them. Email communication [13] and all kind of human social data can be characterized as a network.

Moreover, the spread and democratization of smartphones and mobile devices (Iphone, Ipad, SamSung

Galaxy ...) has generated a new concept in mobile social media called mobile social network or location-based social network. This type of network allows individuals to “check-in” at a physical place and share it with their online friends [20]. Location-based social networks provide rich geographical and temporal information (spatiotemporal data) from people’s location over time (check-in actions). Social network analysis has produced a great attention of various researchers from different specialties. The relationships between vertices are dynamic and varying and evolving over time;

One of the methods used to obtain information about the network is the identification of community structure. This research field represents one of the most major problems in network analysis and in the field of complex networks [22], [23]. Generally, a community is observed as a group of actors in which intra-group relationships are much thicker than those of inter-group ones [24]. Detecting communities is of great importance in epidemiology [1, 2, 3], biology [8, 9, 10, 11, 12], sociology [13, 14] and also recently in computer science [4, 5, 6]. Consequently, studying the evolution of communities would be of great interest to scientists.

Modeling the dynamic social network may be done through two techniques: The first technique is to consider the dynamic social network as one static social network by combining all the interactions over time into one time interval. Nevertheless by using this technique, one may eliminate significant information that is taking place inside the network. The greatest modeling technique for such dynamic social network is to split the network evolution into sequence of consecutive social network where each time interval incorporates the connections that are happening in this precise time interval [21]. This modeling method permits the study of the dynamic network structure over time, the identification of how the network evolves, and finally the forecast of the future network tendency. Besides, this prediction has numerous vital applications, like targeted marketing [14] and recommendation systems [15].

Two methods have been assumed to track the evolution of communities in a dynamic picture: The independent

community detection and incremental community detection. Basically, there exist two approaches to recognize communities incrementally, cost function method, and direct method. The first method tries to detect communities at each time frame without considering the temporal information and association between communities from successive time frames. This method is appropriate for social networks with variable community structures. The incremental community detection algorithms take into account the temporal information directly throughout the mining, where the community detection at a particular time is reliant on the communities mined in earlier time interval. This technique is fitting for networks with community structures that are resistant over time.

In our earlier contribution [16], we have considered the independent community detection approach. In this paper, we suggest an incremental community detection approach that identifies communities by taking into account communities found earlier. This conditional community detection method is more proper for observing more strong communities compared to our former independent approach. The central contribution of our paper is to extend the static Louvain method [25] to calculate evolving communities; where community detection at each time interval originates by the communities identified at the former time interval. Besides, our method is adapted to overlapping and non-overlapping communities. To characterize the variations that are probable to occur for an evolving community, we apply our community evolution detection technique [19] to identify community operations (i.e. continuing, shrinking, growing, splitting, merging, birth and death). This paper is structured as the following: in the next section, we overview some related and previous research in the area of dynamic community detection. Section 3 details the suggested technique. The experimental study and results are given in Section 4. Section 5 demonstrates a summary and future work.

II. RELATED WORKS

There has been an important amount of study about community detection; they are largely classified into two main categories: static methods and dynamic methods. In both categories, there are methods that take into account overlapping communities while others consider non-overlapping methods. We have classified the dynamic methods as the following 1) independent community detection algorithms; 2) incremental community detection algorithms.

2.1 Independent community detection algorithms

The first approach tries to detect communities at each time frame without considering the temporal information and association between communities from successive

time intervals. Moreover, this type of dynamic community detection algorithms are based on applying directly the static algorithms as follows: Split the network into different time frames, apply a classical static algorithm into each time frame and match the communities found in one time frame with the communities of the precedent time frame. The comparison of communities from successive time intervals is founded on using different rules based on the size of their intersection; those rules may be used with the combination of community detection algorithm [26], applying clustering on a graph formed by all detected communities at different time frames [27]; applying heuristic algorithms to match communities based on their intersection [28] or extracting core vertices that are more descriptive of their communities [29]. There are two fundamental matters with the independent community detection method. First, the static algorithms used on each time interval are frequently non-deterministic and henceforth generate dissimilar communities even if the input graph does not alter. This variability generates noise that makes the monitoring very challenging. Additionally, due to the fact that, the communities are mined independently at each time interval without taking into account former associations, this technique is only appropriate for the social networks with significantly dynamic community structures.

