A Framework for Predicting Community Behavior in Evolving Social Networks

Georgia Koloniari
gkoloniari@uom.gr
Applied Informatics Department
Thessaloniki, Greece

Georgios Evangelidis
gevan@uom.gr
Applied Informatics Department
Thessaloniki, Greece

Nikolaos Sachpenderis
sachpenderis@uom.edu.gr
Applied Informatics Department
Thessaloniki, Greece

Ioannis Milonas
imilonas@uom.edu.gr
Applied Informatics Department
Thessaloniki, Greece

ABSTRACT

Detecting and studying communities that are formed in social networks because of the interactions among their users is a very important problem with numerous applications. Such communities, similar to the real world, are dynamic and change along with the evolving social network. The goal of this paper is to propose a complete framework for addressing the problem of predicting the changes and the behavior in general of communities in evolving social networks. The framework consists of different phases that encompass all the steps required for community detection, analysis and behavior prediction. Our approach is based on modeling community evolution by multi-dimensional time series that describe the changes of each community's properties, both structural and content-based properties, through time. The prediction framework is based on rule discovery upon the multi-dimensional time series, so that based on patterns that appear in the evolution of a community's property so far, future behavior can be predicted. In the context of the proposed framework, rule discovery is extended beyond a single time series to multi-dimensional time series. Finally, exploiting the similarity between the behavior of a network's communities, their multi-dimensional time series will be used for community clustering. Thus, rule discovery can also incorporate global rules that appear in clusters of communities as well as on the network level, so as to discover global behavior patterns that characterize all the communities of a network.

CCS CONCEPTS

• **Information systems** → **Social networks**; • **Mathematics of computing** → *Cluster analysis*; *Time series analysis*; • **Computing methodologies** → *Rule learning*.

KEYWORDS

social networks, communities, evolution, prediction, time series, rule discovery

ACM Reference Format:

Georgia Koloniari, Georgios Evangelidis, Nikolaos Sachpenderis, and Ioannis Milonas. 2019. A Framework for Predicting Community Behavior in Evolving Social Networks. In *BCI 2019: 9th Balkan Conference in Informatics, September 26–28, 2019, Sofia, Bulgaria.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The wide use of social networks by numerous users produces a large volume of data and makes the task of their analysis both necessary and challenging. As a social network, we consider any network in which the nodes correspond to the entities that participate in the network, usually social media users, and the edges correspond to some type of relationship or interaction between them.

One of the problems in the context of social analysis that attracts considerable attention is community detection and analysis, that is the study of dense subgroups of nodes that are formed by users that share common interests and/or common characteristics. This problem finds many interesting applications in targeted advertising, recommender systems, etc. Even though many studies have focused on the problem of community detection, their majority ignores one of the most important properties of social networks, that is, their dynamic nature. Social networks change and evolve over time, thus, methods for analyzing static networks are not appropriate to understand their behavior. Communities in social networks also evolve along with the entire social network. Through time, communities change their size, split, merge, new communities appear and others disappear.

Clustering, that is based mostly on the structural properties of a network, such as density and centrality, comprises the main technique for community detection. By observing the clusters that are formed in different time points in a social network, one can study community evolution, evaluate community properties with respect to time, and also discern interesting properties that characterize the lifetime of a specific community. However, both social networks and communities that are formed within them are not only characterized by their structural properties but also by their content as expressed through the interests of the participating users or the user profiles and the content of their interactions.

BCI 2019, September 26–28, 2019, Sofia, Bulgaria

ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

By studying the behavior of communities through time, the goal is to provide predictions for their future behavior. Such predictions should concern both structural and content-based properties, so as to provide added value to applications that exploit communities to design successful marketing or target advertising strategies.

To this end, we propose a unified framework based on rule discovery for predicting the evolution of community properties behavior through time. The proposed framework is based on (i) modeling community properties (both structural and content-based) as time series, so as that the behavior of each community is described through a set of time series, and (ii) applying rule discovery algorithms on these time series to provide appropriate predictions. The discovered rules will follow the form: "if for a specific time interval a property presents a given behavior (left side of the rule), then after some time interval it will exhibit another specific behavior (right side of the rule)". The usefulness of such powerful rules is evident: they provide the potential for predicting the evolution of a community when specific patterns appear in the evolution of some property (appearance of left side of the rule).

