Dynamic Centrality for Directed Co-author Network with Context

Alisa Babskova, Jan Martinovič, Pavla Dráždilová, Kateřina Slaninová

Dynamic Centrality for Directed Co-Author Network with Context

Alisa Babskova[1], Jan Martinovič[1,2], Pavla Dráždilová[1], and Kateřina Slaninová[1,2]

[1] Faculty of Electrical Engineering, VŠB - Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava, Czech Republic
[2] IT4Innovations, VŠB - Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava, Czech Republic
{alisa.babskova.st, jan.martinovic, pavla.drazdilova, katerina.slaninova}@vsb.cz

Abstract. Co-author network is a typical example of dynamic complex network, which evolves and changes over time. One of the ways how to capture and describe the dynamics of the network is determination of *Stationarity* for detected communities in the network. In the paper, we have proposed the modified *Stationarity*, which is focused only on co-authors of a given author and not on the whole community to which the author belongs. Therefore, this modified *Stationarity* is defined for each author in the network and is perceived as dynamic centrality. The relations in homogeneous co-author network are not only set by the number of common publications, but are given by a context to terms used by the author extracted from the article titles. This dynamic centrality calculates with the evaluation by context of directed edges in co-author network. Such modified *Stationarity* gives us information about stability or dynamics of the author's neighbourhood that influences her/him, or about the stability and dynamics of the author's neighbourhood, which the author influences in relation to context.

Keywords: Co-author Network, Directed Network, Context, Dynamic Networks, Stationarity

1 Introduction

Co-author network of computer science bibliography (DBLP)[3] represents an example of dynamic complex network which can be analysed by various methods from the point of view focused on network evolution [2, 3]. It records not only the evolution of the whole network, but we are able to explore the evolution of author communities or selected individual authors during 34 years. It is also possible to investigate DBLP as a heterogeneous bibliographic network which contains multiple types of objects, such as authors, venues, topics and papers, as well as multiple types of edges denoting different relations among these objects. Several researchers [11, 2, 1] deal with analysis of complex networks evolution with the focus on the community analysis and their evolution in homogeneous networks [11] or heterogeneous networks [15]. Other articles [14] are focused on prediction of relations in heterogeneous bibliographic networks.

[3] Computer science bibliography (DBLP) website: http://dblp.uni-trier.de/

One of the methods for evolution analysis of co-author networks is evolutionary clustering for heterogeneous networks [6, 15]. Gupta et al. present an algorithm which performs such an agglomerative evolutionary clustering which is able to show variations in the clusters over time with a temporal smoothness approach. Network evolution can be also analysed on the centrality basis. For example in [8], authors study a model of network evolution where links are created or removed based on the centrality of the nodes incident to the links. Authors of article [7] propose a new centrality framework, called composite centrality (CC). The idea behind the CC-framework is that one first defines a set of characteristics of interest, and then chooses appropriate network (centrality) measures. Authors of paper [4] apply the combined approach of a topic modelling algorithm and a pathfinding algorithm to find whether authors tend to co-author with or cite researchers sharing the same research topics. In citation networks, which are directed and weighted, the authors of [9] present centrality for dynamic networks that measures the number of paths that exist over time in a network. They use this metric to rank nodes by how well connected they are over time to the rest of the network.

Our proposed method for stationarity calculation allows to evaluate stability of author's co-authors. The network is defined by the context. Therefore, the stationarity is also related to the stability of author's research domains. The obtained centrality is intended to be used for dynamic networks, in which the set of co-authors as well as the set of terms vary during time periods. The proposed approach allows us to define the evaluation of the directed edges within the co-authors network. We are able to evaluate the edges from the author towards his/her co-authors and otherwise. These weights are used for calculation of $Stationarity_{in/out}$, which tells us how the author is stable in relation to a set of co-authors and a set of used terms. Dynamic change centrality is defined in [5]. However, this centrality does not work with weighted graph. The change centrality of a node is a measure of the change of its connections over time, taking into account its adjacent nodes, the adjacent nodes of the latter and so on. The weight of changes of near and far neighbours depends on the choosing coefficients of the linear combination.

In this paper, a directed co-author network is constructed using context. Context is created by extracted terms which author used in titles of articles. Different sets of terms used by different authors then give the orientation and the weight in a new evaluation of relations in the co-author graph using the context. Therefore, a dynamics of evaluated directed network is determined not for the whole network or co-author communities, but for individual authors and their co-authors. Our proposed dynamic centrality describes the author influence into his/her neighbourhood ($Stationarity_{out}$) or the influence of his/her neighbourhood into the author ($Stationarity_{in}$).

