Visualization Method of Patent Citation Analysis Based on Improved Hyperbolic Tree

Sheng-Nan ZHANG¹,a and Xiao-Cui MA¹,b,*

¹School of Information Science and Engineering, Shenyang University of Technology, Shenyang, 110023, China
azsncjr@sina.com, b365849859@qq.com
*Corresponding author

Keywords: Visualization, Hyperbolic tree, Citation analysis.

Abstract. Citation analysis is an important content of patent analysis. Aiming at the problems of visualizing the large hierarchical data in the limited display area, an improved hyperbolic tree for the visualization of patent citation analysis is proposed. The method combines the traditional hyperbolic tree with fisheye and other visualization technologies, and it can reveal the multi-level patent information, as well as display the overall layout and local details. Thus, the interaction between the overall and local data may be realized easily, and it is convenient for the users to understand the interested information.

Introduction

The patent citation analysis is an analysis of the reference to the patent and the cited case [1], which plays a very important role in all of the patent analysis. By studying the patent citation relationship and tracking the patent citation information, we can find the core patents, technology trends, and potential competitors in the specific field [2-3]. Because of the great amount of patent information data, the efficient and fast patent citation analysis must use professional and visual analysis tools.

At present, there are many commercial patent analysis tools, but most of them are just the primary level tools for patent information statistics and analysis, and the ones for high level analysis and visualization, such as patent citation analysis are not still perfect [4]. The existing patent citation analysis tools usually use hierarchical tree or hyperbolic tree to visualize the information, and hierarchical tree is a node-link and space-filling visualization technology, while Hyperbolic tree is a tree structure in the hyperbolic plane and it uses an optimization technique called focus + details [5]. With the increasing of patent data, the traditional visualization technology of patent citation analysis is becoming more and more challenging, for example, hyperbolic tree is difficult to reflect the detailed information and its readability is poor in analyzing massive amounts of patents. Although hyperbolic tree improves the space utilization and hides the non-focus information nodes so as to highlight the focus nodes, its ability to reveal multi-level patent information is limited, furthermore, with the increase in the level of citation, the patent data are increased exponentially, there inevitably exists the occlusion in the dense region, and the relevant patent details can’t be fully demonstrated [7-8].

In view of the existing problems, this paper proposes an improved hyperbolic tree method for patent citation analysis. The method combines the traditional hyperbolic tree with fisheye and other visualization technologies, and it can reveal the multi-level patent information, as well as display the overall layout and local details. Thus, the interaction between the overall and local data may be realized easily, and it is convenient for the users to understand the interested information.
Relevant Technology

Hyperbolic Tree

Hyperbolic tree is based on the hyperbolic circular plane to display the hierarchical structure information; it not only can display the whole hierarchy structure, but also has more space to show the current focus parts. Hyperbolic tree technology is mainly composed of three parts: layout, mapping and dynamic transfer.

According to the recursive operation, layout places the root node of the tree in the center of hyperbolic space. Then, the fan region around the root node is divided averagely to the sub nodes according to the above recursive method. By that analogy, we may ensure that each node does not overlap. Mapping uses Pang card (Poincare) projection to map the hyperbolic plane onto Euclidean plane. Dynamic transfer refers to human-computer interaction and information navigation, the effect is shown in Fig. 1.

Fisheye Technology

Fisheye technology is designed based on the characteristics of human vision, while enlarging the focus node, it can reduce the surrounding nodes, and so we can not only highlight the key information, but also reveal the context. The important parameter of Fisheye is DOI, which refers to the important degree of each node that is relative to the focus position. According to the calculated DOI, the visualization information can be controlled.

Visualization Method Based on Improved Hyperbolic Tree

The Basic Idea of Visualization

The basic idea is to combine the traditional hyperbolic tree with fisheye technology. The root node is placed in the center of the whole display area, and the other levels of nodes are displayed by calculating the importance of each node relative to the position of the root node so as to gradually reduce the space of different levels of nodes along the radial direction. Thus, we can not only solve the problem of revealing multi-level patent information, but also enlarge the aggregation nodes and reduce peripheral nodes simultaneously, which will lead to have more space to display the current focus area and greatly improve the utilization rate of space. The improved method also provides a series of interactive means to solve the problems of node occlusion, detail information display, and so on, and it combines hyperbolic tree and line chart, histogram, pie chart to realize the interaction between the whole and the part.

Interactive Process

In order to make a better visual analysis, interactive technology is essential. According to the needs of the hyperbolic tree, six ways of interaction are provided.

Focus shift: The user can move the concerned node to the center of the circular display area, and this node will be allocated the largest area. With the deepening of the citation level, the allocated area is
becoming smaller and smaller. In addition, according to the properties of nodes, the same category of nodes are marked with the same color.

View transformation: The user can choose a simple tree structure to hide the detail information of all nodes. Projection parameter type and the distance between nodes of hyperbolic tree may be modified dynamically so as to rearrange hyperbolic tree re.