Since each community is described through a set of time series, each of which models a different property, the derived rules should not be limited to one time series. Instead, rules on multi-dimensional time series, which given a subset of a community's properties will predict the evolution of one or more of their properties, can also be derived. The rule discovery framework aims to extend these rules beyond a single community to global rules by considering all communities of a social network, but also to clusters of communities based on the similarity of their behavior that can be derived by the similarity of the time series that model their properties.

In particular, the goals of the proposed framework are:

- Determining a set of structural and content-based properties for constructing time series that describe the evolution of a community.
- Discovering rules based on a single community property. For instance, given the time series that corresponds to the evolution of the size of a community for k time points, to predict its behavior, i.e., its value, for a next time interval, that is whether the size will increase, decrease, etc.
- Discovering rules based on a subset of community properties. By considering a given subset of properties as a multidimensional time series, the rule will predict the future behavior of one or more of those properties.
- Clustering communities based on the shape of the time series of their properties. That is, two communities belong to the same cluster if the time series that describe the evolution of one or more of their properties are similar.
- Discovering global rules both based on a single or a subset of community properties.

The rest of the paper is structured as follows. Section 2 presents background knowledge regarding both research on community evolution and also on rule discovery on time series. Section 3 describes the proposed framework, and Section 4 discusses the framework's contributions and novel aspects. Finally, Section 5 concludes with a summary and outlines our future plans for implementing the framework.

2 BACKGROUND KNOWLEDGE

Our framework is based on two main lines of research: (i) community detection and analysis in social networks and (ii) rule discovery for time series prediction. Next, we briefly present the state-of-the-art in both domains of research.

2.1 Community Detection and Tracking

The problem of identifying communities based on connectivity and other structural properties of social networks has been widely studied, mainly, in static networks. [1] reviews various social network analysis techniques and emphasizes the importance of communities and their analysis. [10] presents an analytical categorization of community detection methods according to the clustering technique used. A common category is algorithms that use graph partitioning and divide nodes into clusters of predetermined size aiming to minimize the number of edges connecting different clusters. Another common category comprise the divisive hierarchical algorithms that gradually remove edges between nodes to create isolated clusters. Two of the most important algorithms for community detection, the Girvan and Newman algorithm [12] and the Clauset, Newman and Moore (CNM) algorithm [7] belong to this category. The Girvan and Newman algorithm removes edges between nodes based on edge betweeness centrality, while the CNM algorithm is a fast hierarchical algorithm designed for large networks based on the concept of network modularity.

With regard to community evolution, many studies have dealt with recording and modeling communities over time, so as to draw useful analytic conclusions. In [31], concepts for assessing community evolution are introduced, such as the growth or disappearance rate of a community. At the same time, Palla et al. [24] explore the life expectancy of communities taking into account the weight of intra-community and inter-community edges. In the bibliographic review of [3], several approaches for mapping communities in successive snapshots of the network are presented. In [26], clusters are mapped between successive snapshots, and stable long-lived communities are studied. An evolutionary clustering framework was proposed in [5] and has as a basic idea that clusters should be calculated taking into account the structure of the network at earlier snapshots. Like most of the previous approaches, this is also a costly approach. The Tiles algorithm [25] efficiently discovers communities in dynamic networks by recalculating the participation of a node in a community with every new interaction.

All of the above approaches exploit only structural properties of networks to identify communities and analyze their evolution. A different approach applies classification and clustering on the content of the network. [4] deals with tag clustering and builds networks where nodes are labels and edges correspond to their co-occurrence frequency. Thus, clusters contain associated labels. Similarly, in [28] and [30], semantic analysis is applied on the co-occurence of labels in one and two real social networks respectively, using clustering based on labels. [11] applies classification based on the co-occurence of labels on multimedia social network data, and proposes a model for improving analysis using semantic tools.

