

基于双曲树的专利权利要求可视化分析

刘玉琴

(北京印刷学院绿色印刷包装产业技术研究院, 北京 102600)

摘要: 观察专利权利要求的语言表述, 总结其结构特征, 应用正则表达式进行结构解析。进而, 采用双曲树对转换成树形结构的专利权利要求进行可视化呈现, 方便并增强对专利权利要求结构和内容的理解。构建了专利权利要求解析与可视化工具, 并以苹果公司的图形用户接口专利7166791为例进行实证研究。

关键词: 权利要求; 双曲树; 可视化

中图分类号: TP391

DOI: 10.3772/j.issn.1673—2286.2015.01.009

1 引言

权利要求作为专利的必要组成部分, 划定了专利的技术保护范围, 具有显著的法律属性, 成为各种专利纠纷关注的焦点。因此, 对于专利权利要求的分析和理解在专利信息分析和专利纠纷诉讼中具有重要意义。为准确、高效的进行专利权利要求的理解, 本文构建专利权利要求结构解析和可视化工具, 以世界知识产权组织检索系统为数据源进行数据的获取; 然后, 观察专利权利要求的语言表述, 总结其结构特征, 应用正则表达式进行结构解析; 最后, 将解析后的专利权利要求采用双曲树进行可视化呈现, 以此来提高专利权利要求分析利用的效率。

2 研究背景和方法

2.1 研究背景

一般情况下, 专利权利要求由多个子权利要求构成, 这些子权利要求细分为独立权利要求和从属权利要求。独立权利要求和从属权利要求在语言描述和内

容构成方面不同。同时, 独立权利要求和从属权利要求又存在密切的联系: 从属权利要求是对独立权利要求的进一步限定, 在语言表述上需要引用之前的某一个独立权利要求, 同时也可以作为之后的某一个或几个从属权利要求的技术基础。在语言表述上, 从属权利要求简明扼要, 不再重复之前独立权利要求或从属权利要求的内容。因此, 在专利权利要求数量较多, 从属关系复杂的情况, 增加了对权利要求结构和内容进行阅读和理解的复杂度。

为了提高对专利权利要求的理解和利用效率, 相关的软件工具被设计开发出来。这类软件工具可归纳为两类: 一类是供专利代理人使用的权利要求撰写辅助工具, 如Patent Architect^[1]和Claim Master^[2], 可以检测代理人撰写的权利要求之间的逻辑关系。这类工具只针对代理人撰写的专利, 在专利信息分析中的应用有限。另一类是能够进行权利要求结构解析的工具, 如Patent Tools^[3]和Patent Claims Tree^[4], 提供了权利要求结构解析功能, 并将独立权利要求和从属权利要求以层次树结构进行展示。这两个工具在权利要求数量较多时并不能很清晰的展示权利要求的结构特征, 并且主要用于美国专利的权利要求结构解析。

基于以上分析, 构建一个能够快速的获取各国专利权利要求数据、识别权利要求结构特征、辅助用户理解权利要求内容的软件工具是十分必要的。

2.2 研究方法

基于双曲几何的双曲树算法是操纵大型层次结构数据且在科技信息分析领域应用广泛的可视化方法之一。该方法与用户交互技术融合后, 可以在有限的计算机屏幕空间内清晰的显示和控制层次结构数据。如汤森路透Web of Science^[5]和Thomson Innovation^[6]分别应用该技术进行论文和专利的引证关系可视化表示, 分别提供了论文、专利间一级、二级的前向引证(参考文献或参考专利)、后向引证(引证文献或引证专利)、双向引证关系图, 图中每个节点表示一个论文或专利, 并用不同颜色按照节点文献或专利所属作者、机构、期刊、学科等对节点加以区分, 切换节点文本显示内容。国内学者靖培栋应用双曲树对中图分类号体系进行可视化表示, 设计了相应的软件工具^[7]。

同样, 专利的单个独立权利要求与其从属权利要求构成一个树形结构, 多个独立权利要求与其各自的从属权利要求构成若干树形结构集合。这些树形结构集合以待解析专利为根结点, 转换成一个新的、独立的树形结构, 转换后的、新的树形结构往往具有较多的数据节点, 与双曲树可视化具有较好的拟合性。为此, 本文构建的权利要求结构解析与可视化工具, 首先发现总结从属权利要求与其隶属的独立权利要求之间的语言描述特征, 进而, 采用正则表达式进行权利要求结构的解析; 然后, 基于双曲树算法把权利要求之间的关系以树形结构形象地展示出来。