2.2 Incremental community detection algorithms

The incremental community detection algorithms take into account the temporal information directly during the discovery, where the community detection at an exact time is reliant on the communities identified in preceding time interval. This method is applicable for networks having stable community structures with time. In addition, this method includes both current and temporal information. There exist two techniques to extract communities incrementally which are cost function methods and direct methods.

2.2.1 Cost function method

The cost function method is based on minimizing a cost function, which is first proposed by [16] to trade off between the history quality and the current snapshot quality. The cost function is usually composed of two sub-costs of a snapshot cost (SC) and a temporal cost (TC). Let G_i and C_i be the graph and set of detected communities at time interval i respectively. The general formulation of the cost function is as follows:

$$\text{Cost} = \alpha \text{SC}(G_i, C_i) + (1 - \alpha) \text{TC}(G_{i-1}, C_i) \quad (1)$$

[16] proposed an evolutionary clustering method. They have therefore not been concerned with the evolution of communities over the long term, nor with questions of fusion or division. On the other hand, they sought to ensure that the communities found at time $t + 1$ are coherent with respect to time t . To do this, they have

developed a quality function in two components: the first one is static, and therefore concerns the studied network at instant t , while the other serves to ensure stability, and thus evaluates the distance between the clusters in the previous step and the clusters in the current step. The quality function can thus be expressed in the following way:

$$Q = Q_{\text{instant}} + \alpha Q_{\text{stabilit}} \quad (2)$$

[17] reuse the idea of the core vertices, previously proposed by [18], but use a trick to reduce instability between two detections. They use the Louvain algorithm for the detections on each snapshot, and they initialize this algorithm with the core vertices found in the previous step, which makes it possible to limit the instability. However, it remains important. In this algorithm, the core vertices are defined as those that do not change communities if the same algorithm is repeatedly run on the same slightly modified network

2.2.2 Direct method

Direct methods recognize communities at the recent time interval incrementally by taking into consideration the communities found at earlier interval and upgrade the community structure once different data comes. As an example, Sarkar and Moore [15] proposed the Latent space model with time-based variation to determine communities that are reliable with the network at the recent time and with the communities found at a preceding time. Also, Mucha et al. [7] simplify the Laplazier dynamics technique so as to expand modularity maximization to examine community structure covering various time intervals in evolving social network. As a summary about the methods explained in the above, we observe that cost function methods tries to improve a new quality metric which combines divergence from past, however direct method attempts to update the community structure as new data comes.

From the above review, our contribution is based on extending the Louvain static community detection [25], to incrementally identify communities in a dynamic situation following the direct technique. The reason that we recommend this Louvain method is that it does not demand parameters. Furthermore, different from most of static community detection algorithms that indirectly accept global information is always available, it finds communities with only local information. This locality makes Louvain method primarily recommended in the scenario of large real world networks, where the entire graph is often inaccessible. Lastly, the Louvain community detection method may be just improved to identify overlapping or non-overlapping communities. Consequently, we use the Louvain community detection since none of the others in the literature review fulfill all these desires at the same time. Our experiments confirm

that the proposed algorithm outperforms previous algorithms considerably.

III. PROPOSED APPROACH

Let $G=(V, E, L)$ be a location-based social network, where V with $|V| = N$ is the set of vertices (users), E with $|E| = M$ is the set of edges in the network and $L = \{l(v_1), l(v_2), \dots, l(v_N)\}$ is the set of check-in locations of actors. A dynamic location-based social network is a network in which V, E or L varies over time and illustrated as a series of graphs $\{G_1, G_2, \dots, G_n\}$, where $G_i=(V_i, E_i, L_i)$ describes the graph at time interval i , which incorporates V_i vertices and E_i links. The n_i communities mined at the i^{th} time interval are then represented by $C_i = \{C_i^1, C_i^2, \dots, C_i^{n_i}\}$, where community $C_i^k \in C_i$ is also a graph showed by (V_i^k, E_i^k) .