In one of the few approaches that exploit both social network structure and content, Chakrabarti et al. [6] suggest a method for better classifying hypertext content on the web. In [23], the authors having identified the problem that scientific papers are usually classified on the basis of co-authorship relationships, i.e., based on the structure of the network and not the content of the papers, suggest a classification approach that relies both on co-authorship and content and eventually compare the detected communities. An important contribution to the problem of community detection appears in [32], where both the structure and the content of a social network is integrated into a new representation model. In particular, the Node-Edge Interaction Network is proposed, which contains multiple types of nodes and edges to represent both structure and content information. In addition, methods that allow the efficient update of the interaction network, while the original network evolves through time are also proposed. For community detection in the transformed interaction network random walk based algorithms are applied. Such an approach can be exploited in our framework to identify communities and their evolution so that their properties are modeled as time series.

With regard to predicting the evolution of communities, research is sparse. In [13] and [27], approaches based on structural information attempt prediction of community evolution at later times. However, these approaches treat the problem as a classification problem. They define as classes, possible events such as merging, splitting, etc. for the evolution of a community at the next snapshots, and by training appropriate models draw conclusions about the future structure of the network. In [15], similar to our proposal, time series are used to model the structural and temporal (e.g., age of nodes) properties of communities. By studying communities in different social network snapshots, changes are recorded as predefined events, in a similar manner as approaches [13] and [27]. The ARIMA model is used to predict the values of community properties, and the results are fed into classifiers that predict the next event in the evolution of a community. The approach of Diakidis et al. [9] is the only one in which structure and content-based properties are used to provide predictions, representing each community at one snapshot as a vector of its properties' values. This approach also follows the same idea of classification based on predefined events to predict the next snapshot for each community, and does not provide predictions about the individual properties.

Unlike all approaches that use classification as the predictor, our proposal is not limited to predefined events, but by finding rules it will be able to provide predictions on multiple levels, both for the overall behavior of communities, but also in more detailed degree for the development of specific properties over time, giving the opportunity to extract and correlations between different structural and content properties.

2.2 Rule Discovery in Time Series

Regarding the rule and motif discovery in time series as well as the analysis of the comparative advantages and disadvantages of the different approaches, the literature is relatively limited.

Mannila et al. [22] and Guralnik and Srivastava [14] are perhaps the first to approach the problem of prediction through rules. In particular, Mannila et al. [22] describe the term episode as a collection of events that occur over time relatively close to each other and in a certain order. Many such episodes create a rule. The work concludes with the description of WINEPI algorithm to identify all episodes from a given class of frequent episodes. The method is applicable on alarm data in telecommunication networks, user interface actions, crimes committed by a person, etc. Guralnik and Srivastava [14] deal with dynamic systems in which behavior changes significantly over time in terms of quality characteristics. They examine (a) apriori those points of time where the behavior of the characteristic changes and (b) decide which function best describes the change between time points.

In other studies ([8] and [21]), emphasis is placed on the discovery of local motifs in multidimensional time series as opposed to traditional time series analysis dealing with a single rule. Thus, they relate different patterns in the same time series or different patterns in different time series to discover the rules. In their recent work, M. Shokoohi-Yeta et al. [29] use the shape of a time series for predicting its future state. Specifically, using real number time series, they recognize repetitive patterns in the form of rules, with their left and right parts being a certain distance apart. In conclusion, they propose the MDL (Minimum Description Length) algorithm, which evaluates "candidate rules" on the basis of optimal compression of data. To evaluate the effectiveness of their method, they compare their results with three other methods [29]. But, Keogh and Lin [17] question the validity of their method and of the entirety of approaches that use clustering in a single time series, and show that such clustering can find rules even in random walk data.

In terms of time series clustering, since the mid-1990s, there is a plethora of papers that study all the parameters of the problem, such as time series representations, algorithms for time series clustering, various metric distances, etc. Specifically, to measure the distance between multidimensional time series a number of metrics have been proposed [16], [19]. Liao's literature review [20] mentions three time series clustering approaches: raw-data-based, property-based, and model-based.

