



基于借阅曲线分析的高借阅型馆藏生命周期研究

□陈雅迪* 李娟 梁栋 韩萌

摘要 高借阅型馆藏是大量读者广泛选择的结果,是读者兴趣的集中体现。掌握高借阅型馆藏生命周期的发展规律,有助于优化图书馆馆藏管理的实施策略。本研究以西安交通大学图书馆 2008 年至 2018 年的流通数据为基础,对高借阅型馆藏借阅曲线进行 Fisher 最优分割、特征提取和分类统计,归纳其生命周期的发展规律和影响因素。最后,根据研究结果对图书馆馆藏的评估、推广以及布局调整等方面提出建议。

关键词 借阅曲线 时序特征 Fisher 最优分割 使用分析 生命周期

分类号 G252

DOI 10.16603/j.issn1002-1027.2021.02.005

1 引言

在信息时代,日新月异的现代信息技术带来了多元化的信息渠道,纸本阅读已不再是唯一的知识获取方式。与此同时,图书馆传统的纸质馆藏普遍面临着借阅量下滑的严峻挑战^[1]。为改善纸质馆藏的流通现状,图书馆学界纷纷从资源建设、馆藏布局、宣传推广、服务创新等方面探讨应对策略^[2]。无论何种举措,都需要建立在对馆藏利用价值变化规律的洞悉与把握之上。随着时间的推移,馆藏的利用价值往往处于动态的变化过程中。曾默默无闻的图书可能突然一鸣惊人,曾红极一时的图书也可能逐渐无人问津。在流通数据中,蕴藏着馆藏价值的变迁轨迹;运用数据挖掘技术,图书馆可梳理并归纳馆藏借阅曲线的形态特征,准确把握藏书生命周期的变化规律,以此作为提升馆藏利用率的决策依据。

2 相关研究

“生命周期”的概念源于生物学领域,表征了生物历经孕育、萌芽、成长、成熟、老化、衰亡等生命阶段的规律性过程^[3]。随着其内涵的演化和发展,生命周期理论被广泛应用于各个学科领域。文献信息学领域将文献信息作为一种有机生命体,认为其价值有生命周期,也会经历成长期、成熟期、衰老期直至死亡的不同阶段^[4]。1985 年,美国学者马善德

(Marchand)和霍顿(Horton)提出了“信息生命周期管理(Information Lifecycle Management,ILM)”的概念^[5],图书馆学界也基于 ILM 理论,在资源评估^[6]、存储策略^[7]、归档方法^[8]等方面开展了大量研究。虽然其主题内容有所不同,但核心目的均为针对信息资源所处的不同生命阶段,采取相应的策略,以达到效益最大化。然而,这类研究往往停留于理论构想阶段,缺乏实践调查与科学论证^[9],特别是缺乏对文献信息利用价值的历时性实证研究。

对文献信息价值的历时研究主要集中于文献老化领域。文献老化是指,文献的效用随时间推移逐渐衰减,其价值逐渐过时,越来越少地被用户所利用。自 1944 年戈斯内尔(Gosnell)^[10]提出文献老化概念以来,一系列经典的文献老化数学模型被相继提出,如负指数模型、B-K 模型等。在文献老化数学模型的基础上,人们建立了半衰期、普赖斯指数、文献老化率等测度指标,以实现文献老化程度的定量衡量^[11]。总体来看,传统的文献老化研究往往将视角局限于文献生命周期的衰退期,其建立的数学模型也多为单调下降的函数,较少关注文献自诞生以来其利用价值随着传播、扩散而成长、成熟的前期增长阶段^[12]。