This paper is organized as follows: The proposed approach of creation of context with a detailed description of the process which leads to a construction of new directed co-author network is presented in Section 2.2. Section 3 introduces principles of network's dynamic, dynamic evaluation of the neighbourhood of the selected author and our approach in the directed network created from DBLP and with the context. Experiments for directed co-author network with the context for finding dynamic centrality $Stationarity_{in/out}$ of selected author is presented in Section 4. A conclusions can be seen in Section 5.

2 Network with Context on the DBLP

This section describes the way how to create a homogeneous network of co-authors from the originally heterogeneous network DBLP. The evaluation of edges between the authors is represented by context based on the terms extracted from the article titles.

2.1 Digital Bibliography Library Project

DBLP (Digital Bibliography Library Project) is a computer science bibliography database hosted at University of Trier, in Germany. It was started at the end of 1993 and listed more than 2.3 million publications in May 2014. These articles were published in Journals such as VLDB, the IEEE and the ACM Transactions and Conference proceedings [10]. DBLP has been a credible resource for finding publications, its dataset has been widely investigated in a number of studies related to data mining and social networks to solve different tasks such as recommender systems, experts finding, name ambiguity, etc. Even though, DBLP dataset provides abundant information about author relationships, conferences, and scientific communities, it has a major limitation that is its records provide only the paper title without the abstract and index terms.

2.2 Author's Relationships with Context to Terms

We can create a one-mode graph from a bipartite graph [16], where the bipartite graph captures relations between to different types of groups (authors and their join publications). These author's relations would then be evaluated measuring the intensity of their shared activity. We have added context obtained from a data collection using term extraction [12] for the evaluation of relations.

Our method that we have used for more precise evaluation of the intensity of person's relations was to ascertain the context among authors and the terminology they used in article titles in DBLP.

We use terms for evaluation of the relation between co-authors. We extend standard evaluation of the relation, which is based on the number of the join publications or articles, by a factor that represent context between author and terms selected from the term set.

Term set is understood as a collection of all keywords, which are extracted from titles of articles. A detailed description of *term set* was published in [12].

Let A be a set of all authors in dataset. We define a single author A_i. For A_i, it is evaluated the strength of association with the other co-author. The strength of participation could be computed in a way that we go through all the author's publications while marking all the participated co-authors. Let set P be a set of all publications in DBLP and P_{A_i} be a set of all publications of author A_i.

The $Association$ between the co-authors A_i and A_j can be defined by Jaccard coefficient that reflects mainly the proximity of both co-authors from number of their join publications:

$$Association(A_i, A_j) = \frac{|P_{A_i} \cap P_{A_j}|}{|P_{A_j}| + |P_{A_i}| - |P_{A_i} \cap P_{A_j}|} \tag{1}$$

If this method is applied to all the authors, we obtain weighted undirected graph of co-author network. This approach was inspired by [4].

If we define a set T as the set of all terms in the input text (titles of articles in DBLP) and T_{A_i} as the set of all the terms that could be found in titles of articles of author A_i, then t_k is the term belonging to the author A_i (t_k *in* T_{A_i}). Thus, we define $(t_k \textit{ in } T_{A_i})$ as the number of occurrences of term t_k in the input text T_{A_i}. This number is then approximated by the number of occurrences of term t_k in the all titles of articles $(t_k \textit{ in } T)$. The higher value, the less relevant term t_k becomes. In addition, the result is approximated by T_{A_i}, because there is an assumption that T_{A_i}, which has a high cardinality, lower the importance of the individual terms, while low cardinality indicates that the author has only one subject matter. We can define the *relevance of author's terms* as:

$$R(A_i,t_k) = \frac{(t_k \textit{ in } T_{A_i})}{(t_k \textit{ in } T) + |T_{A_i}| - (t_k \textit{ in } T_{A_i})}. \tag{2}$$

And in normalized form:

$$R_{Norm}(A_i,t_k) = \frac{R(A_i,t_k)}{\max(R(A_i,t_1),\ldots,R(A_i,t_{|T_{A_i}|}))}. \tag{3}$$