[2] presents a detailed coverage of time series clustering. Three types of time series clustering are mentioned: (a) whole time series clustering, where each time series from a set of time series is assigned to a cluster, (b) subsequence time series clustering, where subsequences of a time series are clustered, and (c) time point clustering, where points that are close in time and value belong to the same cluster. The second type of clustering is the subject of too many papers due to Keogh and Lin's claim [17] that clustering at a time series level makes no sense. A complete analysis of the debate on that problem is presented in [33].

3 PROPOSED FRAMEWORK

The proposed framework aims at encompassing all processes required to address the problem of prediction community behavior in evolving social networks. Namely, the framework should cover community detection, tracking community evolution, measuring the different structural and content-based properties through time and finally providing predictions for future behavior in multiple levels. As such, one can discern several phases in the proposed network which we describe in detail below.

Phase I: Data collection and pre-processing.
The observation period of the network is split into appropriate non-overlapping time intervals, and for each of those intervals a

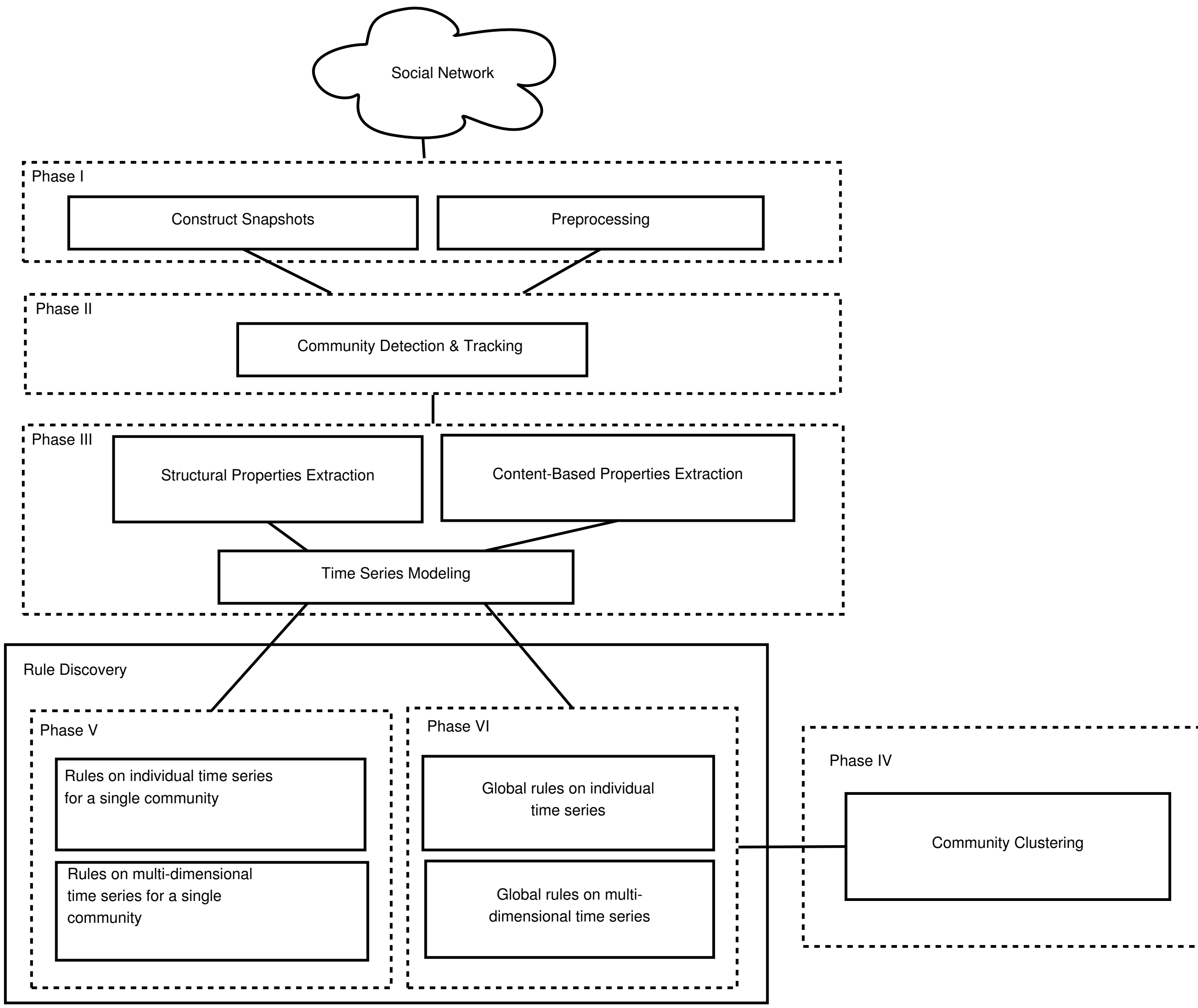

Figure 1: Overview of the proposed Framework

corresponding network snapshot that reflects the state of the network at the given interval is to be constructed. Also, this phase includes any pre-processing required to extract the content-based properties of the communities, based on the type of the content, for instance, when plain text is used, tokenization, stemming, lemmatization, etc, may be required.

Phase II: Community Detection.
For community detection the framework will utilize an evolutionary clustering algorithm that given the previous network snapshot will detect how it evolves at the next snapshot. An appropriate evolutionary clustering algorithm that takes into account both the structure and the content of the network unlike most state-of-the-art approaches will be used.

Phase III: Property Modeling.
This phase will consist of two steps, first determining appropriate structural and content-based properties, and then modeling them as time series by measuring or evaluating their values for different time points. For instance, a content-based community property can regard the popularity (number of occurrences) of the *k*-most popular terms that appear in the content of the community members interactions.

Phase IV: Community Clustering.
By using the full time series of the properties of each communities, this phase will cluster communities based on their time series similarity. For evaluating the time series similarity, we will use distance metrics such as Dynamic Time Warping (DTW) that are appropriate

for measuring the distance between time series of different lengths, as we expect that all communities will not have the same lifetime.

Phase V: Community-Level Rule Discovery.