3 系统介绍

3.1 系统架构

权利要求结构解析与可视化工具数据来源为世界知识产权组织的官方检索系统^[8]。首先, 根据待解析专利的专利申请号或授权号, 由权利要求数据搜索引擎获取包括待解析专利权利要求信息的HTML数据; 然后, 数据清洗引擎应用正则表达式提取出权利要求文本, 进一步根据解析规则将权利要求内容转换为树形结构并存储; 最后, 采用双曲树进行结果展示。整个系统的

结构框架和应用流程如图1所示。

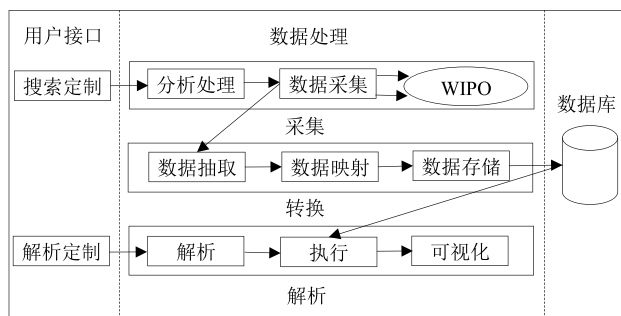


图1 权利要求结构解析与可视化工具框架图

3.2 系统界面

权利要求结构解析与可视化工具的界面设计如图2所示。整个界面分为菜单栏、工具栏和可视化三部分。菜单栏主要包括: 文件、帮助两个菜单项; 工具栏主要包括: 打开按钮、保存按钮、专利号输入文本框、搜索按钮、网络连接按钮、彩色背景按钮、黑白背景按钮、图形样式选择按钮、根节点按钮、截图按钮、报表输出按钮、主页链接按钮、帮助按钮、退出按钮等十四项; 界面下方为解析后的权利要求结构可视化展示区域。

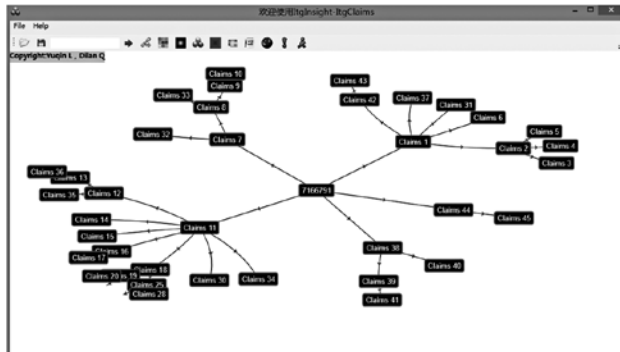


图2 权利要求结构解析与可视化工具主页面

3.3 系统功能

权利要求结构解析与可视化工具的功能如下:

(1) 权利要求数据搜索和清洗。在工具栏的专利号输入文本框中输入待解析专利申请号或授权号, 点击搜索按钮, 系统后台线程与世界知识产权组织专利检索系统进行数据交互, 采集包括权利要求信息的HTML文本。同时, 提取HTML文本中权利要求数据。

(2) 权利要求结构解析和可视化展示。搜索到的

专利权利要求经过后台的解析线程,形成树形结构,然后以双曲树的形式显示在可视化展示区。通过鼠标单击权利要求节点将其布置在屏幕中心。当鼠标滑过某个节点上方时,节点旁显示该节点所代表的权利要求文本。

(3) 查看专利著录项目。单击工具栏的网络连接按钮,浏览器自动打开世界知识产权组织专利数据库,并显示当前待解析专利的作者、机构、国际分类号、题目、申请日、授权日、公告日等著录项目。

(4) WORD输出权利要求文本。单击工具栏的报表输出按钮,系统调用OFFICE VBA将专利的权利要求文本输出到Word文档中。

(5) 解析结果的保存和重复利用。点击工具栏的截图按钮截取并保存可视化展示区域的当前图形,或通过类的序列化将图形的数据结构类保存到计算机硬盘,在后续使用时再通过类的反序列化操作恢复到可视化展示区。

(6) 可视化图形样式设置。应用工具栏中的彩色背景设置、黑白背景设置和图形样式设置,修改背景颜色,图形节点、连线和文字颜色。

4 关键技术实现

4.1 权利要求信息搜索与提取

首先,定义权利要求结构体,用于存储专利的单一权利要求。结构体包括权利要求的序号,权利要求的父权利要求序号和权利要求内容。

其次,选择世界知识产权组织免费开放的专利检索系统,该系统将各国专利的著录项目和权利要求文本翻译为英文,数据全面、便于解析。为此,以C#语言WEB请求类HttpRequest向服务器发送GET请求,与服务器进行信息交互。请求返回包括专利权利要求集合的HTML页面,应用正则表达式进行当前专利所有权利要求信息的提取和清洗。

