OLAP Cube-based Graph Approach for Bibliographic Data

Wararat Jakawat, Cécile Favre, and Sabine Loudcher

Université de Lyon (ERIC LYON 2), France
{wararat.jakawat, cecile.favre, sabine.loudcher}@univ-lyon2.fr

Abstract. There is a growing number of different research fields that are concerned with bibliographic data analysis. In many cases, data of interest can be described as heterogeneous information networks. To explore knowledge from that networks in a multidimensinal way, OLAP (Online Analytical Processing) analysis helps users to access data from different points of views. The ability of OLAP for analyzing classical data is clear. However it must be adapted to provide networked data by considering both nodes and edges. In order to take into account linked data in OLAP on networks, we propose a conceptual graph model to represent bibliographic networks. Then we propose graphs enriched by cubes. Each node and edge of the considered network are described by a cube. It allows the user to quickly analyze the information summarized into cubes. Our proposal also solves the slowly changing dimension problem in OLAP analysis. To illustrate our approach, we integrate three bibliographic databases and a computation process is defined according to user's needs. Then our implementation shows results on a real data.

1 Introduction

Information networks are ubiquitous due to the popular use of Web, blogs and various kinds of online databases. Networks can be homogeneous or heterogeneous networks. Homogeneous networks contain a single object type and a single edge type such as friends network, authors network and movies network. Heterogeneous networks are composed of multiple node and edge types. For example, an author-paper network is an heterogeneous network with two types of nodes (authors and papers) and three types of edges ("written" between authors and papers, "co-author relationship" between authors and the last one relates papers written by the same author). A network can also be a multidimensional network with multiple node attributes and edge attributes [8]. We take the example of bibliographic data. In many fields, bibliographic databases store a collection of information such as the publication title, authors, year, etc. To analyze bibliographic data, there are many types of analysis (Statistics, Data Mining, Graph Theory, OLAP analysis, etc.) to achieve different objectives in bibliometrics (relationship studying, ranking, community mining, etc.). Among these different types of analysis, we are more interested by OLAP analysis (Online Analytical

Processing). OLAP provides the flexibility for navigating into data, for summarizing data at different granularity levels and from different points of view.

Traditional OLAP did a great job on structured data, but OLAP faces challenges in processing networked data and it is called Graph OLAP. In several recent approaches in Graph OLAP, a cube is created for a graph to provide multidimensional and multilevel views [9]. In these approaches, we regret that the interactions among objects are still hidden and that the slowly changing dimension problem is not taken into account [1]. For example, an author, Y. Sun, published a paper when he was at Northeastern University. Then he published another paper when he was at University of Illinois. There are two publications of Y. Sun, one for each university. But from the authors network, if the user does an OLAP operation like a Roll-Up in order to see the institutions network, these two papers will be counted for both universities, and it is an incorrect answer. In this case, networked data is non-summarizable: a higher level network cannot be computed solely from the lower level network without accessing raw data. So OLAP must be adapted to provide networked data by considering both data objects and the interactions among objects. In order to take into account linked data in the OLAP analysis of networks, we previously introduced a framework to be able to analyze various networks built from bibliographic data [15]. In the present paper, we extend the previous framework. Our contributions can be summarized as follows:

- We use the properties of graph theory and we present a conceptual graph model for bibliographic networks. Its content comes from multiple bibliographic databases in a way that allows us to build several different networks such as co-authorships, institutions of author, etc. The conceptual model is mapped easily to support a variety of use cases.
- In order to adapt OLAP to multidimensional networks by considering both nodes and edges, we propose graphs enriched by cubes. Each node or edge is weighted by an OLAP cube. It allows the user to quickly analyze information that has been summarized into cubes and by viewing the graph. It supports Graph OLAP operations such as informational and topological operations and it solves the slowly changing dimension problem.
- We evaluate our proposal on real data set and we provide examples to show how using our tool to analyze data.

The remainder of this paper is organized as follows. Section 2 briefly reviews the related work. Section 3 presents our proposal of graphs with cubes. In Section 4, we present the implementation and we show results on real data set. Section 5 concludes this paper and gives future directions.