Rule discovery will be aimed at discovering repeating motifs over either one time series modeling a single property of a community or over a multi-dimensional time series describing a subset of a community's properties. In the second case, the discovered rules are going to provide predictions either of all the properties that are included in the multi-dimensional time series or for a subset of these properties, in order to detect dependencies and correlations among them.

Phase VI: Global Rule Discovery.
The final phase aims at discovering global rules both on the level of community clusters and on network level, by considering all the communities of a network. Similarly to Phase V, global rules will be based both on single properties and combinations of them. Since similarly to users, network communities also exhibit similarities in their behavior, in this phase by exploiting the results of community clustering from Phase IV, rules that describe groups (i.e. clusters) of communities will be explored. Global rules on network level will be supported by many communities of a given network and therefore can give us information and predictions about the general behavior of the network. By applying the same methodology on data from different networks, the framework will be able to also provide comparative results about their corresponding behavior.

Figure 1 illustrates the overall architecture of the proposed framework and the components that are included in each of its six phases.

To evaluate our framework, we are going to exploit available datasets extracted from social networks that describe user interactions through time (such as [18]). To satisfy the requirements of our study, big datasets are going to be used that contain temporal information and concern a sufficiently large time period so that useful conclusions about the network's evolution can be derived. Moreover, the used datasets should combine both structural and content-based information to describe all aspects of the communities' behavior.

4 DISCUSSION

The study of community behavior in evolving social networks is one of the most important problems in analyzing social networks and benefits many applications in society and economy. Through this study, we understand the behavior of users in social networks, their dynamics and especially their behavior in groups. Our approach, one of the few that not only will study social network behaviors but aims to also predict them, can provide added value and allow the design of strategies and policies that fit better to the needs of communities and, by extension, their users with greater benefits for all involved. More specifically, in terms of both understanding and predicting the dynamics of these communities, the fact that we take into account properties of structure and content and apply our methodology to different social networks enables us to address questions such as:

- whether there is a correlation between the evolution of structural properties and content in the lifetime of a community,
- whether the general popularity of a term in the entire content of a social network is related to the lifetime of a community,
- whether there are general rules describing the behavior of all communities in a social network,
- whether there are common general rules describing community behavior in networks.