3.1 Definitions

In the next, we suggest a mobile modularity metric and average mobile modularity.

Definition 1 (mobile modularity MQ): Supposed the n_i communities identified at the i^{th} time interval are then represented by $C_i = \{ \dots, \}$, the mobile modularity of this partition might be given as

$$MQ_i = \alpha Q(G_i, C_i) + (1 - \alpha)Q(G_{i-1}, C_i) \quad (3)$$

Where $Q(G_{i-1}, C_i)$ is calculating the gain of modularity Q for communities at time interval i using graph from the previous time interval. Though $Q(G_i, C_i)$ is the gain of the static Q modularity for communities found in time interval i .

Definition 2 (Average mobile Modularity Q): Supposed the series of graphs $\{G_1, G_2, \dots, G_n\}$ of a dynamic social network, the average mobile modularity computed on all the time intervals is given as the following

$$\overline{VQ} = \frac{1}{n} \sum_{i=1}^n VQ_i \quad (4)$$

The average Mobile modularity \overline{VQ} , fluctuates from -1 to 1 likewise to static Modularity Q . The comparison of the gain of \overline{VQ} , of several dynamic community detection algorithms permits the estimation of these algorithms on a precise condition. The algorithm which generates higher \overline{VQ} , has better results rather than others on that condition.

3.2 Direct Incremental Louvain method

There exist two different settings of communities: overlapping and non-overlapping communities. In a non-overlapping setting, an individual can belong to only one community, while in an overlapping setting; an individual can belong to various communities. In this section, we explain our suggested incremental community detection method which extends our preceding independent community detection method [30] to incrementally detect communities in a dynamic scenario following the direct approach for non-overlapping and overlapping setting.

3.2.1 Non-overlapping setting

Communities in the current time are obtained based on the communities from the past time interval. The central idea is to substitute the static modularity into the mobile modularity in Louvain algorithm which automatically will modify the initialization of the algorithm in a way that the communities are extracted at each time interval i over the network at time interval i and the network of the previous time interval $i-1$. Similar to the independent community detection method, our proposed incremental method is local and can detect overlapping and non-overlapping communities.

Formally, the Incremental community detection method identifies communities for a dynamic social network with the following process. In the first step, we divide the dynamic social network into 12 time intervals. Since we are working in the context of location-based social network, in each time interval, we construct a weighted social network by changing the weight of the edges of the social network with respect to the temporary information from users' check-ins. Specifically, for each pair of neighbor vertices in a set of vertices V where $v_i, v_j \in V$, $\forall v_i \neq v_j$, the weight of the edge $e(v_i, v_j)$ is changed using a similarity metric between the two vertices by using the following equation:

$$w_{ij} = 1 + \text{sim}(i, j) \quad (5)$$

where $\text{sim}(i, j)$ is the spatio-temporal similarity between i and j .

Our method tries to find a group of time intervals in which a single group of communities is related for all its time intervals. In other words, the final community structure should not only be a good partition for that time interval, but also a realistic partition for the preceding time intervals. To do that, we have proceeded as the following: The first step consists on discovering communities at each time interval using Louvain algorithm using static modularity. Then, we try to combine time intervals with similar community structure to have one aggregate graph. After that, we apply the Louvain algorithm on the aggregate graph. Note that, our incremental community detection method not necessarily considers communities between two successive time intervals but also between discontinuous time intervals. The pseudo code related to the previous process is given in algorithm 1.

3.2.2 Overlapping setting

In the case of overlapping communities, we apply the same process given in algorithm 1 by changing the static modularity to static modularity for overlapping communities as the following:

$$Q_{vv'} = \frac{1}{2M} \sum_{v \neq v'} [(A_{vv'} - \frac{k(v).k(v')}{2M}) \frac{A(v) \cap A(v')}{A(v) \cup A(v')}] \quad (6)$$

Where $A(v)$ designates the membership set of node v (i.e. the communities that node v is a member of them). For vertices v and v' that always belong to the same community (i.e. $A(v) \cap A(v') = A(v) \cup A(v')$), their involvement to the modularity $Q_{vv'}$ is $A_{vv'} - \frac{k(v).k(v')}{2M}$, while the involvement is 0, for nodes v and v' that never belong to the same community (i.e. $A(v) \cap A(v') = \emptyset$). Otherwise, their involvement is within the interval $[0, A_{vv'} - \frac{k(v).k(v')}{2M}]$.