2 Related Work

In a previous paper, we have already surveyed research work that combines OLAP and informational networks [9]. In this paper, we focus only on OLAP and bibliographic networks.

There are different types of design in order to analyze bibliographic data. The first one is a model based on the entity-relationship diagram [3]. The second model is based on the classical multidimensional model used in data warehouses [2, 5]. These models use relational databases to store bibliographic data. So, they deal poorly with edges and making complex queries is not easy because several join operations must be added to answer users' queries. Moreover, traditional OLAP analysis can not be done on a graph.

The concept of Graph OLAP was first proposed by J. Han's team [5, 6, 8]. Chen *et al.* presented basic definitions of OLAP on information networks and they introduced a framework for Graph OLAP [5]. They provided a cube of graph where each cell stores a network instead of a numeric value. Two kinds of OLAP dimensions were defined (informational and topological dimensions) with two kinds of OLAP operations to navigate on the dimensions. The first operation is the informational OLAP. For example, venue and time in author-paper network are two informational attributes. They are used as information dimensions. For instance, they allow to build a network of authors for the ICEIS Conference for all years and another one for the data mining field in 2010. The second operation is the topological OLAP. For example, the network of authors can be generalized by merging all authors of a same institution as one node and building a new graph at the institution level. In this more generalized network, an edge between Stanford and the university of Lyon will aggregate all collaborations occurred between Stanford's authors and the authors of the university of Lyon. However, Chen *et al.* did not mention how to design model for heterogeneous networks. Hence, Yin *et al.* answered this problem by defining a concept of entity dimensions to support two dimensions of heterogeneous networks [10]. There are two novel operations of entity dimensions named Rotate and Stretch, which are able to mine edges between different nodes.

While Qu *et al.* focused on an efficient topological OLAP, they presented two techniques (T-Distributiveness and T-Monotonicity) in order to achieve efficient query processing and cube materialization [6]. Zhao *et al.* defined the concept of multidimensional networks to abstract the real networks. They introduced a new multidimensional model, called Graph Cube [8]. They worked with structure-enriched aggregate networks and they proposed a new type of query, called crossboid query in contrast with traditional queries named cuboid query.

The closest works to those of Han's team are those of Tian *et al.* [7]. Tian *et al.* proposed new operations for summarizing graphs. The first one, called SNAP, can produce a summary graph by grouping homogeneous nodes. Moreover, users can control the different resolutions of summaries by a k-SNAP operation.

According to Chen's framework, only nodes are described by attributes. However, in reality, edges are associated with attributes as well. For example, co-authorship network contains authors as nodes and collaboration relationship as edges. The relationships may be described by time or the papers they wrote together. To solve this problem, Zhang *et al.* and Wang *et al.* proposed models to deal with both node and edge attributes. Zhang *et al.* defined a new multidimensional network that contains node and edge attributes [11]. Node attributes were defined as dimensions in a graph cube while edge attributes were defined as dimensions in a data cube. While, Wang *et al.* proposed a new conceptual model with an hyper graph [13]. Graph aggregation is performed on node and edge attributes. The aggregated graph is a multigraph, where several edges can be between two nodes. It allows users to see the different views.

In a different way, Kaya *et al.* developed three different networks (authors, topics and venues) with a cube-based modeling method [14]. In these networks, each node is represented by a data cube which is analyzed by OLAP operations.

To sum up the short related work about OLAP on bibliographic data, we can add two remarks. The first one is about the slowly changing dimension problem [1]. This problem happens when an object (a fact, a node, etc.) changes its content over time and when this causes a change in the structure. For example, as we said in introduction, the author, Y. Sun, published a paper when he was at Northeastern University then he published another paper when he was at university of Illinois. To the best of our knowledge, the existing approaches in Graph OLAP are not complete with this problem.

