

云南省信息产业专利数据分析与预测

贺星月, 吴建德

(昆明理工大学 知识产权发展研究院, 昆明 650093)

摘要:随着信息产业不断创新发展,专利申请的数量也在不断增加。根据对云南省信息产业专利数据的检索结果,从专利申请量趋势、授权量状况和技术分布三个方面进行了一般的定性分析;并在此基础上,基于二次指数平滑法对短期专利申请量进行了定量分析,建立了较为准确的预测模型。通过对云南省信息产业专利情况的研究,为相关部门制定信息产业未来专利的发展方向与布局提供参考。

关键词:信息产业;专利数据;分析预测

中图分类号:G306 **文献标志码:**A **文章编号:**1671-1807(2018)01-0050-04

在当今时代,信息数量迅猛增长,庞大的信息数据中蕴含着大量的潜在价值,随着信息全球化的进一步扩张,处理信息能力的大小对促进国家、社会发展的作用越来越突出,也越来越受到各级政府的重视。目前新一代信息通信技术不断演进渗透,信息产业不断突破传统边界,内涵不断得到扩增。对于信息产业,更多强调的是其中技术的创新性与新颖性。专利申请作为技术保护的一个有效途径,可以直观的展现产业发展现状,反映技术发展水平。通过对云南省信息产业专利发展情况的描述、收集相关专利数据并进行定性与定量的分析研究,找出其中的发展规律与特点,可以为今后云南省信息产业的发展提供专利方面的帮助。

信息产业是对信息的管理,是通过信息技术,研究信息资源的开发、利用以及收集、生产、处理、储存和经营信息活动的产业。其中不仅有传统意义上的电子信息制造、软件、信息服务和信息通信传输行业,还有新一代的包括云计算、大数据、移动互联网在内的通信技术。

我国信息产业自上世纪 60 年代开始发展,而云南省信息产业自 1992 年才初步形成并起步发展,这使得全省信息产业的发展较为滞后,在全国同时期、同行业中处于较低水平。但随着近年来,一系列国家发展战略的实施,如中国—东盟自由贸易区、大湄公河次区域、经济走廊等给云南省信息产业的快速发展

创造了机会。近年来,信息产业相关专利逐年增长。其中,在电子信息制造基础材料、半导体功能材料、光电子器件、专用电子设备等高端电子信息产品制造领域发展较快,专利数量也较多;同时信息技术在各行各业产生了很多具有云南省特色的软件和信息技术服务产品,相关专利也逐年增加。按照行业统计,2015 年云南省信息产业完成主营业务收入达到 482 亿元^[1]。但是由于产业本身基础薄弱,创新能力不足等原因,导致专利申请数量与全国整体水平仍有很大差距。

目前全国均有关于区域信息产业专利情况的研究分析,如王婵娟等人对安徽省电子信息产业专利情况进行分析研究并提供对策^[2];张慧颖等人应用专利信息对天津市电子信息产业发展现状进行了对比研究^[3];赵富红分析了湖北省电子信息产业专利数据并研究其专利战略^[4]。相类似的文章都对相关地区专利数据进行了宏观的较为单一的专利分析,但是针对云南省信息产业专利数据的分析还较为空白。本文在参考了国内关于专利分析的一般类型基础上,选择了主要的几项专利分析指标进行研究,又基于云南省信息产业专利发展趋势较为稳定的态势,参考王利等人基于指数平滑法的 GPS 卫星钟差预报研究^[5],采用二次指数平滑法构建了关于年度专利申请量的线性预测模型,创新了专利分析的预测方法,并根据预测结果与当前实际情况相结合为未来云南省信息产

收稿日期:2017-10-21

作者简介:贺星月(1993—),女,山西忻州人,昆明理工大学知识产权发展研究院,硕士研究生,研究方向:数字信息管理与保护;吴建德(1979—),男,云南保山人,昆明理工大学知识产权发展研究院 教授、博士生导师,博士,研究方向:数据的分析与挖掘、科技创新与知识产权。