Algorithm 1 : Incremental Louvain algorithm

Input: A weighted aggregate graph G from set of time intervals

Output: A division D

```

1: repeat
2:   Put each vertex of  $G$  in a single community;
3:   Compute the gain a variable modularity  $VQ$  of this division
4:   repeat
5:     for all vertices  $u$  of  $G$  do
6:        $C \leftarrow$  neighbor community maximizing variable modularity value;
7:       if the movement of  $v$  in  $C$  produces a firmly positive value then
8:         change  $v$  from its community to  $C$ ;
9:     until there is no vertex to be transferred;
10:  if the quality measure attained is superior than its initial gain then
11:     $\text{fin} \leftarrow$  false;
12:    display the found decomposition;
13:    combine  $G$  into the graph between communities
14:  if not
15:     $\text{end} \leftarrow$  true;
16: until end;

```

IV. EXPERIMENT RESULTS

4.1 Experimental Datasets

We conduct experiments to evaluate our algorithm on a real social network dataset called Brightkite. It is a location-based social network created in 2007: users are capable to check-in at locations through their mobile devices; Brightkite users can create mutual friendship relations and they can push their check-ins to their Twitter and Facebook accounts. We study a dataset gathered in September 2009 which contains the whole Brightkite user base at that time, with information about 54,190 users. In the dataset, a check-in record is a tuple $\langle \text{userid}, \text{check-in}$

time, latitude, longitude, location id>. Here latitude and longitude signifies the latitude and longitude of the location where the user visited, and check-in time represents the time stamp of the check-in activity.

We are studying active users whose number of check-ins is greater than 50. The goal of selecting this number is to capture essential features of users' behaviors through check-in activities. We set the time intervals to be one month each for a period of one year. Furthermore, we compare our suggested direct incremental community detection method with the cost function incremental community detection technique and independent variation. The comparison between these algorithms is executed from two viewpoints: first, relatively founded on a direct objective for dynamic communities, and then indirectly based on how much they increase the community operation detection framework and forecast precision. In our experiments, we have utilized our community evolution detection method [19] to extract critical community operations. We have experienced for diverse values of α for both objective functions: Static and Mobile modularity Q and we selected to use $\alpha=0.5$ in our experiment.

4.2 Evaluation Measures

4.2.1 Relative evaluation

Figure 1 exhibits a whole comparison based on the quality, size and number of communities over the time respectively. We can evidently observe that incremental approach is constantly discovering communities with important quality regarding both recent and temporal information. Average size of communities is remarkably less for the independent method as it is unsuccessful to identify constant communities that stay for a long time and distinguish several unimportant communities which is predicted as it only use the actual time interval to mine communities. The value of the average community size is bigger for incremental algorithm; hence it outperforms the independent one. The average mobile modularity Q for the independent, incremental /overlapping and incremental/non-overlapping are 0.36, 0.40, and 0.38 respectively. The experimental outcomes demonstrate that the three of the community mining algorithms have slightly equal average mobile modularity Q, however the number of identified communities and their size are dissimilar. Furthermore, the selection of community detection algorithms expresses temporal and evolutionary compartment of communities and the structure of networks from different insight. Henceforward, the different community detection techniques must be chosen based on the actual situation and application.