The second remark is about the visualization of a multidimensional and multilevel view over graphs. For example, a cube, with a venue dimension and time dimension, can contain a cell for ($ICDE$, 2008) and another one for ($DOLAP$, 2008) cell. In the first Graph OLAP approaches, in each cell there is a graph showing collaborations between authors for this venue and this year. Between two authors, we can see the collaborations only according to the venue and the year, we don't see a global view of the collaborations. Furthermore, Wang *et al.* proposed a graph with multiple edges. However, their approach needs to summarize a set of graphs with multiple edges and it is a complex task. In contrast, Zhang *et al.* used a single graph as input rather than a set of graphs. Kaya *et al.* presented three networks where each node is represented by a cube.

Thanks to the related work, we can say that we want to :

– introduce a conceptual model for bibliographic networks based on graph theory and not on the entity-relationship diagram.
– take into account the structure of the network in order to do topological OLAP operations and not only classical or informational OLAP operations.
– deal with heterogenous networks and not only homogeneous networks.
– consider both node and edge attributes.
– have a global view of the network with multidimensional information.
– take into account the slowly changing dimension problem.

To extend OLAP on information networks, this paper presents graphs enriched by cubes. The global idea is that each node or each edge is couple with a cube according to user's requirements. This graph model supports OLAP operations for analysis.

3 Graphs Enriched by Cubes

In this section, we introduce graphs enriched by cubes. We first give an overview of the overall process, then we explain the approach to create graphs with cubes. Afterwards, we extend the OLAP operations to graphs enriched by cubes.

3.1 The Process

Figure 1 illustrates all of the components of our architecture in three steps.

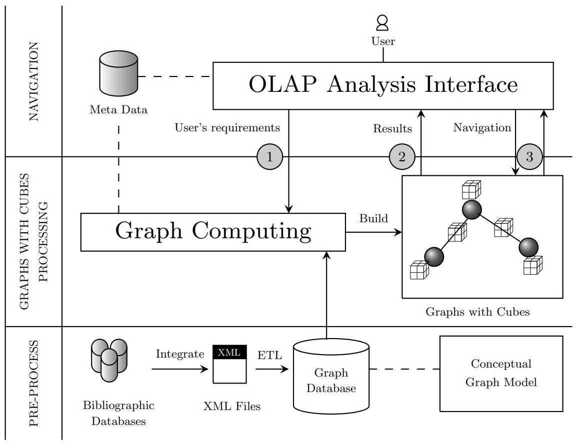


Fig. 1. The Architecture

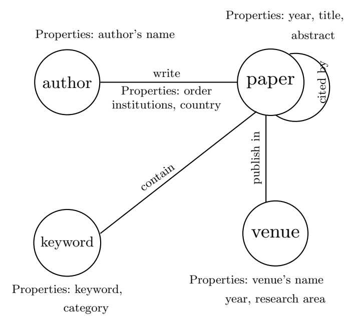


Fig. 2. The conceptual graph model

PRE-PROCESS. We first access bibliographic databases to extract data into XML files. After integration, ETL process is used to extract and load data into a graph database. In order to build an heterogeneous multidimensional network $G = (V, E, A_V, A_E)$ where V is the set of multiple nodes, E the set of multiple edge types, A_V and A_E respectively the set of attributes describing nodes and edges, we introduce a conceptual graph model (Figure 2). The conceptual graph model contains four types of nodes (author, paper, venue, keyword) and four types of edges among these nodes. Each node and edge are described by attributes. More details of nodes and edges are presented in Figure 2.

In reality, bibliographic data may have two problems. First, an entity concerns many different values in the same property. For example, author named Bin Yang works at Aalborg university and Fudan university in the same time. Secondly, a property value is changing over time such as a change of institution. For instance, Yzhou Sun published a paper in 2009 when he was at university of Illinois (Urbaba-Champaign), whereas his other publications were published for Northeasten university. In order to keep this information correctly, we design institution as an edge property between author and paper. It is useful to track changes over time.