业专利布局提供参考。

1 云南省信息产业专利数据分析

1.1 数据检索

初期发展速度十分缓慢,专利信息数据的采集全面性、准确度与可利用程度低;而自 2006 年以来,随着知识产权保护意识的提升,专利申请数量增加,发展趋势稳定增长;因此本文选取佰腾专利检索数据库收集从 2006 年 1 月 1 日至 2016 年 12 月 31 日公开的专利文献信息,经过整理共得到 5 014 件相关专利数据。选取十一年的数据在时间跨度上较大,可以得出更加明显合理的发展规律。

1.2 专利申请量趋势

由于信息产业的重要性与必要性,政府大力扶持发展,技术的不断创新发展使得信息产业专利数据逐年增加。图 1 显示,2006 年至 2015 年之间,专利申请数量以及增长率均逐年递增,整体呈上升趋势。2006 年全省信息产业专利数量仅为 81 件,到 2016 年,已达到 1 024 件。其中自 2011 年以来,增长率的提升更为明显,但 2016 年申请量几乎与前一年持平,原因在于 2016 年的数据截止至检索日前仍有公开不全面与公布滞后性等因素,导致该年数据收集不完整的缺陷存在。

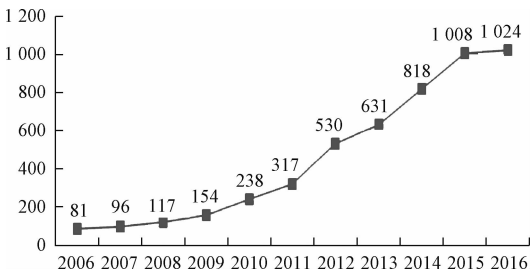


图 1 云南省信息产业专利申请量趋势

表 1 云南省信息产业专利技术分布表

	A	B	C	D	E	F	G	H
申请量	381	426	119	8	185	334	2 364	1 197
所占比例	7.60%	8.50%	2.37%	0.16%	3.70%	6.66%	47.15%	23.87%

2 云南省信息产业专利申请量预测

本文采用二次指数平滑法建立了专利申请量的预测模型,该模型由美国经济学家布朗提出^[6]。指数平滑法只需确定一个近期观测值的权数,则可推算出其余时期权数,在一次指数平滑法的基础上再作一次平滑进行迭代预测。它具有计算简单、样本要求量少、结果稳定等特点,适用于较平稳时间序列的短期预测。

1.3 专利授权量分析

通过分析数据可知,云南省信息产业专利申请中发明与实用新型所占比重为 1:1.39,由于发明所要求的创新水平高,且授权需经过实质审查,因此申请周期和申请难度均高于实用新型。如图 2 所示,发明专利的授权量占总申请量的 46.23%,而实用新型的授权量为 100%。这也从侧面反映了发明专利申请的质量有待提高,需要通过不断提高创新能力,减少授权率与申请率之间的差距。

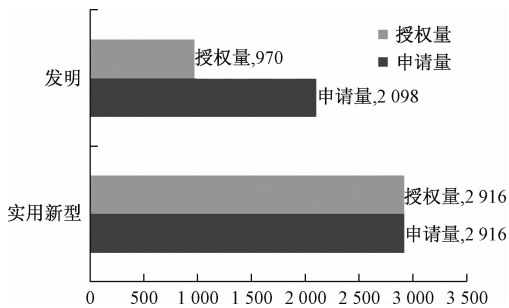


图 2 云南省信息产业专利授权量

1.4 专利技术分布分析

云南省信息产业 5014 件专利申请中,技术分布主要集中在 G、H 分类号中。其中,G:物理;共有专利 2364 件,占总体申请量的 47.15%。其中又主要聚集在分类号 G01 与 G06。G01 有专利 758 件,占云南省总申请量的 15.11%,占分类号 G 的 32.06%;G06 有专利 781 件,占云南省总申请量的 15.58%,占分类号 G 的 33.04%。H:电学;共有专利 1 197 件,占云南省总申请量的 23.87%。其中 H04:涉及专利 581 件,占云南省总申请量的 11.59%,占分类号 H 的 48.53%;H01:涉及专利 212 件,占云南省总申请量的 4.22%,占分类号 H 的 17.7%(见表 1)。

