

时间维度的云服务发展态势研究^{*}

王伟军 鲍丽倩 刘 凯

(华中师范大学信息管理学院 武汉 430079)

摘要:【目的】借助云服务概念簇数量的变化探究云服务当前的发展状况及其未来趋势。【方法】逐一抓取近12年Google搜索引擎收录云服务概念簇的词条数，采用时间序列分析法进行研究。【结果】云服务可分为陡增型、脉冲型和波动型三类。陡增型云服务会继续增长，但增长趋势会减缓；脉冲型云服务出现增长趋势的可能性不大；波动型云服务未来的发展走向不稳定。【局限】只从时间维度来分析云服务的发展趋势，如果要获取更加可靠的结果，需要结合政治、经济、社会等环境因素共同考量。【结论】云服务从实验室迈入市场，今后可能向着服务集中化管理方向发展，特定行业会有个性化定制的云服务。

关键词:云服务 XaaS 时间序列分析 预测

分类号:G350

1 引 言

早在20世纪60年代，麦卡锡(John McCarthy)就提出预测：计算迟早有一天会变成一种公共基础设施^[1]，而这些应用服务统称为云服务。1999年，Salesforce.com通过网站向企业提供企业级应用，提出云计算和软件即服务的理念，开创了新的里程碑^[2]。IBM和Google于2007年开始在云计算领域进行合作，云计算作为一种全新的商业和应用计算方式被正式提出，并迅速成为学术界和产业界研究的热点^[3]。据IDC预测，云计算服务全球增长率为26%，而中国的云计算市场保持在30%–40%的增长率。Gartner甚至认为云服务将在2014年取代PC而成为市场中的主要计算形式^[4]。

云服务应用已经深入人们生活的各个领域，正在引发IT业的新变革。它的发展加速企业信息化、推助电子商务的发展、优化互联网资源配置，对互联网产业、信息技术及电子商务领域影响深远，所以分析云服务近段时间的发展情况和预测今后的发展态势是极其必要的。

2 研究现状

云计算近几年发展迅猛，与此同时以“-aaS”为后缀的云服务也取得了爆炸式增长，被统称为“XaaS”概念簇^[5]，它概括了所有与云计算有关的服务，为大众所熟知的则是IaaS、PaaS和SaaS(简称SPI)。IaaS(Infrastructure as a Service): 基础设施提供商投资重要的技术、服务和数据中心，将IT作为服务提供给客户。PaaS(Platform as a Service): 平台提供商向使用者提供集成的开发环境、服务器和硬件等资源。使用者可以在平台上进行软件开发生命周期的一系列活动，包括软件开发、测试、部署和维护，还可以将应用程序与其他用户协作共享。SaaS(Software as a Service): 是软件提供商将应用软件统一部署在远程数据中心的服务器或服务器集群上供用户使用^[6]。

在产业界，云服务从出现就吸引了各界的关注。2007年，亚马逊的AWS、IBM和Google联合进行的蓝云计划，亚马逊开IaaS先河推出了“弹性云”模式^[7]。2008年，Google的App Engine^[8]和微软的Azure^[9]都推出了PaaS的模式，允许公司创建个性化

收稿日期：2013-11-29

收修改稿日期：2014-02-18

*本文系国家自然科学基金项目“基于用户偏好感知的SaaS服务选择优化研究”(项目编号：71271099)和湖北省自然科学基金创新群体重点项目“基于云计算的知识集成与服务研究”(项目编号：2011CDA116)的研究成果之一。

的应用，也允许独立软件厂商或者其他第三方机构针对垂直细分行业创造新的解决方案。Salesforce、NetSuite、Google 的 Gmail 及 SPSCcommerce.net 更是将 SaaS 发挥到极致，凸显提供软件的快捷和便利。