GRAPHS with CUBES PROCESSING. A graph enriched by cubes may be used easily to perform OLAP operations on a network and it provides multiple network views at different levels of granularity. It takes a single graph rather than a set of graphs. With user's requirements, the first graph enriched by cubes is built. For example, the user chooses as fact the co-authorship. Co-authorship is a network where nodes are authors and an edge between two of them indicates they coauthored papers. Formally, we use the concept of path associated with co-authorship network. There are different paths of co-authorship from the conceptual model such as $author - write - paper$ or $author - write - paper - publish - venue$. Each path gives the different dimensions. For example, dimensions as keywords, the year and the venue of papers are taken from $author - write - paper - publish - venue$ path. Therefore, dimensions can be derived from node and edge attributes. Then, a first graph enriched by cubes like in Figure 3 (d) is built. The network is the co-authorship network enriched by a cube for each edge in order to count the number of papers written by two authors according to keywords, years and venues. It has no sense to build a cube for each node (author) because the fact is the co-authorship. While the fact is the scientific production, the network is the authors network and cubes for nodes and cubes for edges are created. It has a sense to count the scientific production of an author or between two authors. In order to view the constructed network from different perspectives, dimensions of cubes allow to perform multidimensional analysis over networks. For enriched graph computing, we propose a new algorithm (see section 3.2 for more details). Finally, graphs with cubes are sent as the result to OLAP analysis interface.

NAVIGATION. The OLAP interface manages both the user's needs and interactions, the input and the output of graphs with cubes during analysis. The OLAP interface uses meta-data in order to know the relationships between facts, measures, dimensions, nodes, edges, etc. It helps users to specify the first enriched graph to start OLAP analysis. Then the interface allows users to explore graphs and cubes from different views with OLAP operations.

In the next two subsections, we give more details about the algorithm for computing graphs enriched by cubes and about the OLAP operations.

3.2 How to Build Graphs Enriched by Cubes?

The graph enriched by cubes construction involves two algorithms: BUILDGRAPH for computing the aggregated graph and BUILDCUBES for constructing cubes on nodes or edges.

BUILDGRAPH (Algorithm 1) starts with the user's requirements with a fact F, a measure M, a set of dimensions D. We first generate a set of paths, P, which depends on fact, measure and dimensions (line 1) from G. Subsequently, the algorithm creates the structure of nodes V'_f, which contains node names according to the fact and a list of paths id. Then, we traverse the set of paths. For each path, we create a new node V'_f, if there is no such value (line 4-5). Otherwise, we simply update a path id for the node V'_f (line 7). After the loop, the algorithm creates the structure of edges E'_f. Each edge contains edge name coming from two any nodes and a list of measure's values. For each v'_f in V'_f, we compare the list of measure's values with the adjacent v'_f by using intersection operator. If the comparison result is not empty, we create a new edge $e'_f(v'_f, v'_f + 1)$.

After the creation of the aggregated graph $G' = (V', E', V'_P)$ where V' is a set of generalized nodes, E' is a set of generalized edges and V'_P is a set of paths associating with nodes, cubes are computed by BUILDCUBES. Due to the limited of space, the idea of BUILDCUBES is that if the fact needs cubes on nodes, the algorithm scans though V'_f. Otherwise, it scans though E'_f. The measure's value is computed from each path. Its value puts into cell that belongs to its dimensions. Each E'_f, we can find a list of measure's values. We don't keep a set of paths id in edges because more than one path have the same measure's value. For example, *paper*44 is got from the path 1 and the path 3 because it is written by two authors.

Figure 3 illustrates both algorithms. BUILDGRAPH takes as input the user's parameters. As previously, the fact is the co-authorship, the measure is the number of papers, the dimensions are the year, the venue and the keywords. In order to obtain the first graph, a set of paths is generated like $author - write - paper - publish - venue$. In our example, there are 13 paths (Figure 3a). The next step is to compute a set of nodes. We get a list of authors with their paths (Figure 3b). Then, any two authors who wrote papers together, are added to a list of edges (Figure 3c). The number of papers on edges is computed by using intersection operators. For instance, J. Han published paper33, paper47, paper44 and etc. Y. Sun published paper44, paper10, paper47 and etc. A set of papers between them is computed by $\{paper33,\ paper47,\ paper44, ...\} \cap \{paper44,\ paper10,\ paper47, ...\} = \{paper44\ ,\ paper47, ...$. Due to needing edges cubes of co-authorship network, the output graph of co-authorship network is built by selecting a set of nodes from edges like in Figure 3d. Authors who only write papers alone are not in the network.