2.1 模型选取

指数模型的计算公式:

$$\begin{cases} S_t^{(1)} = ay_t + (1-a)S_{t-1}^{(1)} \\ S_t^{(2)} = aS_t^{(1)} + (1-a)S_{t-1}^{(2)} \end{cases} \tag{1}$$

$$\text{预测模型公式: } F_{t+T} = A_t + B_t T \tag{2}$$

$$\begin{aligned} A_t &= 2S_t^{(1)} + S_t^{(2)} \\ \text{其中: } B_t &= \left(\frac{a}{1-a} \right) (S_t^{(1)} - S_t^{(2)}) \end{aligned} \tag{3}$$

一次指数平滑公式中： $S_t^{(1)}$ 代表 t 时期的一次指数平滑值； a 代表平滑系数，其取值范围是 $(0,1)$ ； Y_t 代表 t 时期的观测值。该模型的关键在于确定平滑系数和初始值。实践中，时间序列所呈现的变化趋势不同，会使得相应的平滑系数 a 的取值也有所不同。一般按照波动程度大小，习惯于将取值区间分为三段 $[0.1,0.4]$ 、 $[0.4,0.6]$ 、 $(0.6,0.9]$ 。对于设置初始值，一般采用第一期的实际值或前几期的平均值作为初始平滑值。

2.2 具体预测分析

2.2.1 初始平滑值 S_0

本文所需分析的数据量较小，因此选用前三期数据的平均值作为初始平滑值可得： $S_0 = (Y_1 + Y_2 + Y_3)/3 = 98$

2.2.2 选择平滑系数 a

由表 2 数据可知，云南省信息产业专利数据变化趋势为逐年明显上升，为了进一步精确选择的合理性，本文选取了多个不同平滑系数取值区间相应的平滑值，分别计算各个平滑值的平均绝对误差进行误差分析，确定 a 的取值。误差计算公式为：

$$A \cdot D = \frac{\sum |Y_t - S_t^{(1)}|}{n}$$

当 $a = 0.3$ 的平滑值的平均绝对误差 $A \cdot D = \frac{1596.09}{11} = 145.01$ ；当 $a = 0.6$ 的平滑值的平均绝对误差： $A \cdot D = \frac{593.22}{11} = 53.93$ ；当 $a = 0.9$ 的平滑值的平均绝对误差： $A \cdot D = \frac{106.98}{11} = 9.73$ 。

由计算结果比较可得， $a = 0.9$ 时的平滑值的平均绝对误差最小为 9.73，因此选用 $a = 0.9$ 较为合适。接下来根据公式(1)进行计算可得表 2 结果。表 2 中记录了十一年间的实际观测值与经过一次平滑与二次平滑之后的申请量预测值。再利用折线图将实际值与 $S_t^{(2)}$ 值进行了拟合比较。通过图 3，可以直观的显示依据观测模型计算所得与实际值的拟合程度较高，因此预测值贴近于实际的观测值，模型选取较为合适。

2.3 预测公式

根据表 2 中的计算结果，建立预测公式。由于 2016 年数据检索时公开不全面问题的存在，导致检索量小于实际专利申请量，因此人工剔除掉 2016 年数据以减少误差，采用 2015 年的 $S_t^{(1)}$ 和 $S_t^{(2)}$ 计算数值求解 A_t 、 B_t ：