在学术界，国内外已经有学者对云服务发展趋势做了相关研究，主要视角可集中分为以下几类：

(1) 通过云服务的市场份额、市场价值等经济因素的变化来预测未来走势。Pring 等据此预测 SaaS 成为云服务中的领先指标，并且将会被云应用所代替，IaaS 及设施提供者对云服务的影响有望超过目前预期^[10]。

(2) 对云服务相关概念研究现状考察后作出预测。Leavitt 通过对云服务基础结构、实施、应用分类、性能和可靠性、安全和隐私等问题的阐述来确定未来的发展趋势，预测在未来 5 年内，云计算服务将成为中小型企业外包服务的解决方案，而大型企业则需要建设大型数据中心来获得高负载能力^[11]。Armbrust 等则从云服务的十点挑战和机遇入手，预测了三大走势：应用软件匹配付费许可模式，基础软件运行于虚拟主机，硬件系统设计于同一容器尺度内^[12]。

(3) 云服务在图书馆学领域的探索。刘炜提出一套由云服务检索方式、云服务器构架、云共享相结合的检索服务方式^[13]。

(4) 云服务相关技术的预测。冯宇彦等从云服务的技术架构切入，通过 4 种适配器将物理机、虚拟机、网络设备和存储设备整合，建立统一的资源池^[14]。

上述研究成果固然丰富，涉及经济学、计算机科学、情报学等领域，不过遗憾的是研究大多只是对 SPI 进行预测，而无法得知其他概念簇的发展态势。在 XaaS 向商务、社会领域融合的背景下，以技术为基础的 SPI 并不能代表整个云服务市场，因此，现有的研究无法对云服务进行更为细致的趋势预测。针对这一问题，本文将云服务概念簇展开进行详尽的研究。

3 研究方法与数据

3.1 研究方法

时间序列分析是一种重要的现代统计分析方法，研究一组按某种(如时间)顺序排列的动态数据的统计规律^[15]。20世纪 60 年代后期，在 Box 和 Jenkins^[16]

提出一套比较完善的建模理论及方法后，时间序列分析日趋完善，成为自然科学、社会科学领域中不可缺少的数据分析工具。目前，其理论和方法已被广泛应用到工程技术、地震、气象、水文、生物医学、经济管理以及军事科学等诸多领域。但是没有相关研究利用时间序列法分析云计算相关概念的发展趋势。作为互联网环境下诞生的新知识理念，云服务的发展动态及其未来走势对互联网产业、信息技术及电子商务等相关领域影响深远。而时间序列分析可以根据不同时间段云服务数据中所反映的发展过程、周期、方向和趋势，进行类推或延伸，借以推测云服务未来可能达到的发展水平。基于此，本文利用时间序列法对云服务近十几年的发展情况进行梳理并研究其未来走势，以期能够对云服务研究有所裨益。

3.2 数据来源

搜索引擎的词条创建数可以直观体现每个 XaaS 的产生时间、关注度和变化趋势，借此来预测概念簇未来的发展趋势。为确保样本来源范围足够广、样本量足够大，本研究选择 Web 搜索引擎作为数据采集工具。首先使用“SaaS、PaaS、IaaS 和 Ecosystem”4 个关键词对 Google、Bing、Yahoo、Lycos、Baidu 这 5 种知名的搜索引擎进行预测试。按搜索引擎的不同划分 5 个组，检全率用 4 个关键词的某个搜索结果条目之和与 5 组全部结果条目总量的比例进行表示，而检准率的确定则采用人工判定的方式计算前 10 页中命中条数与前 10 页总条数的比例。比对测试结果发现，Yahoo、Lycos 和 Baidu 搜索结果广度不够而 Bing 的精度不够，同时考虑到 Google 对学术数据库也有较好的涵盖，因此最终选取 Google 作为唯一的数据来源采集工具。