3.3 OLAP Operations

Operations on cubes (e.g., roll-up, drill-down, slice, etc.) are supported to explore different multidimensional views and allow interactive queries and analysis. As we said before, two different types of operations are introduced in Graph OLAP [5]. The first one is an informational OLAP, and it uses informational attributes. This operation doesn't change the structure of the network. For example, venue and time are two informational attributes with their respective hierarchies $\{year, decade, all\}$ and $\{conference, area, all\}$. The second one, topological OLAP, implies a new structure of the network; if we do a topological Roll-Up, the network is generalized by some merging nodes. This operation uses topological attributes. In the authors network, for instance, the hierarchy $\{institution, country, all\}$ associated with the node attribute author can be used for merging authors from a same institution into a generalized node. In graphs enriched by cubes, we can perform both informational and topological OLAP. Informational OLAP operations are classically done, so we don't give details in the present. The most difficult problem we have to solve is how to support

Algorithm 1 BUILDGRAPH

Input: An heterogeneous multidimensional network $G = (V, E, A_V, A_E)$, a fact F, a measure M, a set of dimensions D
Output: A graph $G' = (V', E', V'_P)$.

1: Generate a set of paths (P) according to F, M and D
2: V'(value of node fact (V_f), a set of paths id $(pids)$) = null
3: **for** each $p \in P$ **do**
4:    **if** $p(V_f)$ not in Node(V_f) **then**
5:       Create $V'(V_f, p)$
6:    **else**
7:       add p at Node(V_f)
8:    **end if**
9: **end for**
10: E'(values of edge fact E_f, $a\ list\ of\ measure's\ value$) = null
11: **for** each $i = 0$ to V'.size **do**
12:    **for** each $j = i + 1$ to V'.size-1 **do**
13:       list1 = get the values of measure from node[i].pids
14:       list2 = get the values of measure from node[j].pids
15:       **if** list1 $\cap$ list2 != null **then**
16:          Create $E'(V_i\ and\ V_j,\ list1\ \cap\ list2)$
17:       **end if**
18:    **end for**
19: **end for**
20: Build graph according to fact.

topological OLAP operations over networks. This problem is even more difficult if we take into account the slowly changing dimension over time. A higher level of network cannot be computed from lower level of network without accessing raw data. Networked data is often non-summarizable. The idea of keeping a set of paths into nodes in the previous algorithms allows us to solve this problem.

Topological roll − up. Figure 4b shows an example of a topological roll-up of the co-authorship network to the institutions network. While all authors of a same institution are merged as one node, edges are created when any two institutions published papers together. In case of many institutions of an author in the same time, the author is counted into all his institutions. After the roll-up, in the more generalized network, new cubes have to be computed. In our example, co-authorship network involves edge cubes, whereas institutions network needs both cubes on nodes and edges. To build the institutions network, we use both BUILDGRAPH and BUILDCUBES. Before computing a set of nodes (line 2 in algorithm 1), we need to filter paths instead of generating a set of paths (line 1 in algorithm 1). We have to filter paths because all nodes of data set are collected in V', but some nodes may not be in co-authorship network (because some papers are written by only one author). The step of path filtering is called when the previous network needs cubes on edges. Then we compute a new set of nodes from line 2 in algorithm 1. Refer to example in Figure 4b, nodes are

pid	author-write-paper-publish-venue
1	P. Zhao-write-paper44-publish-KDD
2	P.Zhao-write-paper47-publish-CIKM
3	Y. Sun-write-paper44-publish-KDD
4	Y. Sun-write-paper10-publish-KDD
5	Y. Sun-write-paper47-publish-CIKM
6	J. Han-write-paper33-publish-KDD
7	J. Han-write-paper47-publish-CIKM
8	J. Han-write-paper44-publish-KDD
9	J. Han-write-paper33-publish-KDD
...	...