Most approaches that address the problem of community analysis on evolving social networks focus on community detection and modeling community evolution but without providing any predictions on their future behavior. On the contrary, the proposed work aims at developing a framework for predicting the future behavior of both structural and content-based community properties.

The prediction methodology is based on modeling the community properties as time series and discovering rules for predicting the future shape of the corresponding time series. The few studies that focus on the problem of community evolution prediction model the problem as a classification problem considering as classes potential events that occur in the evolution of a community. Therefore, such approaches are limited to predicting only the appearance of a predefined set of events, i.e., the predefined classes, in the lifetime of a community. On the other hand, our proposed framework is not limited by predefined events and is able to provide more detailed predictions with the use of appropriate rules, not only on the behavior of a community as a whole, but separately for both specific structural and content-based properties of the community as well.

Another novel characteristic of the proposed framework is that rule discovery goes beyond considering each community property separately. Instead, it focuses on multi-dimensional time series that describe sets of community properties, both on a single community level and also on all the communities of a network. Rule discovery in multiple levels offers opportunities to study the evolution of the communities in depth and also derive comparative conclusions between the evolution of different properties based on both structure and content and how they influence each other but also on how they influence and comprise a community's behavior as a whole.

Finally, to the best of our knowledge, community clustering based on the shape of the time series that describe their properties evolution has not been applied in the bibliography before. Combined with rule discovery for clusters of communities, the framework is able to determine different community types and derive conclusions and predictions for their future behavior.

5 CONCLUSIONS

In this paper, we present a proposed framework that envisions addressing all phases involved in providing predictions for the future behavior of communities in evolving social networks. The basic idea of the framework is to model community behavior by using time series for describing the behavior of individual community properties through time. To better encompass all aspects of social behavior by the communities, the framework will consider both structural and content-based properties of the network and will provide predictions in different levels, for a single property, for a set of properties of a single or clusters of communities.

We are currently developing a prototype of the proposed framework. We are exploring appropriate evolutionary clustering algorithms for detecting evolving communities while also experimenting with rule discovery in a single time series on corresponding social network data.

REFERENCES

[1] Mariam Adedoyin-Olowe, Mohamed Medhat Gaber, and Frederic T. Stahl. 2014. A Survey of Data Mining Techniques for Social Media Analysis. *JDMDH* 2014 (2014). http://jdmdh.episciences.org/18

[2] Saeed Reza Aghabozorgi, Ali Seyed Shirkhorshidi, and Ying Wah Teh. 2015. Time-series clustering - A decade review. *Inf. Syst.* 53 (2015), 16–38. https://doi.org/10.1016/j.is.2015.04.007

[3] Thomas Aynaud, Eric Fleury, Jean-Loup Guillaume, and Qinna Wang. 2013. *Communities in Evolving Networks: Definitions, Detection, and Analysis Techniques.* Springer New York, New York, NY, 159–200. https://doi.org/10.1007/978-1-4614-6729-8_9

[4] Grigory Begelman and Frank Smadja RawSugar. 2006. Automated Tag Clustering: Improving search and exploration in the tag space. In *WWW 2006.*

[5] Deepayan Chakrabarti, Ravi Kumar, and Andrew Tomkins. 2006. Evolutionary clustering. In *Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Philadelphia, PA, USA, August 20-23, 2006.* 554–560. https://doi.org/10.1145/1150402.1150467

[6] Soumen Chakrabarti, Byron Dom, and Piotr Indyk. 1998. Enhanced Hypertext Categorization Using Hyperlinks. In *SIGMOD 1998, Proceedings ACM SIGMOD International Conference on Management of Data, June 2-4, 1998, Seattle, Washington, USA.* 307–318. https://doi.org/10.1145/276304.276332