表 2 二次指数平滑数据

年份 (观察期 t)	申请量 (观测值 Y)	$S_t^{(1)}$ (一次指数平滑) ($a = 0.9S_0^{(1)} = 98$)	$S_t^{(2)}$ (二次指数平滑) ($a = 0.9S_0^{(2)} = 98$)
2006	81	82.7	84.23
2007	96	94.67	93.62
2008	117	114.77	112.65
2009	154	150.07	146.33
2010	238	229.2	220.91
2011	317	308.22	299.5
2012	530	507.82	487
2013	631	618.68	605.51
2014	818	798.07	778.86
2015	1 008	987	966.19
2016	1 024	1 020.3	1 014.89

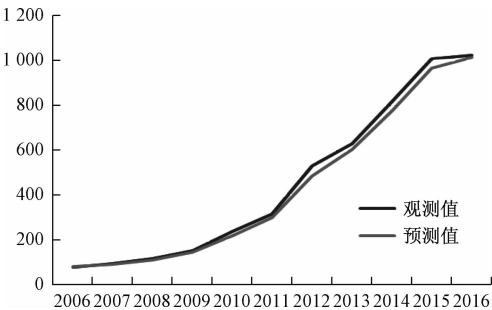


图 3 专利申请观测值与预测值拟合图

$$A_t = 2S_t^{(1)} - S_t^{(2)} = 2 \times S_{11}^{(1)} - S_{11}^{(2)} = 1007.81$$

$$B_t = \left(\frac{a}{1-a}\right)(S_t^{(1)} - S_t^{(2)}) = \frac{0.9}{0.1}(S_{11}^{(1)} - S_{11}^{(2)}) = 187.29$$

因此可得：预测经过 T 时期后的信息产业专利申请量的预测公式为：

$$F_{10+T} = 1007.81 + 187.29T \tag{4}$$

如计算 2018 年预测值为：此时 $T = 3$

$$F_{10+2} = 1007.81 + 187.29 \times 3 = 1569.68$$

据此，得出公式(4)为预测模型，根据公式，可以计算出短期内某一年度云南省信息产业专利数据的大致预测数据值。依据该值可预估计划本年度相关行业投入比例，制定年度信息产业规划等。但仍然存在缺陷，由于 2016 年的数据量有所欠缺，因此造成最后一期的指数平滑数据并不能完全正确的表示出数据量，但原因在于专利数据公开的不完整性，本文提供的预测计算方法依旧有效。

3 结论与建议

当前云南省大力推进信息产业技术创新，建立完善的信息产业整体布局，期待在新机遇下结合地理优势，打造东南亚信息辐射中心。但是仍然不可

忽视目前全省信息产业处于前期发展阶段的实际情况,整体发展水平也有待进一步提升。本文通过专利的角度,侧面对云南省信息产业情况进行分析,结合实际,提出今后信息产业专利发展布局的建议,从而更好的促进云南省信息产业规划目标的实现。

通过申请量趋势分析,可知信息产业专利申请量自2006年以来一直保持增长态势,增长率也稳步提升。但是结合前几年数据和预测模型得出的数据估计,若每年专利申请增长量保持在300个以内,仍然与全国同时期其他地区存在差距。因此,要加大对产业技术创新激励,提升产业整体创新意识,鼓励企业院校人才钻研开发,形成产业内部的有利竞争格局。专利保护内容的多样化,新颖性提高,会使得申请量大幅提升,这对信息产业发展的促进作用更明显。

数据显示目前实用新型专利授权率高于发明专利,实用新型接近百分之百的授权率,但是发明专利授权率仅占申请量的一半左右。也即对于发明专利更应重视其申请的有效性。在保证专利数量增长的前提下,也要兼顾质量的提高。发展竞争不单单以数量为标准,授权率的大小也应做为考核的对象。同时培训专门人才学习交流,在申请环节严格按照专利申请内容和书面格式的要求撰写、递交申请书,正确完成专利申请流程步骤,减少不必要的纠错程序,缩短授权周期,从而提高专利申请的授权率。

