Keyword Map Visualisation

Matej KLOSKA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
matej.kloska@gmail.com

Nowadays, people are creating more digital documents – data elements – than ever before. If we want to search and navigate in those documents, we need efficient way how to interpret relations between documents. Information visualisation has become a large field and various "subfields" are beginning to emerge. The question is, whether there is an inherent relation among the data elements to be visualised. Proper relations creation and information visualisation implies high success rate in looking for desired information in documents.

If we work with a large number of data elements, we often need efficient representation of content – e.g., to describe each document with an appropriate set of keywords. Keyword sets alone do not guarantee quality of maps, i.e., the ease of navigation in information space. Quality in this sense strongly depends on interconnections of keywords between documents. There are several techniques how to create keyword connections. Most of them are based on clustering techniques.

Another problem is, how to properly visualise keywords and relations between them. The importance to support users and provide adequate user experience is also crucial problem, which we have to keep in mind. The more fail proof interface, the more valuable maps.

We propose a method for keyword map visualisation for educational system ALEF with support of content managing system COME^{2}T (Collaboration- and Metadata-oriented Content Management EnvironmenT; see Figure 1).Our method will be implemented and evaluated using existing system COME^{2}T [1]. The COME^{2}T allows easy administration of lightweight semantics for the provided content – digital documents and user-created annotations. It was designed as is being used a content management system for educational system ALEF.

Graph structure highly affects visual quality of output map. Basically, there are two key issues targeted to a structure. First of them is *graph size*. The higher the count of vertices in map, the more difficult to work with such graph. To solve this issue, we introduce structural clustering in order to virtually reduce number of visible vertices on

* Supervisor: Marián Šimko, Institute of Informatics and Software Engineering

Spring 2014 PeWe Workshop, March 21, 2014, pp. 99–100.

screen. Structural clustering is more suitable in comparison to traditional semantics clustering that is based on clustering of keywords' semantics.

Second issue related to graph structure is *graph density*. Dense graph is a graph with the number of edges close to the maximum number of edges. On the other hand, sparse graph is the opposite – low number of edges. The higher the number of graph edges the lower the readability for the user.

We have identified three features that would improve user experience and make navigation easier - keyword map overview, navigation bar and coloring of nodes and edges

The keyword map overview will help user in every moment know where in map (s)he exactly is. It would be particularly useful in the case of zoom and pan where only a small part of map is shown on screen.

Currently, there is no other simple way how to provide user simple feedback on position in map in COME^{2}T, especially when clustering is applied. Navigation bar would enhance track of navigation in map when user wade in any cluster. It is similar to well-known address bar in web browsers or file explorers. In our case, bar provides information about users' map exploration path with respect to clusters hierarchy.

The third proposed feature is colouring of map elements. Visual feedback is key feature for every user. We would like to provide interface for custom colouring of nodes and edges in addition to automatic one based on defined rules and analysis of graph. Automatic colouring is easy to implement and use because of predefined edges types.

Since we are still in a phase of implementing the proposed utilities to the system COME^{2}T, we here describe the evaluation plan. Our method would be evaluated in several phases. In first phase we will evaluate the impact of user interface and visualization library changes on time required for creation defined maps for provided courses. We expect, that time required for creation will be smaller than for the early method.In second we will evaluate the impact of supporting features like navigation bar and keyword map overview when a user comes back to previously created maps. We expect that supporting features will help a user orientate in map easier and recognize work already done. Final evaluation of method will be evaluated using the same tasks as in evaluation made before implementation of our method.

Extended version was published in Proc. of the 10th Student Research Conference in Informatics and Information Technologies (IIT.SRC 2014), STU Bratislava, 9-14.

Acknowledgement. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/0675/11.

References

[1] Šimko, M., Franta, M., Habdák, M., & Vrablecová, P. (2013). Managing content, metadata and user-created annotations in web-based applications. Proc. of the 2013 ACM Symposium on Document Engineering – DocEng '13, 201. doi:10.1145/2494266.2494270.