[7] Aaron Clauset, M E J Newman, and Cristopher Moore. 2005. Finding community structure in very large networks. *Physical review. E, Statistical, nonlinear, and soft matter physics* 70 (01 2005), 066111. https://doi.org/10.1103/PhysRevE.70.066111

[8] Gautam Das, King-Ip Lin, Heikki Mannila, Gopal Renganathan, and Padhraic Smyth. 1998. Rule Discovery from Time Series. In *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), New York City, New York, USA, August 27-31, 1998.* 16–22. http://www.aaai.org/Library/KDD/1998/kdd98-003.php

[9] Georgios Diakidis, Despoina Karna, Dimitris Fasarakis-Hilliard, Dimitrios Vogiatzis, and George Paliouras. 2015. Predicting the Evolution of Communities in Social Networks. In *Proceedings of the 5th International Conference on Web Intelligence, Mining and Semantics, WIMS 2015, Larnaca, Cyprus, July 13-15, 2015.* 1:1–1:6. https://doi.org/10.1145/2797115.2797119

[10] Santo Fortunato. 2010. Community detection in graphs. *Physics Reports* 486, 3-5 (2010), 75 – 174. https://doi.org/DOI:10.1016/j.physrep.2009.11.002

[11] Eirini Giannakidou, Ioannis Kompatsiaris, and Athena Vakali. 2008. SEMSOC: SEMantic, SOcial and Content-Based Clustering in Multimedia Collaborative Tagging Systems. In *Proceedings of the 2th IEEE International Conference on Semantic Computing (ICSC 2008), August 4-7, 2008, Santa Clara, California, USA.* 128–135. https://doi.org/10.1109/ICSC.2008.73

[12] Michelle Girvan and Mark E. J. Newman. 2002. Girvan, M. and Newman, M. E. J. Community structure in social and biological networks. Proc. Natl Acad. Sci. USA 99, 7821-7826. *Proceedings of the National Academy of Sciences of the United States of America* 99 (07 2002), 7821–6. https://doi.org/10.1073/pnas.122653799

[13] Bogdan Gliwa, Piotr Bródka, Anna Zygmunt, Stanislaw Saganowski, Przemyslaw Kazienko, and Jaroslaw Kozlak. 2013. Different approaches to community evolution prediction in blogosphere. In *Advances in Social Networks Analysis and Mining 2013, ASONAM '13, Niagara, ON, Canada - August 25 - 29, 2013.* 1291–1298. https://doi.org/10.1145/2492517.2500231

[14] Valery Guralnik and Jaideep Srivastava. 1999. Event Detection from Time Series Data. In *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, August 15-18, 1999.* 33–42. https://doi.org/10.1145/312129.312190

[15] Nagehan Ilhan and Sule Gündüz Ögüdücü. 2015. Predicting Community Evolution based on Time Series Modeling. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2015, Paris, France, August 25 - 28, 2015.* 1509–1516. https://doi.org/10.1145/2808797.2808913

[16] Leonidas Karamitopoulos, Georgios Evangelidis, and Dimitrios Dervos. 2010. PCA-based Time Series Similarity Search. In *Data Mining - Special Issue in Annals of Information Systems.* 255–276. https://doi.org/10.1007/978-1-4419-1280-0_11

[17] Eamonn J. Keogh and Jessica Lin. 2005. Clustering of time-series subsequences is meaningless: implications for previous and future research. *Knowl. Inf. Syst.* 8, 2 (2005), 154–177. https://doi.org/10.1007/s10115-004-0172-7

[18] Jure Leskovec and Andrej Krevl. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. http://snap.stanford.edu/data.