The $ContextScore(A_i,A_j)$ of undirected edges is calculated by following Eq.4 for all terms in $(T_{A_i} \cup T_{A_j})$:

$$\begin{aligned} ContextScore(A_i,A_j) &= ContextScore(A_j,A_i) = \\ &= Association(A_i,A_j) \sum_{t_k \in (T_{A_i} \cup T_{A_j})} R_{Norm}(A_i,t_k)\, R_{Norm}(A_j,t_k) \end{aligned} \tag{4}$$

These equations form an evaluation in undirected graphs, but do not describe sufficiently the situation in the co-author network. Relationships between co-authors are not equal in both directions. Due to this reason, we have created an evaluation for directed edges. The undirected relation weight includes relevancies of both authors in the evaluation of common relation (times $Association$). In directed graph, the relation weight includes only the relevance of one author (times $Association$) to define the influence (or power of the influence) of one author to another. Relevancy then represents the scope of his/her interest within the all terms.

Then the *ContextScoreD* of directed edges is calculated by the next Eq. 5 for all terms in T_{A_i}:

$$ContextScoreD(A_i,A_j) = Association(A_i,A_j) \sum_{t_k \in T_{A_i}} R_{Norm}(A_i,t_k) \tag{5}$$

Similar situation is for the insufficient evaluation of edges due to time periods. A evaluation of edges in directed or undirected graphs depends on the time. Relationships between co-authors are not same in different time periods. So, we created an evaluation for edges in specified time periods. We calculate only with publications in the specified time period. The definition of the $Association(A_i,A_j)$ is extend to definition of the $Association(A_i,A_j,t_0,t_{max})$ where t_0 is the begin of the selected time period and the t_{max}

is the end of the selected time period and $P_{A_i}^{t_0,t_{max}}$ is the set of publications of author A_i in time period $< t_0, t_{max} >$.

$$Association(A_i, A_j, t_0, t_{max}) = \frac{|P_{A_i}^{t_0,t_{max}} \cap P_{A_j}^{t_0,t_{max}}|}{|P_{A_j}^{t_0,t_{max}}| + |P_{A_i}^{t_0,t_{max}}| - |P_{A_i}^{t_0,t_{max}} \cap P_{A_j}^{t_0,t_{max}}|} \tag{6}$$

The *ContextScoreP* is calculated by the next Eq. 7 for selected time period and for terms used by author A_i in this time period $T_{A_i}^{t_0,t_{max}}$:

$$ContextScoreP(A_i, A_j, t_0, t_{max}) = \sum_{t_k \in T_{A_i}^{t_0,t_{max}}} ContextScoreP(A_i, A_j, t_k, t_0, t_{max}) \tag{7}$$

3 Dynamic network analysis

Dynamic network analysis (DNA) varies from traditional social network analysis. DNA could be used for analysis of the non static information of nodes and edges of social network. DNA is a theory in which relations and strength of relations are dynamic in time and the change in the one part of the system is propagated through the whole system, and so on. DNA opens many possibilities to analyse and study the different parts of the social networks. It is possible study behaviour of individual communities, persons or the whole graph of the social network. We focus to analyse the behaviour of neighbourhood (exactly adjacent vertices) of selected author extracted from the network during a time period. The proposed approach which use dynamic metrics is inspired by work of Palla et al. [13].

The *AutoCorrelation* function $C(A_i, t_v, t)$ is used to quantify the relative overlap between two following neighbourhoods $N(A_i, t_v) = \{A_j; ContextScore(A_i, A_j, t_v) > 0\}$ of the same author A_i at t time steps apart:

$$C(A_i, t_v, t) = \frac{|N(A_i, t_v) \cap N(A_i, t_v + t)|}{|N(A_i, t_v) \cup N(A_i, t_v + t)|} \quad i = 1, \ldots, |A|, \tag{8}$$

where $|N(A_i, t_v) \cap N(A_i, t_v + t)|$ is the number of common nodes (members) in $N(A_i, t_v)$ and $N(A_i, t_v + t)$, and $|N(A_i, t_v) \cup N(A_i, t_v + t)|$ is the number of nodes in the union of $N(A_i, t_v)$ and $N(A_i, t_v + t)$.

Palla et al. [13] evaluate communities in the network using *AutoCorrelation* function. However, we are interested in dynamics of individual nodes in the network and their neighbourhood rather then dynamics of different communities in the network. Therefore, we have defined *Stationarity* for $N(A_i, t_v)$, a set of the all neighbour nodes.