再次,分割权利要求集中的单一权利要求,并记录权利要求的序号。

通过观察总结权利要求的语言表述,发现每个权利要求作为前面权利要求的从属权利要求时,都会出现明确的引用序号。比如:一个方法专利的独立权利要求的典型语言描述为:A method of……;从属权利要求的典型语言描述为:A method as recited in claim

X……或者A method according to claims X……或者The method of claims X……等。

发现这些特征后,编写正则表达式提取每一个单一权利要求的父权利要求序号,并存储在一个权利要求结构体中。如果权利要求为独立权利要求,设定其父权利要求序号为0。这样,每件被解析的专利由若干个权利要求结构体构成。提取的权利要求集合,存储在C#结构类TreeView中。

4.2 双曲树算法

传统树型结构在层次增多或者节点增多时,需占用大量空间,而计算机屏幕所能够提供的可视化空间有限,使得查找某个节点或者获得整个结构信息变得不方便。为此,美国Xerox Palo Alto研究中心提出了基于双曲几何的双曲树算法来解决这一问题^[9]。双曲树算法实现的主要原理是将树结构在双曲空间进行布局,然后映射到欧式空间的庞莱卡圆盘进行显示。

在技术实现上,定义复数类用以表示双曲空间中点的位置,复数的实部、虚部分别与双曲空间中点的横纵坐标对应,将双曲空间坐标转换到欧式空间的坐标,定义扇形类用于表示双曲空间中的扇形区域;双曲树节点类,实现节点在双曲空间的布局。

4.3 图形绘制

应用微软Windows Presentation Foundation (WPF) 框架进行可视化图形的绘制,分别定义节点绘制函数和连线绘制函数实现图形绘制。

首先,应用节点绘制函数绘制矩形框和文本。

其次,为了实现节点间平滑的曲线连接效果,应用C#语言自带的贝塞尔曲线构造类QuadraticBezierSegment进行节点连接绘制,并在绘制连线时附加箭头,以表示节点间权利要求的引证关系。

4.4 权利要求可视化

专利权利要求可视化的最终表现形式如图3所示。首先,将待解析专利的专利申请号或授权号作为根节点置于屏幕中心,距离其最近的一层节点为该专利的所有独立权利要求;对同一独立权利要求下的从属权利要求按照序号大小进行排序,形成第二层分支结构;依

次就同一从属权利要求下的子从属权利要求进行更深层级的排序,形成第三、四……层的分支结构。以带有箭头指向的连接线指示权利要求之间的引用关系。同时,以每个权利要求的内容作为节点的备注信息,当鼠标滑过节点时,显示备注信息。通过这种方式使用户可以迅速便捷的获取、理解专利独立权利要求与从属权利要求内容,相互之间的引用依赖关系。特别是在专利权利要求数量多,结构复杂的情况下,更有助于专利权利要求内容和结构的理解。

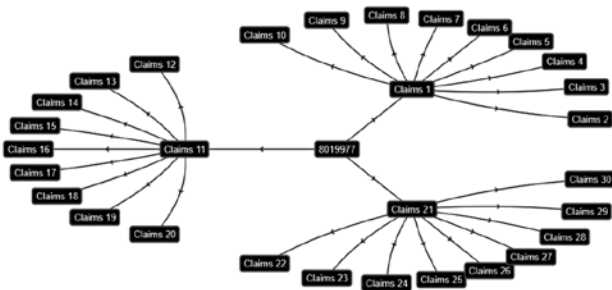


图3 权利要求可视化表现形式

5 实证应用

以美国苹果公司的授权专利7166791为例,应用本文构建的权利要求结构解析和可视化工具进行实证分析。该专利是一篇关于苹果公司操作娱乐设备的图形用户接口设计专利,拥有权利要求45项,被引用次数20次,在中、美、欧、日、韩等多个国家地区进行了专利申请,是图形用户接口技术中非常重要的专利之一。

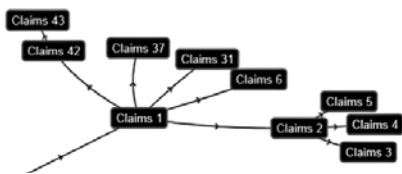


图4 专利7166791权利要求结构

首先,输入待解析专利授权号进行权利要求数据采集和清洗,并对结构进行解析和可视化呈现,如图4所示。发现该专利有5项独立权利要求,序号分别为1、7、11、38、44,从属权利要求之间的引用关系更是复杂,引用层级最多达到6级。