[19] Zheng-xin Li, Ke-wu Li, and Hu-sheng Wu. 2016. Similarity measure for multivariate time series based on dynamic time warping. In *Proceedings of the 2016 International Conference on Intelligent Information Processing, ICIIP 2016, Wuhan, China, December 23-25, 2016.* 15. http://dl.acm.org/citation.cfm?id=3028857

[20] T. Warren Liao. 2005. Clustering of time series data - a survey. *Pattern Recognition* 38, 11 (2005), 1857–1874. https://doi.org/10.1016/j.patcog.2005.01.025

[21] Elena Lutsiv. 2007. Association Rules Discovery in Multivariate Time Series. In *Proceedings of the SYRCoDIS 2007 Colloquium on Databases and Information Systems, Moscow, Russia, May 31 - June 1, 2007.* http://ceur-ws.org/Vol-256/submission_3.pdf

[22] Heikki Mannila, Hannu Toivonen, and A. Inkeri Verkamo. 1997. Discovery of Frequent Episodes in Event Sequences. *Data Min. Knowl. Discov.* 1, 3 (1997), 259–289. https://doi.org/10.1023/A:1009748302351

[23] Dario De Nart, Dante Degl'Innocenti, Marco Basaldella, Maristella Agosti, and Carlo Tasso. 2015. A Content-Based Approach to Social Network Analysis: A Case Study on Research Communities. In *Digital Libraries on the Move - 11th Italian Research Conference on Digital Libraries, IRCDL 2015, Bolzano, Italy, January 29-30, 2015, Revised Selected Papers.* 142–154. https://doi.org/10.1007/978-3-319-41938-1_15

[24] Gergely Palla, Albert-László Barabási, and Tamás Vicsek. 2007. Quantifying social group evolution. *Nature* 446 (05 Apr 2007), 664 EP –. https://doi.org/10.1038/nature05670

[25] Giulio Rossetti, Luca Pappalardo, Dino Pedreschi, and Fosca Giannotti. 2017. Tiles: an online algorithm for community discovery in dynamic social networks. *Machine Learning* 106, 8 (2017), 1213–1241. https://doi.org/10.1007/s10994-016-5582-8

[26] Nikolaos Sachpenderis and Georgia Koloniari. 2018. Determining interesting communities in evolving social networks. In *Proceedings of the 22nd Pan-Hellenic Conference on Informatics, PCI 2018, Athens, Greece, November 29 - December 01, 2018.* 249–254. https://doi.org/10.1145/3291533.3291567

[27] Stanislaw Saganowski, Bogdan Gliwa, Piotr Bródka, Anna Zygmunt, Przemyslaw Kazienko, and Jaroslaw Kozlak. 2015. Predicting Community Evolution in Social Networks. *Entropy* 17, 5 (2015), 3053–3096. https://doi.org/10.3390/e17053053

[28] Patrick Schmitz. 2006. Inducing Ontology from Flickr Tags.. In *Collaborative Web Tagging Workshop at WWW 2006.* Edinburgh, Scotland.

[29] Mohammad Shokoohi-Yekta, Yanping Chen, Bilson J. L. Campana, Bing Hu, Jesin Zakaria, and Eamonn J. Keogh. 2015. Discovery of Meaningful Rules in Time Series. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, August 10-13, 2015.* 1085–1094. https://doi.org/10.1145/2783258.2783306

[30] Lucia Specia and Enrico Motta. 2007. Integrating Folksonomies with the Semantic Web. In *The Semantic Web: Research and Applications, 4th European Semantic Web Conference, ESWC 2007, Innsbruck, Austria, June 3-7, 2007, Proceedings.* 624–639. https://doi.org/10.1007/978-3-540-72667-8_44

[31] Masashi Toyoda and Masaru Kitsuregawa. 2003. Extracting evolution of web communities from a series of web archives. In *HYPERTEXT 2003, Proceedings of the 14th ACM Conference on Hypertext and Hypermedia, August 26-30, 2003, Nottingham, UK.* 28–37. https://doi.org/10.1145/900051.900059

[32] Chang-Dong Wang, Jian-Huang Lai, and Philip S. Yu. 2014. NEIWalk: Community Discovery in Dynamic Content-Based Networks. *IEEE Trans. Knowl. Data Eng.* 26, 7 (2014), 1734–1748. https://doi.org/10.1109/TKDE.2013.153

[33] Seyedjamal Zolhavarieh, Saeed Aghabozorgi, and Ying Wah Teh. 2014. A Review of Subsequence Time Series Clustering. *The Scientific World Journal* 2014 (2014), 1–19. https://doi.org/10.1155/2014/312521