Provided that we consider for each moment an unitary relation weight $w(A_i, A_j, t_v) = 1$ then Eq. 8 can be modified to Eq. 9:

$$C(A_i, t_v, t) = \frac{|N(A_i, t_v) \cap N(A_i, t_v + t)|}{|N(A_i, t_v) \cup N(A_i, t_v + t)|} = \tag{9}$$
$$= \frac{\sum\limits_{A_j \in (N(A_i, t_v) \cup N(A_i, t_v + t))} w(A_i, A_j, t_v) w(A_i, A_j, t_v + t)}{\sum\limits_{A_j \in (N(A_i, t_v) \cup N(A_i, t_v + t))} (\max(w(A_i, A_j, t_v), w(A_i, A_j, t_v + t)))^2}$$

Then consider, that the time axis is equidistantly divided into the years, for example $t_0 = 2000, t_1 = 2001, \ldots, t_{max} = 2014$ and t is 1 year. The *Stationarity* of neighbourhood of author A_i is defined as the average *AutoCorrelation* between subsequent states:

$$\zeta(A_i) = \frac{\sum_{t_v=t_0}^{t_{max}-1} C(A_i, t_v, t)}{t_{max} - 1 - t_0}, \tag{10}$$

where t_0 denotes the begin of the observation, t_{max} is the end of the observation and t is a step. Thus, $(1-\zeta)$ represents the average ratio of members changed in the period [13].

Authors of the paper [13] found that the auto-correlation function decays faster for the larger communities, showing that the membership of the larger communities is changing at a higher rate. In contrast, they said that small communities change at a smaller rate with their composition being more or less static. The *Stationarity* was used to quantify static aspect of community evolution.

We extend our approach for the directed network with context which is created from terms. We look on the *Stationarity* of neighbourhood from directed point of view. The directed edges evaluated by $ContextScoreD(A_i, A_j)$ describe the influence power of author A_i into author A_j. AutoCorrelation is defined by a number of neighbours of the selected node. Due to this reason, the original definition would be $C_{in} = C_{out}$. However, this approach is not sufficient. Therefore, we have decided to eliminate a specific amount of edges by the following rules:

- determine $diff_{ij} = |ContextScoreD(A_i, A_j) - ContextScoreD(A_j, A_i)|$ for all $i, j = 1, \ldots, |A|$.
- create distribution of the differences $diff_{ij}$ and select the value $bound = 0.01$ as threshold.
- if $diff_{ij} < bound$ then same edges remain with $ContextScoreD(A_i, A_j)$ and $ContextScoreD(A_j, A_i)$ else delete weaker directed edge and stronger edge has a new weight $w(strongerEdge) - w(weakerEdge)$.

We have left the edges in both directions, if the authors influence each other by the nearly same power. If one of the authors influences the other *bound* more, the stronger edge has been left during the reduction.

Then we definite the AutoCorrelation in directed way by the Eq.11 The *Auto-Correlation* function $C_{in/out}(A_i, t_v, t)$ is used to quantify the relative overlap of directed weighted edges between two neighbourhoods $N_{in/out}(A_i, t_v) = \{A_j; ContextScore(A_{j/i}, A_{i/j}, t_v) > 0\}$ of the same author A_i at t time steps apart:

$$C_{in/out}(A_i, t_v, t) = \frac{|N_{in/out}(A_i, t_v) \cap N_{in/out}(A_i, t_v + t)|}{|N_{in/out}(A_i, t_v) \cup N_{in/out}(A_i, t_v + t)|} \quad i = 1, \ldots, |A|. \tag{11}$$

The *StationarityD* of neighbourhood of author A_i in directed graph is defined as the average *AutoCorrelation* between subsequent states:

$$\zeta_{in/out}(A_i) = \frac{\sum_{t_v=t_0}^{t_{max}-1} C_{in/out}(A_i, t_v, t)}{t_{max} - 1 - t_0}, \tag{12}$$

The more increases $C_{in}(A_i, t_0, t)$, the more dynamically is A_i influenced by its neighbourhood. If $\zeta_{in}(A_i) = 1$ then the influence of neighbourhood into A_i is more static in time. If $\zeta_{in}(A_i) < 1$ then the influence of neighbourhood is more dynamic.

The more increases $C_{out}(t)$, the more dynamic is the influence of A_i into its neighbourhood. If $\zeta_{out}(A_i) = 1$ then the influence of A_i into its neighbourhood is more static. If $\zeta_{out}(A_i) < 1$ then the influence of A_i into its neighbourhood is more dynamic.