Viewpoint control: In the traditional hyperbolic tree, all nodes are in the periphery of the citation tree, which not only leads to the low utilization rate of space, but also leads to the occurrence of occlusion between nodes. In order to solve the above problems, by using the viewpoint control, nodes can be switched between the thumbnail display and the full display.

Details-on-demand: When inputting the character in the text box of the upper right of the screen, hyperbolic tree nodes and matching characters will be highlighted synchronously so as to trace the citation context.

Brushing: The specific details of the node, such as patent name, application date, inventor, and so on, are displayed when moving the mouse over a node.

Overview-detail: There are two windows in the improved hyperbolic tree, namely, overview window and detail window. In the detail window, a line chart, histogram, pie chart are added so as to use them to display the relevant competition, technology and field analysis respectively when clicking the node.

Algorithm Implementation

The algorithm consists of two main steps: the node layout in hyperbolic space and the mapping from hyperbolic space to Euclidean space. Specific steps are as follows:

Step1: Set the root node of the tree to the origin of the hyperbolic plane (0, 0).

Step2: Fan area of the root node is averagely divided into its sub nodes, then the sector of each sub node is also equally divided to the lower sub nodes, in turn, until all nodes are distributed the sector. Let the fan area of the node $m$ be $(p_k, q_k)$, then the fan section of the $k$ sub node is calculated as follows.

$$ (p_k, q_k) = (p_k + k \times q_k, q_k / n) $$

Where, $p$ is the starting angle of the node sector, $q$ is the size of the sector, $n$ is the total number of sub nodes.

Step3: Use Poincare projection to map the points and lines of the hyperbolic space onto the Poincare disc of the European space. Since the hyperbolic tree is constructed in the circular area, the polar coordinate system is more favorable for numerical calculation, the plane coordinates are transformed into polar coordinates. Assuming that each node is represented by a complex number $z$. Formula is as follows. Coordinate transformation with changes of $\rho$ axis length change and the angle to represent the changes of $z$. $\overline{\rho}$ is the complex conjugate of $\rho$.

$$ Z_i = \frac{\theta_1 + \rho}{1 + \rho_i} $$

where, $\rho = \frac{\theta_1 \rho_1 + \rho_2}{\theta_1 \rho_1 + 1}, \rho = \frac{\theta_1 \rho_1 + \overline{\theta_1 \rho_2}}{\theta_1 \rho_1 + 1}, \rho_1, \rho_2$ are the axis lengths of two arbitrary nodes in curve, $\theta_1, \theta_2$ are the corresponding angles, $\rho$ is the length of the pole axial, $\overline{\rho}$ is the complex conjugate of $\rho$, $\theta$ is the angle of the pole.

Step4: Calculate the DOI, the formula is as follows.

$$ DOI = (x \mid f_p = y) - API(x) - D(x, y) $$
Where, API is a priori importance of node $x$, which is assigned to the node in order to show its relative importance in the overall framework structure, $D(x, y)$ is used to calculate the distance between the node $X$ and the current focus $Y$, finally DOI is described and displayed.

**Application Examples**

A visualization system of patent citation analysis is designed based on the improved method. The functions of this system mainly includes: checking the citation relationship between patents, displaying the details of the nodes, realizing the overall and local interaction, changing the hyperbolic tree projection parameters, adjusting the hyperbolic tree structure, and highlighting the concerned parts.

This paper is based on the United States patents and aims at the new energy vehicles, the patents and their citation data from 2014 to 2004 are retrieved.

Check the cited cases: By searching for a patent, we can view the cited cases of the patent as the root node of the hyperbolic tree, and can adjust hyperbolic tree and call functions according to the demand, the effect is shown in Fig. 2.

![Figure 2. Interactive expression of hyperbolic tree.](image)

Find competitors: It can be found that a large number of patents in a company's patent is cited by another company, which is a large number of follow-up studies on the company's patent. Can see a number of companies competing for a very intense, the effect is shown in Fig. 3.

Check the company's patent layout: After the discovery of a patent, the company has carried out a large number of patent applications around the patent. Through the broken line chart can quickly find the company's future trend of the company in the next few years, so that companies can increase their competitiveness, the formation of core technology group, the effect is shown in Fig. 3.

Find core patent: It can be found that the core of the field, which have high reference rate, and then can understand the technical core of the field. Through the analysis of the hyperbolic tree in the sub areas, we can quickly find the core patents in the field, the effect is shown in Fig. 3.
Figure 3. Interactive expression of hyperbolic tree and other visualization tools.

Conclusions

Through combining the hyperbolic tree and other visual technologies, this paper presents an improved hyperbolic tree for the visualization of patent citation analysis, and it not only solves the traditional visual problems, but also displays the search results more clearly and abundantly. The method may be used in the visualization of other large hierarchical data.

Acknowledgement

This research was financially supported by general project of scientific research in colleges and universities of Liaoning Province (L2014049).

Liaoning Provincial Department of education project (W2015311).

References