随后采用‘XaaS’ AND ‘as a service’的检索词形式进行搜索，X 取值分别由 A 至 Z，通过手工逐一搜集每个 XaaS 的中文名称、含义、搜索量、可发现的最早提出日期、最早提出者及出处并导入数据库。最终，整理出与云计算紧密相关的 XaaS 概念形式共 152 个。最后利用笔者开发的蜘蛛爬虫工具，采集了 Google 上所有概念词从诞生至今的词条数量，作为时间序列分析的数据来源。创建词条数可以直观表明该概念的使用频率及深度，从而代表了它在云服务领域的流行趋势。

3.3 数据预处理

考虑到时间序列随时间动态、整体的统计规律性，对数据进行预处理是非常有必要的。预处理过程既能够使序列随时间变化的特征体现更加明显，有利于模型选择，也能使得数据满足模型的要求。

由于部分云服务概念比较生僻，致使网络搜索词条数出现为零的情况，为保证结果的相对准确，对三

年以上为无记录的概念进行删除。此外，SPI 模型服务与其他 XaaS 服务数量级相差很大(后文专门研究该模型的序列图)，也在研究云服务整体发展脉络时予以删除，最后剩余 93 个 XaaS 云服务概念。将其导入 SPSS 21.0 中进行时间序列分析，在对有缺失值的数据进行修补，以及将数据资料定义为相应的时间序列之后，得出观察数据的时间序列图，如图 1 所示：