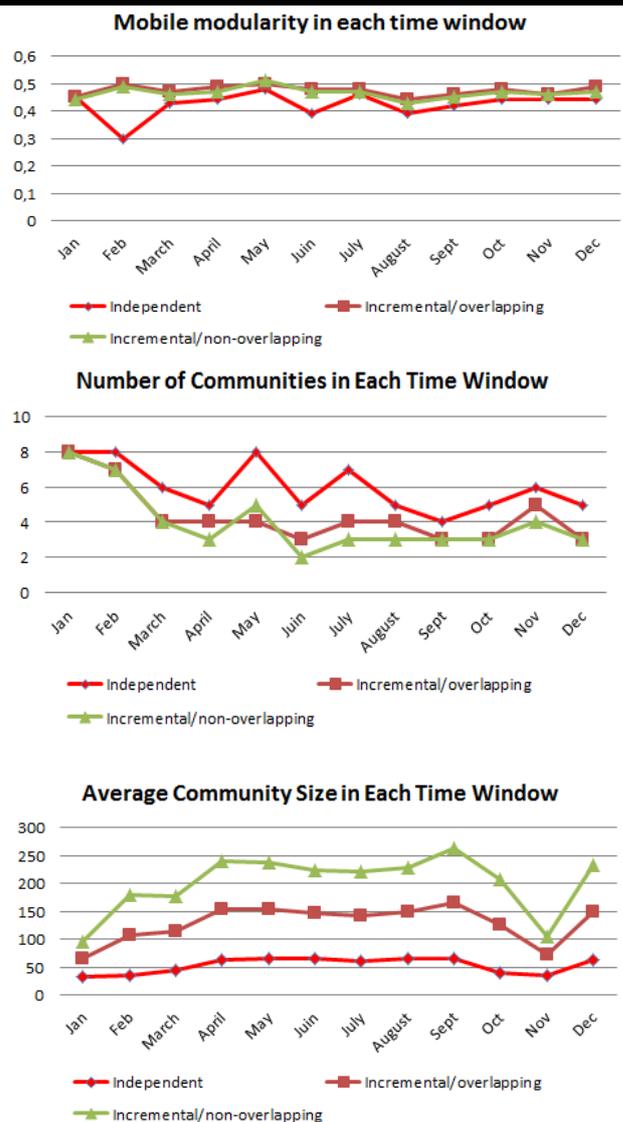


Fig. 1: Relative Evaluation for Each Time interval

4.2.2 Indirect evaluation

The effect of diverse versions of dynamic community detection techniques on the recognized operations is demonstrated in Table 1, where the complete number of operations for each detected class during the 12 time intervals is given.

Table 1 : Indirect evaluation

Algorithm	Birth	Death	Continuing	Splitting	Merging
Incremental/ Direct/overlapping	12	12	29	4	10
Incremental/ Direct/non-overlapping	13	13	24	6	12
Independent	32	30	45	10	15

Another description of community analysis and operations detected for community mining algorithms is better illustrated in Figure 2. We are giving a color to each similar community at each time interval and white color is assigned to communities that appear in one time interval. Additional, solid, dashed, and dotted arrows describe extracted continuing, splitting, and merging operations correspondingly. The independent approach is too dynamic, identifying communities that fluctuate a lot between time intervals, and consequently, triggering numerous arbitrary produced operations, e.g. 32 Birth, 30 Death. The Incremental method for overlapping setting accurately describes the constant communities over time intervals by incorporating the temporal information, and simultaneously, identifies other types of operations carefully.