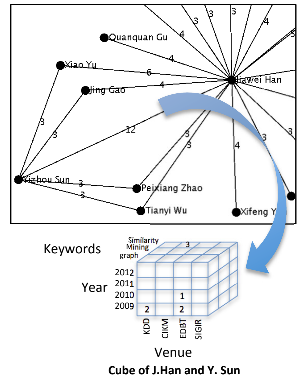
(a) Set of paths

P. Zhao	{1,2,...}
Y. Sun	{3,4,5,...}
J. Han	{6,7,8,...}
...	...

(b) Set of nodes with list of path id

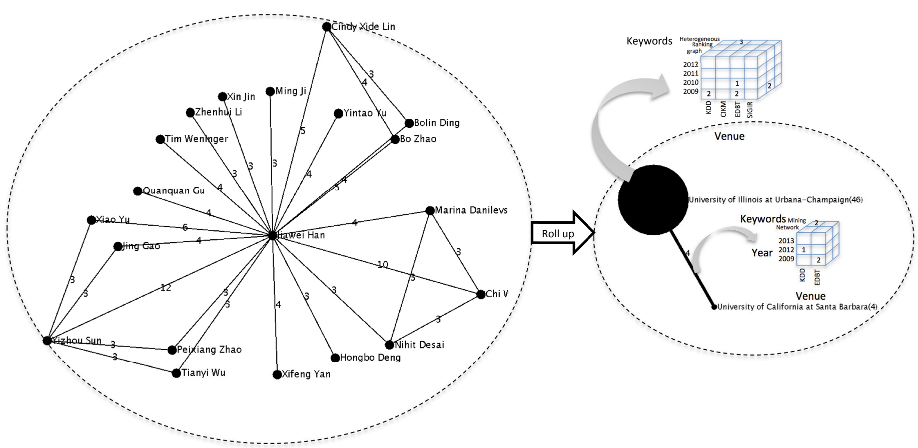
P. Zhao, Y. Sun	{paper44, ...}
P. Zhao, J. Han	{paper47, ...}
Y. Sun, J. Han	{paper44, paper47, ...}
...	...

(c) Set of edges with list of measure



(d) Graphs enriched by cubes

Fig. 3. Computation of a co-authorship network



(a) Co-authorship network (number of papers > 2)

(b) Institution network (number of papers > 2)

Fig. 4. Roll up from the co-authorship network to the institution network

grouped into institutions. For example, university of Illinois contains path6 and path7 because J. Han and P.S. Yu belong to this university.

Slice. Traditional slice operation selects one particular dimension from a given cube and provides a new sub-cube. In our context, slice operation can not be like the classical one, it should be adapted to graphs. The slice operation selects a part of the graph and provides a new sub-graph. For example, if a whole co-authorship network is too big to be comprehensive, the user can focus on a smaller subgraph more interesting to analyze information clearly.

4 Experiments

In this section, we present how we have implemented our solution and we give some examples of analysis.

4.1 Implementation

The implementation has been done as follows:

- We get data from three bibliographic databases. First we use the well known database DBLP. But in order to complete our conceptual graph model, we access on ACM and Microsoft Research databases for taking keywords, institutions and research areas. In theses three sources, we keep only three research areas (data mining, databases and information retrieval) and we pick only a few representative conferences for the three areas (PODS, EDBT, KDD, DOLAP, ASONAM, SIGIR and CIKM). At the end, we build a data set which contains 4,727 papers and 8,238 authors since 2009.
- The ETL process is used to fulfill data into the model. After cleaning, data is loaded into an unified structure with a graph model.
- A new type of NoSQL databases, called graph databases, is used to implement our conceptual graph model. We choose Neo4j[1] because it is an open-source software, it supports the properties of our graph model.
- Finally, an OLAP interface analysis is developed in Java and tested on a Mac OS X version 10.9.2 with Intel core i5 2.4 GHz and 8 GB of Ram. For graph visualization, we use the GraphStream[2] library because it is a library to model and analyze the dynamic of graphs and it is an open source library.