4 Experiments

In general, $Stationarity$ described in Section 3 determines the dynamics or the statics of the author's neighbourhood. We have generated weighted directed co-author graphs for each year from 1980 to 2014 in the experiments. The first phase of the experiments was focused on directed edges weighted by context $ContextScoreP(A_i, A_j, t_0, t_{max})$, see Eq. 7. Such obtained graphs contained two edges with opposite directions and with different evaluation between each two co-authors. As the edge evaluation is given by context, we can claim that the edge direction for each author express his/her influence to his/her co-authors (out) or the influence of his/her co-authors to him/her (in). Based on the consideration about the graph reduction and the reduction of less important edges due to obtaining a real image of the author's neighbourhood, we have defined $diff_{ij} = 0,01$ and have removed the edges according to the method described in Section 3. Then, we have calculated for selected authors $AutoCorrelation_{in/out}$, see Eq. 11 and $Stationarity_{in/out}$, see Eq. 12.

Based on the idea that the edge direction between the authors defines the influence of his/her publication activity to his/her co-authors, we have to explain the meaning of $Stationarity_{in}$ and $Stationarity_{out}$. $Sationarity_{in}$ in this case means the measure of stability or dynamics of the neighbourhood, which has the influence to the selected author. $Sationarity_{out}$ in this case means the measure of stability or dynamics of the neighbourhood, to which the selected author has the influence. Tab. 1 shows the values of $AutoCorrelation_{in}$ and $AutoCorrelation_{out}$ for the selected author A6. As we can see, the values did not much changed for the author A6 during the years 2008-2013. Moreover, the values of $AutoCorrelation_{in/out}$ are very close for each year. That means that neighbourhood N_{in} and N_{out} has changed in a relatively similar way in time. In addition, we can notice that the values of $AutoCorrelation_{in/out}$ were for the author A6 the lowest in years 2011 - 2012. This means that the author has changed the group of co-authors during this time period. Therefore, his/her neighbourhood seems to be more dynamical.

AutoCorrelation	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013
C_{in}	0.24242	0.23333	0.22222	0.11363	0.22727
C_{out}	0.24137	0.20689	0.20454	0.09302	0.22727

Table 1. $AutoCorrelation_{in/out}$ for selected author A6 and selected times

We have concentrated on several selected authors during last years in the experiments. Table 2 and Table 3 show the selected authors and their $Stationarity_{out}$ and $Stationarity_{in}$ for different time periods. The values of $Stationarity_{out}$ and $Stationarity_{in}$ determine a neighbour stability of selected authors for particular time periods.

The selection of the authors was not random; we have selected the authors, whose publication activity is known for us, and for whom we also know their co-authors. We also have selected the authors with different publication activity. Due to this selection, we have ensured the suitable test data collection which we are able to intuitively assess and verify.

However, in our experiments, we do not use IDs nor author names, but we have done anonymisation by our own identification A1-A6.

We can see the values of $Stationarity_{in}$ and $Stationarity_{out}$ for the author A1 in Tab. 2 and Tab. 3. The values are absolutely identical, which in our evaluation means that the dynamics of the neighbourhood that influences the author A1 and the dynamics of the neighbourhood that the author A1 influences is the same. Since the values are small, we are talking about a relatively dynamic neighbourhood of the author A1. Very similar situation is for the author A2 with the difference that his/her neighbourhood is more stable then the neighbourhood of the author A1. Observing the author A3, see Tab. 3, we can find the gradual increase of the values of $Stationarity_{out}$. This can be interpreted as a possible stabilisation of the neighbourhood, to which has the author A3 influence. It can be possible to predict its better stabilisation in the future. Considering the neighbourhood, which has the influence to the author A3, we can see in Tab. 2 that $Staionarity_{in}$ stays nearly on the same value during the analysed time period. That means that the co-author community of the author A3 that influences him/her is permanently dynamic and do not stabilises.