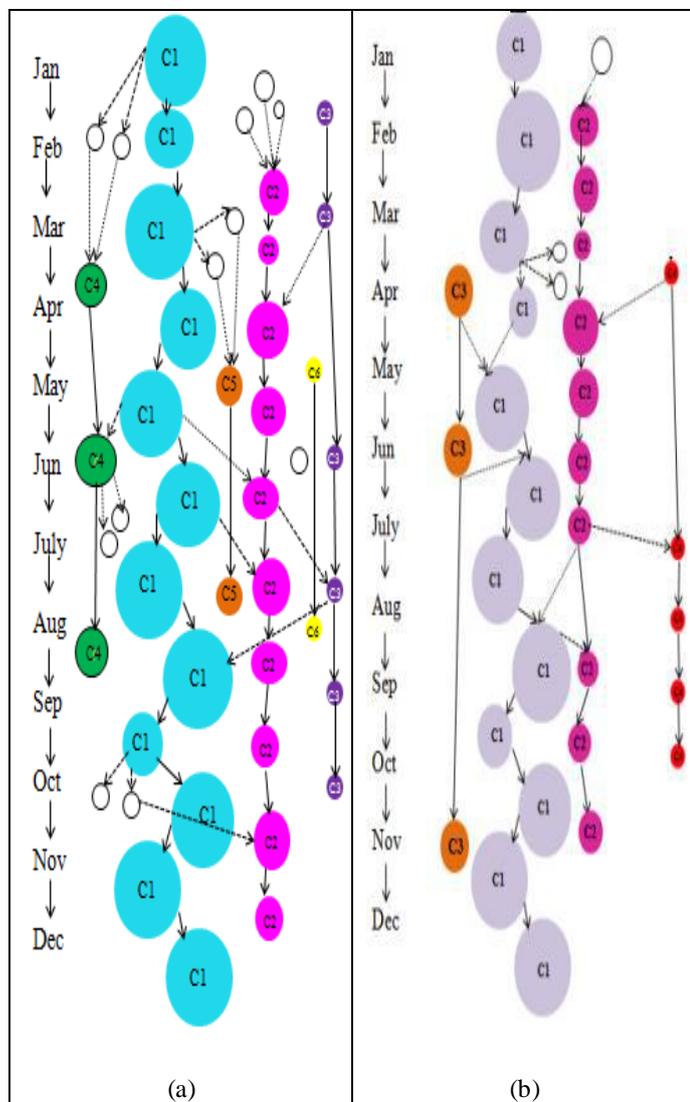


Fig. 2: (a) operation detected for independent algorithm; (b) Operation detected for incremental direct overlapping algorithm

V. CONCLUSION

Community detection and community evolution over time

are seen as one of the greatest significant problems in social network analysis. The early techniques tried to fix this issue by discovering communities at each time interval individually without taking into account communities at other time intervals. In this paper, we have briefly illustrated different dynamic community detection techniques. We then suggest an Incremental direct Louvain community detection method for both overlapping and non-overlapping communities. The suggested method is then compared with its corresponding independent version. Our experiment is using a real-life datasets in order to evaluate the consistency and viability of our method. The results showed that our method attains flexibility, able to bring in more meaningful communities, provides more vision into the dynamic of communities and enhances the time and computational complexity. There are various future directions that need to further investigation. First, it would be remarkable to study how to expand our current method by forecasting the future changes of the communities based on the current and earlier operations. Second, it would also be interesting to test our suggested methods to identify communities in other kinds of networks (e.g., collaboration networks, communication networks, and so on).

REFERENCES

- [1] Kretzschmar, M. & Morris, M. (1996). Measures of concurrency in networks and the spread of infectious disease. *Math. Biosci.*, 133:165–195.
- [2] Meyers, L. A., Newman, M. & Pourbohloul, B. (2006). Predicting epidemics on directed contact networks. *Journal of Theoretical Biology*, 240:400–418.
- [3] Read, J. M. & Keeling, M. J. (2003). Disease evolution on networks: the role of contact structure. *Proc. R. Soc. Lond. B*, 270:699–708.
- [4] Aizen, J., Huttenlocher, D., Kleinberg J., & Novak, A. (2004). Traffic-based feedback on the web. *Proc. Nat. Acad. Sci.*, 101(Suppl.1):5254–5260.
- [5] Flake, G., Lawrence, S., & Giles, C. L. (2002). Coetzee. Self-organization and identification of web communities. *IEEE Computer*, 35(3).
- [6] Getoor, L. & Diehl, C. P. (2005). Link mining: a survey. *SIGKDD Explor. NewsL.*, 7(2):3–12.
- [7] Mucha, J.P., Richardson, T., Macon, K., Mason, A., & Onnela, J.P. (2010). Community structure in Time-Dependent, multiscale, and multiplex networks. *Science*, 328(5980):876–878.
- [8] Croft, D. P., James, R., Thomas, P., Hathaway, C., Mawdsley, D., Laland, K., & Krause, J. Social structure and co-operative interactions in a wild population of guppies (*poecilia reticulata*). *Behavioural Ecology and Sociobiology*, In Press.