4.2 Example of Analysis

We first use the example of the co-authorship network introduced before Figure 5 shows the co-authorship network in three areas since 2009. Each edge of this network has a cube. In order to reduce the graph, we filter edges in order to keep only edges with a number of papers over than 10.

For example, look at the edge between Iadh Ounis and Craig Macdonald; these authors published 29 papers together. We consider these papers like a cube with two dimensions. It could be interesting to have two ways of visualization. The first way is to focus on time, having the count of papers per year. Each year has the count of papers by venues. The second way is to focus on the venue, having the count of papers per venue's name. Each venue has the count of papers by year.

Now, if we want to roll up the co-authorship network of the conference KDD between 2009 and 2013 to the institutions network (for the same conference and years) with a topological OLAP operation, we obtain Figure 6. The institutions are filtered with a number of papers over than 10. For example, the university of Illinois at Urbana - Champaign published 31 papers in the KDD conference from

[1] http://neo4j.com/

[2] http://graphstream-project.org

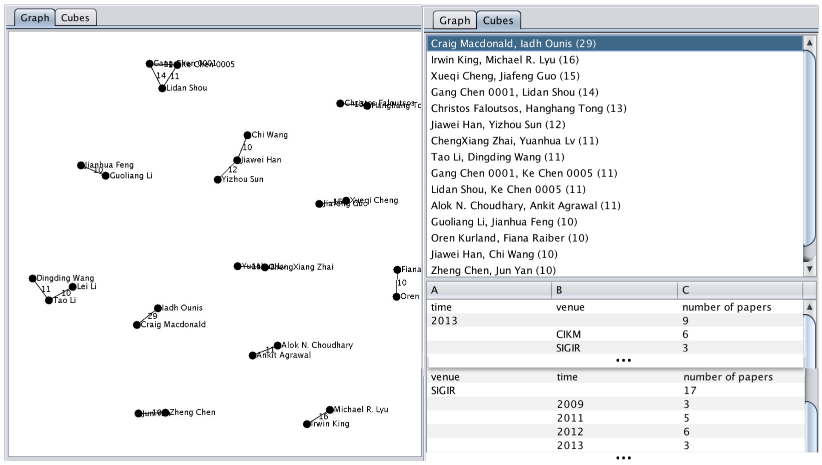


Fig. 5. The co-authorship network (on three areas and all years) with a number of papers over than 10

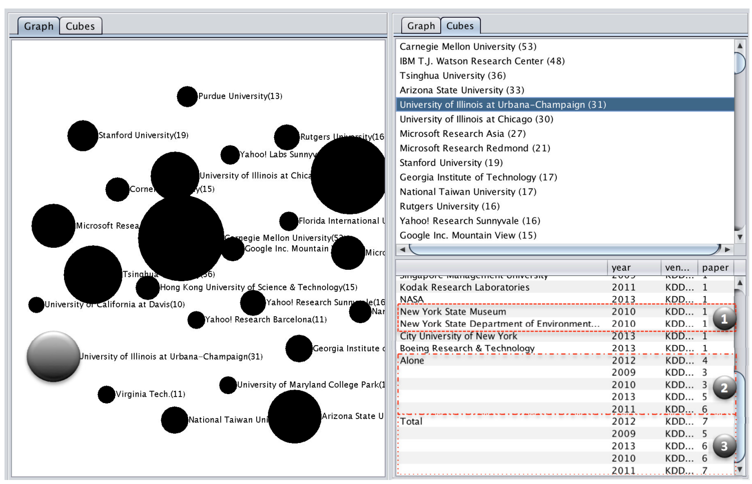


Fig. 6. The institution network for the KDD conference (2009-2013) with a number of papers over than 10

2009 to 2013. Look at the big number 1 in Figure 6, it means that the university of Illinois at Urbana - Champaign has one collaboration with the New York State Museum and one with the New York State department of Environment in 2010 by publishing in the KDD conference. The big number 2 shows the number of papers written by several authors but all belonging to the same university (Illinois at Urbana - Champaign).
Furthermore, in the interface, the user can slice to consider only a sub-graph.