Author	$\zeta_{in} 2005-2010$	$\zeta_{in} 2006-2011$	$\zeta_{in} 2007-2012$	$\zeta_{in} 2008-2013$
A1	0.17334	0.17123	0.16423	0.16728
A2	0.22614	0.23612	0.22833	0.21554
A3	0.24186	0.27344	0.27344	0.29395
A4	0	0	0.08	0.18
A5	0	0	0	0.033
A6	0.22051	0.21232	0.180505	0.20777

Table 2. $Stationarity_{in}$ for selected authors and selected time period

5 Conclusion

In the paper, the authors proposed a modified method for determination of $Stationarity$ in a directed network. As the edge evaluation by $ContextScoreD$ means the knowledge scope, which one author can provide the other author, the $Stationarity_{out}$ during the time corresponds with the influence power, which one author could have to the other

Author	ζ_{out} 2005 − 2010	ζ_{out} 2006 − 2011	ζ_{out} 2007 − 2012	ζ_{out} 2008 − 2013
A1	0.17334	0.17123	0.16423	0.16728
A2	0.21446	0.22445	0.22375	0.21554
A3	0.22282	0.24504	0.23852	0.24236
A4	0	0	0.08	0.18
A5	0	0	0	0.075
A6	0.19422	0.19068	0.16129	0.19462

Table 3. $Stationarity_{out}$ for selected authors and selected time period

co-authors. Contrary to the previous statement, $Stationarity_{in}$ during the time corresponds with influence power from the other co-authors to the given author. Presented experiments show that the stability measure of the selected authors is low and the set of co-authors change in time. Moreover, the influence power of the author to his/her neighbourhood differs from the influence power from his/her neighbourhood to the author. We intent to focus on other types of weighted directed networks, in which is important to determine $Stationarity$ of its members in the future.

Acknowledgments

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) and the national budget of the Czech Republic via the Research and Development for Innovations Operational Programme, by the project New creative teams in priorities of scientific research (reg. no. CZ.1.07/2.3.00/30.0055), supported by Operational Programme Education for Competitiveness, and co-financed by SGS, VSB - Technical University of Ostrava, Czech Republic, under the grant No. SP2014/154 Complex network analysis and prediction of network object behavior.

References

1. L. Backstrom. *Group formation in large social networks: membership, growth, and evolution*, pages 44–54. ACM Press, 2006.
2. A.-L. Barabâsi, 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*, 311(3):590–614, 2002.
3. I. Belykh, M. di Bernardo, J. Kurths, and M. Porfiri. Evolving dynamical networks. *Physica D: Nonlinear Phenomena*, 267(Complete):1–6, 2014.
4. Y. Ding. Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *Journal of informetrics*, 5(1):187–203, 2011.
5. P. Federico, J. Pfeffer, W. Aigner, S. Miksch, and L. Zenk. *Visual Analysis of Dynamic Networks Using Change Centrality*, pages 179–183. IEEE, 2012.
6. M. Gupta, C. C. Aggarwal, J. Han, and Y. Sun. Evolutionary clustering and analysis of bibliographic networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, pages 63–70. IEEE, 2011.

7. A. C. Joseph and G. Chen. Composite centrality: A natural scale for complex evolving networks. *Physica D: Nonlinear Phenomena*, 267:58–67, 2014.
8. M. D. Koenig and C. J. Tessone. Network evolution based on centrality. *Physical Review E*, 84(5), 2011.
9. K. Lerman, R. Ghosh, and J. H. Kang. Centrality metric for dynamic networks. In *Proceedings of the Eighth Workshop on Mining and Learning with Graphs*, pages 70–77. ACM, 2010.
10. M. Ley. The dblp computer science bibliography: Evolution, research issues, perspectives. *LNCS*, 2476:1–10, 2002.
11. Y.-r. Lin. Facetnet : A framework for analyzing communities and their evolutions in dynamic networks. *Social Networks*, pages 685–694, 2008.
12. S. Minks, J. Martinovic, P. Drázdilová, and K. Slaninová. Author cooperation based on terms of article titles from dblp. In *IHCI2011*, 2011.
13. G. Palla, A.-L. Barabási, and T. Vicsek. Quantifying social group evolution. *Nature*, 446:664–667, 2007.
14. Y. Sun, R. Barber, M. Gupta, C. C. Aggarwal, and J. Han. Co-author relationship prediction in heterogeneous bibliographic networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, pages 121–128. IEEE, 2011.
15. Y. Sun, J. Tang, J. Han, M. Gupta, and B. Zhao. Community evolution detection in dynamic heterogeneous information networks. In *Proceedings of the Eighth Workshop on Mining and Learning with Graphs*, MLG '10, pages 137–146, New York, NY, USA, 2010. ACM.
16. K. A. Zweig and M. Kaufmann. A systematic approach to the one-mode projection of bipartite graphs. *Social Network Analysis and Mining*, 1(3):187–218, 2011.