- [9] Cross, P. C., Lloyd-Smith, J. O., & Getz, W. M. (2005). Disentangling association patterns in fission-fusion societies using African buffalo as an example. *Animal Behaviour*, 69:499–506.
- [10] Fischhoff, I. R., Sundaresan, S. R., Cordingley, J., Larkin, H. M., Sellier, M.-J., & Rubenstein, D. I. (2006). Social relationships and reproductive state influence leadership roles in movements of plains zebra (*Equus burchellii*). *Animal Behaviour*.
- [11] Rubenstein, D. I., Sundaresan, S., Fischhoff, I., & Saltz, D. (2007). Social networks in wild asses: Comparing patterns and processes among populations. In A. Stubbe, P. Kaczensky, R. Samjoo, K. Wesche, and M. Stubbe, editors, *Exploration into the Biological Resources of Mongolia*, volume 10. Martin-Luther University Halle-Wittenberg.
- [12] Sundaresan, S. R., Fischhoff, I. R., Dushoff, J., & Rubenstein, D. I. (2006). Network metrics reveal differences in social organization between two fission-fusion species, Grevy's zebra and onager. *Oecologia*.
- [13] Joshua, R. T., Dennis, M. W., & Bernardo, A. H. (2003). Email as spectroscopy: Automated discovery of community structure within organizations. In *Communities and technologies*, pages 81–96.
- [14] Kempe, D., Kleinberg, J., & Tardos, E. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '03*, pages 137–146.
- [15] Purnamrita, S. & Andrew, W. M. (2005). Dynamic social network analysis using latent space models. *SIGKDD Explor. NewsL.*, 7:31–40.
- [16] Chakrabarti, D., Kumar, R., & Tomkins, A. (2006). Evolutionary Clustering, ACM, pp. 554–560, [Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining].
- [17] Wang, Q., & Fleury, E., (2010). Mining time-dependent communities. In *LAWDN – Latin American Workshop on Dynamic Networks*.
- [18] Wang, Y., Wu, B., Du, N., (2008). Community evolution of social network : feature, algorithm and model". *Science And Technology*.
- [19] Boujlaleb, L., Idarrou, A., & Mammass, D., (2017). Tracking community evolution in social networks. *Journal of Theoretical and Applied Information Technology*, E-ISSN 1817-3195 / ISSN 1992-8645.
- [20] Zheng, Y., (2011). Location-based social networks: Users". In *Computing with Spatial Trajectories*, Zheng, Y and Zhou, X, Eds. Springer.
- [21] Berger-Wolf, T. Y., & Saia, J. (2006). A framework for analysis of dynamic social networks. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '06*, pages 523–528.
- [22] Bruhn, J., (2011). *The Sociology of Community Connections*". Springer Science+Business Media B.V.
- [23] Gallaher, A., & Padeld, H., (1980). *The Dying community*. Advanced seminar series. University of New Mexico Press.
- [24] Newman, M. E. J., (2006). Finding community structure in networks using the eigenvectors of matrices". *Physical Review E*, vol. 74, no. 3, pp.
- [25] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics : Theory and Experiment* 10 (P10008).
- [26] Palla, G., Barabási, A.L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature* 446:664–667.
- [27] Falkowski, T., Bartelheimer, J., & Spiliopoulou, M., (2006). Mining and visualizing the evolution of subgroups in social networks". In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, WI '06*, pages 52–58.
- [28] Takaffoli, M., Sangi, F., Fagnan, J., & Zaiane, O. R., (2010). A framework for analyzing dynamic social networks". In *Proceedings of the 7th Conference on Applications of Social Network Analysis, ASNA '10*.
- [29] Wang, Y., Wu, B., & Du, N. (2008). Community evolution of social network : feature, algorithm and model". *Science And Technology*.
- [30] Boujlaleb, L., Idarrou, A., & Mammass, D. (2017). Detecting Communities by Integrating Perspective Communities as Semantic Information". In *Proceedings of the 4th International Symposium on Social Networks Analysis, Management and Security (SNAMS-2017)* co-located with FiCloud.