进一步,发现该专利权利要求的内容并不按照序号1-45的次序进行排列,而是按照引用关系相互交叉,形成若干的权利要求引用关系路径。如图5所示

的独立权利要求11下的从属权利要求,其引用路径包括: 36<-13<-12<-11; 14<-11; 15<-11; 17<-16<-11; 22<-21<-20<-19<-18<-11; 24<-23<-19<-18<-11; 29<-27<-26<-25<-18<-11; 28<-25<-18<-11; 30<-11; 34<-11等十个路径。



图5 专利7166791独立权利要求11的从属权利要求结构

对于处在某一引用路径上的任一权利要求,对其内容的阅读和理解,需要按照路径指示的权利要求依次进行。如图6中引用层级最多的从属权利要求29,其内容表述如下:

A hierarchically ordered graphical user interface as recited in claim 27, wherein when the selected item is at least one of the songs in the list of all songs of the selected album, further comprising: a fifth order interface having a song information list corresponding to the selected song that includes, a song title of the selected song, a song artist label of the selected song, and a song album title corresponding to the selected album.

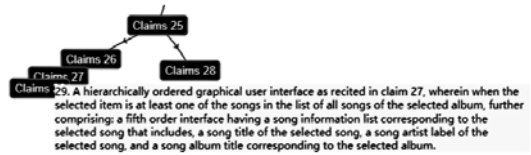


图6 专利7166791从属权利要求29内容表述

对于其内容的阅读和理解单纯从其表述或按照1-28的序号对在前权利要求进行阅读和理解,都无法准确的理解其内容。但是,通过对整个引用路径中的权利要求11、18、25、26、27依次进行阅读和理解,则可以比较清晰的知道其内容为:在与娱乐设备进行交互操作中,用户选择专辑中的某一歌曲时,设备向用户显示的内容包括歌曲名称、艺术家名称和专辑名称等信息。

通过以上方式的操作可以快速的对该专利的任一权利要求进行解读,从而提高专利权利要求理解和利用的效率。

6 结语

专利权利要求在专利信息分析和专利纠纷中具有重要作用,本文以世界知识产权组织专利检索系统为数据源,构建专利权利要求结构解析和可视化工具。在总结权利要求语言表述特征的基础上,编写正则表达式进行权利要求内容的提取和结构解析,应用双曲树可视化技术对权利要求进行呈现,以此提高对专利权利要求分析和利用的效率。在下一步的研究中,将论文设计的软件工具转换为网络版,在互联网上进行发布,向社会大众提供免费开放的服务,是后续将要开展的工作。

参考文献

- [1] Patent Architect[EB/OL].[2013-05-15].<http://www.patentarchitect.com/UsersManual.pdf>.
- [2] Claim Master[EB/OL].[2013-05-15].<http://www.patentclaimmaster.com/>.
- [3] Patent Tools[EB/OL].[2013-05-15].http://www.pattools.com/claim_tree.html.
- [4] Patent Claims Tree[EB/OL]. [2013-05-15].<https://chrome.google.com/webstore/search/Patent%20Claims1>.
- [5] Web of knowledge[EB/OL].[2013-05-15].<http://apps.webofknowledge.com/>.
- [6] Thomson Innovation[EB/OL].[2013-05-15].<https://www.thomsoninnovation.com>.
- [7] 靖培栋.《中图法》可视化研究[J].中国图书馆学报,2009(3):42-46.
- [8] WIPO[EB/OL].[2013-05-15].<http://worldwide.espacenet.com/advancedSearch?DB=EPODOC>.
- [9] Lamping J, Rao R.The Hyperbolic Browser: A Focus + Context Technique for Visualizing Large Hierarchies[J].Journal of Visual Languages and Computing,1996,7(1):33-35.

作者简介

刘玉琴,男,1979年生,博士后,北京印刷学院绿色印刷包装产业技术研究院高级工程师,研究方向:专利分析、信息可视化,E-mail:liuyuqin2004@126.com。

Construction of Resolution and Visualization System of Patent Claims

LIU YuQin

(Academic of Printing and Packaging Industrial and Technology, Beijing Institute of Graphic Communication, Beijing 102600, China)

Abstract: This paper summarizes the claims structure features with claims descriptions, and studies the visualization method of patent claims which have been transformed into tree structure based on the hyperbolic tree. Then resolution and visualization system of patent claims is build. At last an empirical illustration about Apple's graphical user interface patent 7166791 is proposed.

Keywords: Patent claims; Hyperbolic tree; Visualization

(收稿日期: 2015-01-30)
编辑: 王立学