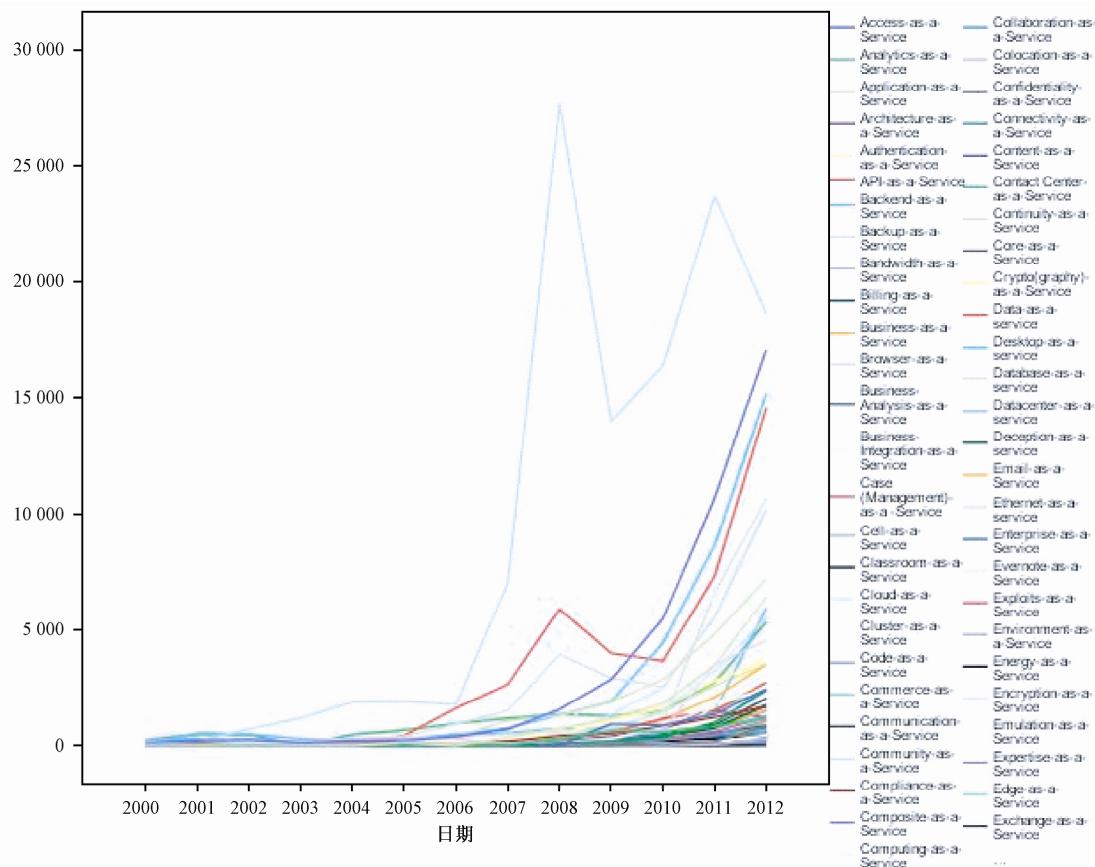


图 1 XaaS 时间序列总图

纵观图 1 中词汇的走向，XaaS 概念簇诞生于 2000 年，经历 2000 年—2005 年的孕育期后，于 2006 年开始爆炸式增长。该趋势完全符合云服务的生长模式，这是因为云服务是依托云计算的大数据支撑应运而生的，而且云服务必须要基于分布式系统和可扩展性机器等环境才能发挥效用，其取用方便、费用低廉等优点也令增长成为必然。

根据序列图曲线呈现不同的外形特点，可以将 XaaS 划分为三种类型：

(1) 陡增型。陡增型指的是序列图曲线首先基本

与时间轴平行，而后在某一个时间段出现大幅度上升。这类 XaaS 在 2000 年—2005 年之间几乎为零增长，从 2006 年—2008 年起呈现几何倍数增长，在短短几年内达到高峰。以 Access as a Service 为代表的云服务构成此类云服务概念簇的主体。陡增型揭示出 XaaS 蓬勃生长的原因，即云计算发展使得应用类服务变得快捷和便利，各种应用性服务大量出现且迅猛增长。

(2) 脉冲型。脉冲型是指序列图呈现 1—2 个脉冲，走向存在明显的起伏波动，突出表现为一个阶段的

上涨之后出现下滑趋势。结合此类概念的涵义及背景, 可以发现脉冲型云服务的两大特点: 一是与计算机软硬件紧密相关, 如 Database as a Service; 二是与特定企业紧密相关, 如 Yahoo as a Service 便直接以企业命名。互联网产品更新换代速度较快, 相关服务风靡一时之后的销声匿迹, 就符合这种起伏。专有名词命名的云服务, 由于用户的类型和数量非常有限, 也会出现脉冲式的增长态势。