云南省信息产业专利技术集中分布在G、H分类号中,所以在电子信息制造等处理设备行业发展速度最快,今后仍然要继续重点突破,学习国内外先进技术,在现有基础上打造一批国内先进水平的本地优秀专利。目前云南省在信息处理和信息服务行业,发展程度要低于电子制造业,因此,今后应鼓励相关信息服务行业发展,采用引进社会项目,政府投资等方式吸引开发资源进入,加速该行业专利规模的扩大。同时政府在产业中起到促进作用,通过制定适宜的专利布局,为信息产业中的各行业提供政策性的指导与

建议,引导产业的全面整体发展。

文章所采用的预测方法在结合实际数据的情况下,可以较为合理的预测出未来几年专利数量的发展情况,可为相关部门制定专利规划、专利发展方向提供参考。但是模型仍然存在不足,需要不断修正以减少误差值。

参考文献

- [1] 云南省信息产业发展规划(2016—2025)[R]. 云南:云南省人民政府办公厅,2016.
- [2] 王婵娟,宋黎晓,宗剑钊.安徽省电子信息产业专利分析及对策研究[J].科技管理,2013(3):18—20.
- [3] 张慧颖,金笑非.应用专利信息的天津市电子信息产业发展现状及对策分析[J].重庆理工大学学报:自然科学,2015,29(2):110—116.
- [4] 赵富红.湖北省电子信息产业专利战略研究[D].武汉:华中科技大学,2009.
- [5] 王利,张勤,黄观文.基于指数平滑法的GPS卫星钟差预报[J].武汉大学学报:信息科学版,2017,42(7):995—1001.
- [6] 李建华,周挺进.基于指数平滑法的环境噪声污染预测模型及应用[J].环境科学与管理,2012,37(7):50—53.
- [7] 纪跃芝,范小娜.利用指数平滑法预测经济变量[J].长春工业大学学报,2003,24(2):61—63.
- [8] 秦超.基于专利数据挖掘的混合动力车柴油发动机专利分析[J].企业科技与发展,2014(4):4—8.
- [9] 袁有楼,成思.广东省专利申请数据预测与分析[J].科技管理研究,2007,27(8):259—261.
- [10] 何静,马虎兆.天津市电子信息产业专利信息分析及对策研究[J].工业技术经济,2009,28(1):13—17.
- [11] 马芳,王效岳.基于数据挖掘技术的专利信息分析[J].情报科学,2008,26(11):1672—1675.
- [12] 王慈光.二次指数平滑法中确定初始值的简便方法[J].西南交通大学学报,2004,39(3):269—271.
- [13] 李品秀,杨泽祥.云南省信息产业发展战略[J].印度洋经济体研究,2000,14(1):61—64.
- [14] 刘彤,侯元元,黄裕容.基于专利分析的国内外新能源汽车汽车技术发展现状对比研究[J].科技和产业,2016,16(8):74—79.
- [15] 何静,蔡利超.广东省智能机器人产业的分析与思考——基于专利的视角[J].科技管理研究,2017,37(15):155—163.

Analysis and Prediction of Patent Data of Information Industry in Yunnan Province

HE Xing-yue, WU Jian-de

(Yunnan Research Institute of Intellectual Property Rights Development, Kunming University of Science and Technology, Kunming 650093, China)

Abstract: With the development of information industry, the number of patent applications has also increased. Based on the patent data retrieval results of information industry in Yunnan Province, made the general analysis from three aspects: the patent application quantity, the authorization quantity and the technical distribution. On this basis, According to the Quadratic Exponential Smoothing Method, analyzed the short-term patent application quantity. Established an accurate prediction model. And from the research on the present situation of information industry patent, provided references for the related departments to design the future direction and layout.

Key words: information industry; patent data; analysis and prediction