There are several groups of authors in co-authorship network. Suppose that the user needs to consider only the interest group; with a slice operation, the user can select the sub co-authorship network. Finally, a roll-up operation is done on this sub-graph.

5 Conclusion

In this paper, we wanted to enhance decision support on networks by combining OLAP on networked data. So we first presented a conceptual graph model to support OLAP on bibliographic networks. In order to consider both nodes, edges and multidimensional information for the analysis, we proposed the graphs enriched by cubes. The graphs enriched by cubes perform multidimensional views of a heterogeneous graph rather than a set of graphs. Cubes are provided for nodes or edges according to the user's requirements (fact, measure, dimensions, etc.) In order to compute graphs enriched by cubes, we proposed the algorithms which, in addition, solve the slowly changing dimension problem in OLAP analysis. Then we adapted the OLAP operations to graphs enriched by cubes. We showed an implementation with real data sets from three bibliographic databases. We focus our approach on bibliographic data in this paper, and it can be applied to other use cases as well.

References

1. Waqas, A., Zimányi, E., Wrembel, R.: Temporal Data Warehouses: Logical Models and Querying. In: EDA'15, vol. RNTI-B-11, pp. 33–48 (2015)
2. Ferrara, A., Salini, S.: Ten challenges in modeling bibliographic data for bibliometric analysis. In: Scientometrics, vol. 3, pp. 765–785 (2012)
3. Mallig, N.: A relational database for bibliographic. Journal of Informatics, pp. 564-580 (2010)
4. Trifonova, T. G.: Warehousing and OLAP Analysis of Bibliographic Data. Intelligent Information Management, vol. 3, pp. 109–197 (2011)
5. Chen, C., Yan, X., Zhu, F., Han, J., Yu, P.S.: Graph OLAP: Towards online analytical processing on graphs. In: ICDM'08, pp. 103–112 (2008)
6. Qu, Q., Zhu, F., Yan, X., Han, J., Yu, P.S., Li, H.: Efficient Topological OLAP on Information Networks. In: DASFAA'11, Part I. LNCS, vol. 6587, pp. 389–403 (2011)
7. Tian, Y., Hankins, R.A., Patel, L.M.: Efficient Aggregation for Graph Summarization. SIGMOD Conference, pp. 567–580 (2008)
8. Zhao, P., Li, X., Xin, D., Han, J.: Graph cube: on warehousing and OLAP multidimensional networks. In: SIGMOD'11, pp. 853–864 (2011)
9. Loudcher, S., Jakawat, W., Morales, E.P.S., Favre, C.: Combining OLAP and information networks for bibliographic data analysis: a survey. Scientometrics, vol. 103(2), pp. 471-487 (2015)
10. Yin, M., Wu, B., Aeng, Z.: HMGraph OLAP: a Novel Framework for Multidimensional Heterogeneous Network Analysis. In: DOLAP'12, pp. 137–144 (2012)
11. Zhang, J., Hong, X., Peng, Z., Li, Q.: Nestedcube: Towards Online Analytical Processing on Information-Enhanced Multidimensional Network. WAIM Workshops, pp. 128–139 (2012)

12. Beheshti, S.M.R., Benatallah, B., Motahari-Nezhad, H.R.: A Framework and a Language for On-Line Analytical Processing on Graphs. In: WISE'12, pp. 213-227 (2012)
13. Wang, Z., Fan, Q., Wang, H., Tan, K.L., Tan, Agrawal, D., El Abbadi, A.: Pagrol: Parallel graph olap over large-scale attributed graphs. ICDE'14, pp. 496-507 (2014)
14. Kaya, M., Alhajj, R.: Development of multidimensional academic information networks with a novel data cube based modeling methord. Information Sciences, pp. 211–224 (2014)
15. Jakawat, W., Favre, C., Loudcher, S.: OLAP on Information Networks: A New Framework for Dealing with Bibliographic Data. SoBI'13 in conjunction with ADBIS, pp. 361-370 (2013)