(3) 波动型。此类序列图走势起伏频繁, 而且相隔年份的数量相差悬殊。本类概念簇走向极不稳定, 序列图至少出现三次以上的升降趋势。波动型的云服务并不常见且多与商业相关, 如 Business Integration as a Service。对词条进行深入分析后, 发现引发波动的主要原因在于这类云服务模式易受到当时政治、经济、贸易等外部环境因素的影响, 从而起伏较大。从图 1 中可见, 陡增型占比最大, 其次是脉冲型, 波动型最小。具体的数量分布如表 1 所示:

表 1 三种类型的数量分布

类型	涵义	走势特点	数量	比例
陡增型	出现快速增长	持续上升	51	54.84%
脉冲型	出现增长又降低	有 1~2 个起伏	34	36.56%
波动型	不断出现增长又降低	起伏频繁	8	8.6%

4 基于时间序列法的预测分析

4.1 模型识别

由云计算的序列图观察得到 XaaS 没有呈现季节性变化趋势, 不需要进行季节差分处理, 采用指数平滑中的无季节性模型——Brown 线性趋势模型和 Holt 线性趋势模型及 ARIMA 模型。

指数平滑法最先由 Holt 在 1958 年提出, 它最初只应用于无趋势、非季节作为基本形式的时间序列的分析, 后经 Brown、Winter 等统计学家的深入研究和发展, 使指数平滑设计的数据内部构成更丰富^[17]。

Brown 线性趋势的基本原理是将一次和二次平滑值之差加在一次平滑值上, 从而对趋势进行修正, 计算公式如下^[17]:

$$\begin{cases} S'_t = a x_t + (1-a) S'_{t-1}, \\ S''_t = a S'_t + (1-a) S''_{t-1}, \end{cases} \quad \begin{cases} a_t = 2S'_t - S''_t, \\ b_t = \frac{a}{1-a}(S'_t - S''_t) \end{cases} \quad (1)$$

其中: S'_t 为一次指数平滑值, S''_t 为二次指数平

滑值; m 为预测超前期数, F_{t+m} 为第 m 期预测值。

Holt 平滑法原理与布朗平滑法相似, 只是不用二次指数平滑, 而是对趋势直接进行平滑, 计算公式如下^[17]:

$$\begin{cases} T_t = a x_t + (1-\alpha)(T_{t-1} + b_{t-1}) \\ b_t = \beta(T_t - T_{t-1}) + (1-\beta)b_{t-1} \\ \hat{x}_{t+\tau} = T_t + b_t \tau, \tau = 1, 2, \dots \end{cases} \quad (2)$$

其中, t 为当期, τ 为预期超前期数, T_t 表示趋势估计, b_t 表示利用 t 期的数据对趋势增量 b 的估计, x_t 是第 t 期的实际观察值, $\hat{x}_{t+\tau}$ 利用前 t 期数据, 对第 $t+\tau$ 期的预测值, α, β 是平滑常数, 满足 $0 < \alpha, \beta < 1$ 。

ARIMA(求和自回归移动平均模型)是由 Box 和 Jenkins 于 20 世纪 70 年代初提出的著名时间序列预测方法^[16], 所以又称为 B-J 模型、博克思-詹金斯法。ARIMA 模型的优点在于: 该方法做出的预测比传统计量建模方法做出的预测更加可靠, 特别在短期预测方面^[18]。它包含三个主要的参数——自回归阶数(p)、差分阶数(d)和移动平均阶数(q), 一般模型形式记为 ARIMA(p, d, q)。

笔者从这三类中选择有代表性的云服务进行时间序列分析, 依据序列图的分析和专家建模分别采用不同的模型, 如表 2 所示:

表 2 模型选择及预测数据表

类型	概念簇	模型	预测值	增长率	趋势
陡增型	Backup as a Service	Brown	17 763	66.01%	↑
	ERP as a Service	ARIMA (0, 2, 0)	1 074	76.07%	↑
	Platform as a Service	ARIMA (0, 2, 0)	169 442	47.3%	↑
	Software as a Service	Brown	2 970 000	177.57%	↑
脉冲型	Database as a Service	Brown	8 336	-18.28	↓
	User Interface as a Service	Brown	628	234.33%	↑
	Infrastructure-as-a-Service	ATIMA (0, 2, 0)	131 045	-7.72%	↓
波动型	Filtering as a Service	Holt	13	62.50	↑
	Business Integration as a Service	Brown	39	-15.22	↓

4.2 模型检验

(1) 指数平滑法。“模型描述”可知当前模型所使

用的分析变量和方法。“模型拟合”给出了包括平稳 R 方在内的 8 个拟合优度统计量。“模型统计量”表格显示平稳 R 方统计量的取值都大于 0(0.613, 0.274, 0.135, 0.470, 0.116), 说明当前 Brown/Holt 线性模型要优于基本的均值模型。关于残差的自相关(ACF)和偏自相关(PACF)序列图都没有明显趋势特征(拖尾或截尾)。由观测序列, 拟合序列在图中高度相近的特点, 可以判断使用的模型是较为合理的。预测索引词条数分别为 Backup as a Service(17 763), Database as a Service(8 336), User Interface as a Service(628), Filtering as a Service(13), Business Integration as a Service(39)。

(2) ARIMA。模型参数输出中 t 统计量的显著性(sig 列)显示模型参数都非常显著地为 0。残差序列图也显示都没有明显趋势特征(拖尾或截尾)。由观测序列, 拟合序列在图中高度相近的特点, 可以判断 ARIMA 使用的模型是较为合理的。预测索引词条数为 ERP-as-a-Service(1 074)。

4.3 模型预测

SPI 作为基础云服务模型, 不能忽视它对整个云服务的影响。为了探测 SPI 模型是否指引云服务的未来的走向, 同样也对 SPI 模型进行时间序列分析。观之 SPI 模型中三个序列图走向(见图 2), 除 Software as a Service 从 2007 年开始上涨趋势明显外, Infrastructure as a Service 和 Platform as a Service 在 2010 年之后才凸显涨幅。由表 2 的预测数据可得: 就增长率而言, 预测 SPI 模型中检索词条最高的增长率 177.57%(Software as a Service), 同比上一年增长率下调了 316.87%。这一增长率并不是最高的, User Interface as a Service 增长率就达到 234.33%。但是从搜索词条数量上来看, SPI 的优势明显, 说明 SPI 模型已经深入人心, 仍然占有高比重。综上, SPI 模型虽然在词条数量上独占鳌头, 但其增长幅度却明显减缓, 更多新兴的云服务增长势头强劲, 很可能会反超 SPI 成为主要云服务。

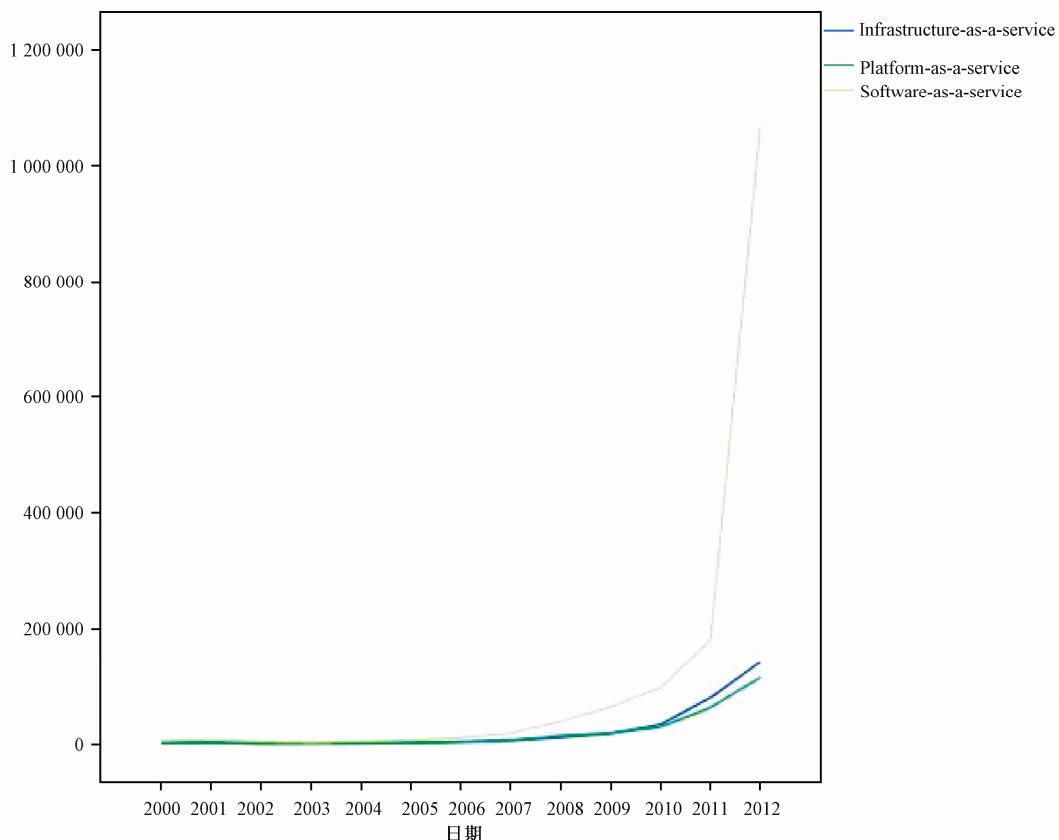


图 2 SPI 时间序列图

陡增型云服务接下来仍会呈现增长趋势，但可能增长幅度会减缓；脉冲型云服务未来发展前景并不明朗，可能会出现先降低再上升的趋势，也可能下滑，甚至停滞不前；波动型云服务由于受到环境的影响，同样发展不稳定。云服务的增长趋势说明云服务经历了从新的技术到应用形式的蜕变，企业也开始从“以产品为中心”向“以用户为中心”过渡。今后云服务可能向着更集中的方向发展，特定行业会有量身定做的云服务。

5 结语

(1) SPI 中 Infrastructure as a Service 将会呈现下降趋势，而 Platform as a Service 和 Software as a Service 则会出现持续增长趋势。在索引词条数目上，Software as a Service 占绝对优势，跟随其后的是 Platform as a Service。追溯其中缘由，可能是因为 IaaS 已经具体化并开始分化成其他应用类的 XaaS，虽说这个概念词条数量下降，但带动了其他云计算服务蓬勃发展。而 SaaS 与中小企业联系紧密，导致词条数目剧增。IaaS 需要解决的问题是数据安全和保密及系统可靠性，带宽不足和访问异常将会影响 PaaS 的顺利使用和共享。所以，SPI 要想顺利走下去，还需要攻克许多难关。

(2) 云服务的发展引起了图书情报界的关注和重视。SaaS 的迅猛增长预示着更多软件提供商将不断尝试软件提供模式，由于技术成熟，费用降低，更多的图书馆将有能力通过 SaaS 实现业务的自动化管理。云服务的发展同样会对存储带来革新：分布式和非结构化的数据存储模式，将操作系统、服务和用户数据都保存在云计算的存储体系中，只需在一个地方处理数据，就可以实现共享。

云计算的出现彻底打破了地域的概念，需要在网络资源存储、资源共享和网络安全监管等方面形成完备的监督和法律体系。在云计算与大数据日益结合的情况下，用户隐私保护是不容回避的一大挑战。如果在发挥应用虚拟化和数据资源整合的优势的前提下能够切实保障信息的安全性，那么云服务将会如同计算机一样引领潮流，彻底改变人们的生活。

参考文献：

- [1] Wikipedia.JohnMcCart [EB/OL].(2013-12-26). [2014-02-18]. [http://en.wikipedia.org/wiki/John_McCarthy_\(computer_scientist\)](http://en.wikipedia.org/wiki/John_McCarthy_(computer_scientist)).
- [2] Salesforce [EB/OL]. [2014-02-18]. <http://www.salesforce.com/company/>.
- [3] IBM.Google and IBM Announce University Initiative to Address Internet-Scale Computing Challenges [EB/OL]. (2007-10-08). [2014-02-18].<http://www-03.ibm.com/press/us/en/pressrelease/22414.wss>.
- [4] Gartner.Gartner Says the Personal Cloud will Replace the Personal Computer as the Center of Users'Digital Lives by 2014 [EB/OL].(2012-03-12). [2014-02-18].<http://www.gartner.com/newsroom/id/194731>.
- [5] Burns C.The Cloud Service Explosion [EB/OL].(2012-08-13). [2014-02-18]. <http://www.networkworld.com/supp/2012/enterprise4/081312-ecs-cloud-services-261195.html>.
- [6] 黎春兰, 邓仲华.云计算环境下的信息资源共享研究 [J].中国信息界, 2011 (12):66-70.(Li Chunlan, Deng Zhonghua. Sharing of Information Resources in the Cloud Computing Environment [J].China Information Times, 2011 (12): 66-70.)
- [7] Sims K.IBM Introduces Ready-to-Use Cloud Computing [N/OL]. [2014-02-18]. <http://www-03.ibm.com/press/us/en/pressrelease/22613.wss>.
- [8] Wikipedia.Google App Engine [EB/OL]. (2014-02-11). [2014-02-18].http://en.wikipedia.org/wiki/Google_App_Engine.
- [9] Wikipedia.Windows Azure [EB/OL]. (2014-02-11). [2014-02-18]. http://en.wikipedia.org/wiki/Windows_Azure
- [10] Pring B, Brown R H, Frank A, et al.Forecast:Sizing the Cloud;Understanding the Opportunities in Cloud Services [R/OL]. [2009-03-18]. https://img2.insight.com/graphics/uk/content/microsites/cloud/forecast_sizing_the_cloud_un_1665_25.pdf.
- [11] Leavitt N.Is Cloud Computing Really Ready for Prime Time [J].IEEE Computer Society, 2009, 42(1):15-20.
- [12] Armbrust M, Fox A, Griffith R, et al.A View of Cloud Computing [J].Communications of the ACM, 2010, 53(4): 50-58.
- [13] 刘炜.图书馆需要一朵怎样的“云”？ [J].大学图书馆学报, 2009, 27(4):2-6.(Liu Wei. How Libraries Uprising with the Cloud Computing [J].Journal of Academic Libraries, 2009, 27(4): 2-6.)
- [14] 冯宇彦, 赵占雪, 陈婉婷.云服务的解决方案探索 [J]. 信息技术与标准化, 2011(9):68-75.(Feng Yuyan, Zhao Zhanxue, Chen Wanting.The Exploration of the Solution for

- Cloud Service [J]. Information Technology & Standardization, 2011 (9): 68-75.)
- [15] 王燕.应用时间序列分析 [M].北京:中国人民大学出版社, 2005.(Wang Yan.Applied Time Series Analysis [M]. Beijing: China Renmin University Press, 2005.)
- [16] Box G E P, Jenkins G M, Reinsel G C. Time Series Analysis:Forecasting and Control [M]. Wiley, 2008.
- [17] 杜强, 贾丽艳.SPSS 统计分析——从入门到精通 [M].北京: 人民邮电出版社, 2012. (Du Qiang, Jia Liyan.SPSS Statistical Analysis—From Entry to Master [M].Beijing: Post & Telecom Press, 2012.)
- [18] 陈远, 王菲菲.基于时间序列的电子商务市场预测系统研发 [J].情报科学, 2009, 27(12): 1821-1833.(Chen Yuan, Wang Feifei.E-commerce Market Forecast System R&D Based on Time-series [J].Information Science, 2009, 27 (12): 1821-1833.)

作者贡献声明:

王伟军, 刘凯: 提出研究思路, 设计研究方案;
鲍丽倩: 数据分析和论文起草;
刘凯: 数据采集和清洗;
王伟军, 鲍丽倩, 刘凯: 论文修订。
(通讯作者: 王伟军 E-mail: wangwj@mail.ccnu.edu.cn)

Development Trends of Cloud Services in Time Dimension

Wang Weijun Bao Liqian Liu Kai

(School of Information Management, Central China Normal University, Wuhan 430079, China)

Abstract: [Objective] Explore the current status and development trends in future with the change of numbers of cloud services. [Methods] Grab entries of cloud service concepts of Google search engine in last 12 years, using time series analysis. [Results] Cloud services can be divided into steep-type, pulse-type and wave-type categories. Steep-type cloud service will continue to grow, but the growth trend will slow down. It is unlikely for pulse-type cloud service to grow. The development trend of pulse-type cloud service is instability. [Limitations] Analyze trends of cloud services only from the time dimension. [Conclusions] From laboratory to market, cloud services may develop in the direction of centralized management services in the future, there will be industry-specific customization of cloud services.

Keywords: Cloud service XaaS Time series analysis Forecasting