从研究对象来看,已有研究多针对电子资源进行讨论,对纸质图书的研究相对较少,且大多为定性

* 陕西省图书馆学会课题“流通数据分析及读者阅读行为研究”(编号:191017)研究成果之一。

通讯作者:陈雅迪,ORCID: 0000-0003-3874-1958,邮箱:chenyd@mail.lib.xjtu.edu.cn。



分析。在已有的图书老化的实证研究中,一般利用入藏年份、半衰期等静态指标对某一大类图书进行分析。如邓奇强等通过统计 TP 类图书的借阅量,获取其半衰期及老化曲线,在此基础上筛选老化图书,降低其在个性化推荐中的推荐强度^[13];彭跃峰运用文献老化的负指数模型,对华南农业大学图书馆各类图书的老化曲线进行拟合,并获取其半衰期,对其老化规律及影响因素进行分析^[14];段峰等人分别应用同时法与历时法对山东大学图书馆的借阅数据进行分析,计算出图书老化的半衰期,应用于调整馆藏布局的策略^[15]。这类研究主要聚焦于藏书生命周期的老化阶段。尽管也有研究考虑到在知识传播过程中,图书借阅量随时间推移而增长的前期过程,并建立了相应的数学函数模型^[16],但当研究对象细化到藏书个体时,实际的借阅曲线往往千差万别,不能被某一种数学模型精确概括。

近年来,在文献老化领域的范畴下,兴起了针对文献引文曲线的分析研究。通过对文献的历时性引文分析,归纳出不同形态的引文曲线,进而发现其对应的生命周期特征。针对文献引文曲线的研究将时间分布的因素加入到传统的文献老化测度指标中^[17],为文献生命周期的研究提供了一种崭新的思路^[18]。

本研究借鉴了引文曲线的分析方法,将图书年度借阅量在时间轴上的分布绘制为借阅曲线;通过对借阅曲线形态的分类讨论,探究其生命周期的特征及产生机理。本研究的独特性在于,并未事先假定图书借阅曲线服从某种数学模型的分布,而是利用 Fisher 最优分割法提取借阅曲线的生命阶段特征,从而实现对大量图书借阅曲线的自适应分类。Fisher 最优分割法是一种常用的有序数据样本分割方法,通过分析有序数组内的离差分布,获得对有序数组的自适应分割,目前已在河流汛期分期^[19]、图像分割^[20]、地层划分^[21]、经济发展阶段分析^[22]等方面广泛应用。

3 借阅曲线的特征提取方法设计

理想情况下,借阅曲线有三种基本类型:单调递增的“上升型”借阅曲线,表明该图书借阅量逐年增加,越来越受到读者的喜爱;单调递减的“下降型”借阅曲线,表明该图书的借阅量逐年下降,其影响力日趋衰减;走势平稳的“稳定型”借阅曲线,代表历年来

读者对该图书的需求变化不大。但在实际中,不同图书借阅曲线的形状往往千差万别;每种借阅曲线是由上述典型曲线排列组合形成的复合曲线,反映出其影响力发展的不同阶段。由于借阅曲线是按时间先后顺序排列构成的,在进行阶段划分时,其前后次序不可更改。因此,“对借阅曲线的分段问题”本质上属于有序数据的最优分割问题。为了能识别每种图书借阅曲线的发展阶段,采用 Fisher 最优分割法对其进行分段统计。

3.1 分割方法

若某图书共有 n 年的流通记录,其历年的年度借阅量按时间升序排列为数据序列 X :

$$X = \{x_1, x_2, \dots, x_n\} \quad (1)$$

若该图书的借阅曲线包含 K 个发展阶段,则应将 X 划分为 K 段:

$$\{x_{i_1}, x_{i_1+1}, \dots, x_{i_2-1}\} \{x_{i_2}, x_{i_2+1}, \dots, x_{i_3-1}\} \dots \{x_{i_K}, x_{i_K+1}, \dots, x_{i_{K+1}-1}\} \quad (2)$$

不考虑重复的划分方法,则共有 C_{n-1}^{K-1} 种分法。根据 Fisher 最优分割法,存在一种划分,使得每段内部的差异之和最小,段间的差异之和最大,该方法即是对有序数据的最优分割。

定义某段 $x(i, j) = \{x_i, x_{i+1}, \dots, x_j\}$ 的离方差之和为段直径 $d(i, j)$, 即:

$$d(i, j) = \sum_{a=i}^j [x_a - \bar{x}(i, j)]^2 \quad (3)$$

公式(3)中, $\bar{x}(i, j)$ 为段内样本的均值。段直径越小,表明在该段内图书的各年度借阅量的差异越小;段直径越大,表明在该段内图书的各年度借阅量的差异越大。将误差函数 W 定义为段内离方差之和,则 Fisher 最优分割的最佳划分方法应当满足:

$$W = \min \left[\sum_{j=1}^K d(i_j, i_{j+1} - 1) \right] \quad (4)$$

通过遍历分割点的位置,使公式(4)的值最小,即可获得图书年度借阅序列的最优分割。为了降低计算复杂度,采用动态递归策略优化计算过程。递推公式如下:

$$W(n, 2) = \min[d(1, i-1) + d(i, n)], 2 \leq i \leq n \quad (5)$$

$$W(n, K) = \min[W(i-1, K-1) + d(i, n)], K \leq i \leq n \quad (6)$$

在求解 X 的最优 K 段分割时,以公式(6)的值



最小为目标,求分割点 i_K ;再通过递归,最终得出 X 在段 $\{1, \dots, n\}$ 上的最优 K 段分割方案。

3.2 确定最优分割数

分割次数越多,每段数据内的差异越小,当 $K = n$ 时,每段数据内的离方差为 0。但分割次数并非越多越好:当 K 增加到一定程度后, W 的变化将趋于平缓,此时不必再增加分割次数。如果将 $W(K)$ 视为一个有序序列,对其进行最优 2 段分割,获得分割点 a ,则最优分割数 $K = a$ 。以《追风筝的人》一书为例,如图 1 所示,对序列 $W(K)$ 进行 Fisher 最优 2 次分割,其分割点位于 $K = 3$ 处,因此,该图书借阅曲线的最优分割次数为 3 次。

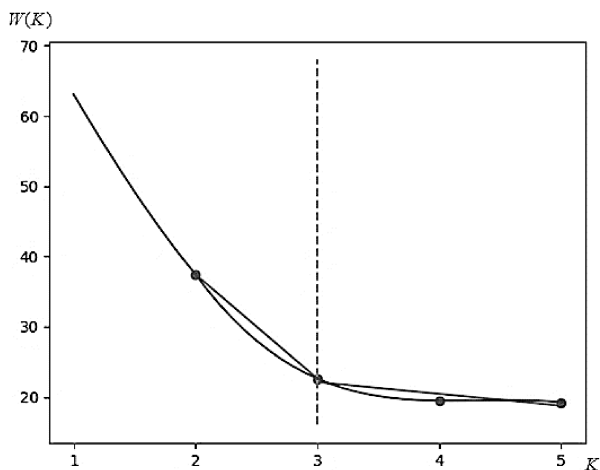


图 1 根据 $W(K)$ 确定最优分割次数

3.3 借阅曲线的特征提取

为了表征各分段曲线的特征,避免因数据波动带来的干扰,对分割后的子序列进行线性拟合,以所得线段的斜率作为子段的特征。

对公式(2)的各分段依次应用最小二乘法进行线性拟合:

$$f(t) = kt + b \quad (7)$$

获得由各分段的斜率构成的特征:

$$F = \{k_1, k_2, \dots, k_K\} \quad (8)$$

3.4 借阅曲线的特征评价

在实际中,人们对图书的借阅趋势常用定性概念进行评价,例如“近年来,某图书的借阅量显著增加”。这种描述符合人们的评价习惯,但存在主观性及不确定性,不宜用精确的定量描述方式界定。因此,引入了模糊辨识方法^[23],采用 5 级指标建立特征评价模糊集:

$$U = \{\text{显著下降、下降、平稳、上升、显著上升}\} \quad (9)$$

为建立馆藏借阅曲线的趋势特征与各评价类别的映射关系,构造如图 2 所示的多态函数作为该模糊集的隶属度函数。

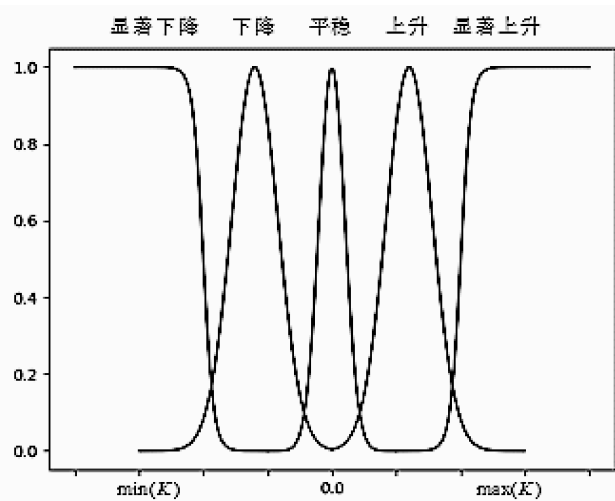


图 2 借阅趋势评价模糊集隶属度

其中,特征值越接近于 0 的借阅曲线,属于模糊子集“平稳”的程度越高,故以对称中心为 $k=0$ 的高斯函数作为其该模糊子集的隶属度函数;与最大特征值越接近的借阅曲线,其“显著上升”的隶属程度越大,故选择单调递增的 Z 型函数作为其隶属度函数;特征值越接近最小特征值的借阅曲线,其“显著下降”的隶属程度越大,选择单调递减的 Z 型函数作为其隶属度函数。对模糊子集“上升”和“下降”,根据实际特征值的分布统计,将其隶属函数分别定义为以正半轴特征均值为中心的高斯函数,和以负半轴特征均值为中心的高斯函数。将实际的特征数据向量压缩至 $[0, 1]$ 区间,并利用最小二乘法估计隶属度函数的参数,建立模糊集合的隶属函数。根据最大隶属度法则进行模糊推理,获得馆藏的趋势特征在模糊集合 U 上的评价结果。

4 高借阅型馆藏借阅曲线的分类

由于高借阅型馆藏是大量读者广泛选择的结果,是读者兴趣的集中体现,其生命周期也具有更加明显的特征。如能揭示高借阅型馆藏生命周期的变化规律,则可针对性地调整业务实施举措,助力馆藏利用率的提升。因此,本文选取年度借阅量排名前 1% 的高借阅型图书作为研究对象,对其借阅曲线进行分类统计。

为充分展现馆藏借阅量随时间变化的规律,从



西安交通大学图书馆的流通日志中,抽取 2008—2018 年的数据作为实验来源。在统计时间段内,图书馆的年购书数量和复本政策没有明显变化。从流通日志中遴选出年度借阅量排名前 1% 的图书。在去除“通用借书证”“IPAD”“MAC”等非纸质图书的其他类型资料后,共获得 15816 种高借阅型图书,及其产生的 1135792 条借阅记录。将每种图书的年度借阅序列绘制为借阅曲线,按照图 3 所示的分析流程,应用 Fisher 算法对借阅曲线进行自适应分割和特征评价。

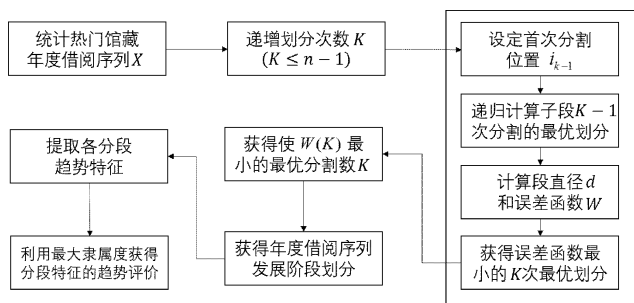


图 3 高借阅型馆藏趋势分析流程

如图 4 所示,共获得 23 种典型借阅曲线。根据其特征在时间轴上的分布情况,将其归纳为 6 类,分为“上升型”“下降型”“平稳型”“先上升后下降型”“先下降后上升型”“多峰波动型”曲线。

由图 4 可知,在近十年的统计时间窗内,呈“先上升后下降型”的借阅曲线最多,占总量的 55.57%;呈“平稳型”的借阅曲线最少,仅占总量的 0.4%。因此,高借阅型馆藏的借阅趋势并非一成不变,而是多呈波动态势。对不同类型借阅曲线的图书上架时间进行分析,获得借阅曲线随图书上架时间的分布图。如图 5 所示,“上升型”图书的上架时间大多分布于 2015 年之后,且随着时间的推移所占比重不断增加;“下降型”图书的上架时间大多分布于 2008 年之前,且随着时间的推移所占比重不断减少;“先上升后下降型”图书的上架时间多分布于 2006 至 2015 年之间,总体上呈现为两边窄中间宽的“纺锤形”;多峰型曲线的上架时间分布相对均匀;“平稳型”和“先下降后上升型”图书的上架时间主要分布于统计时间窗右侧附近。

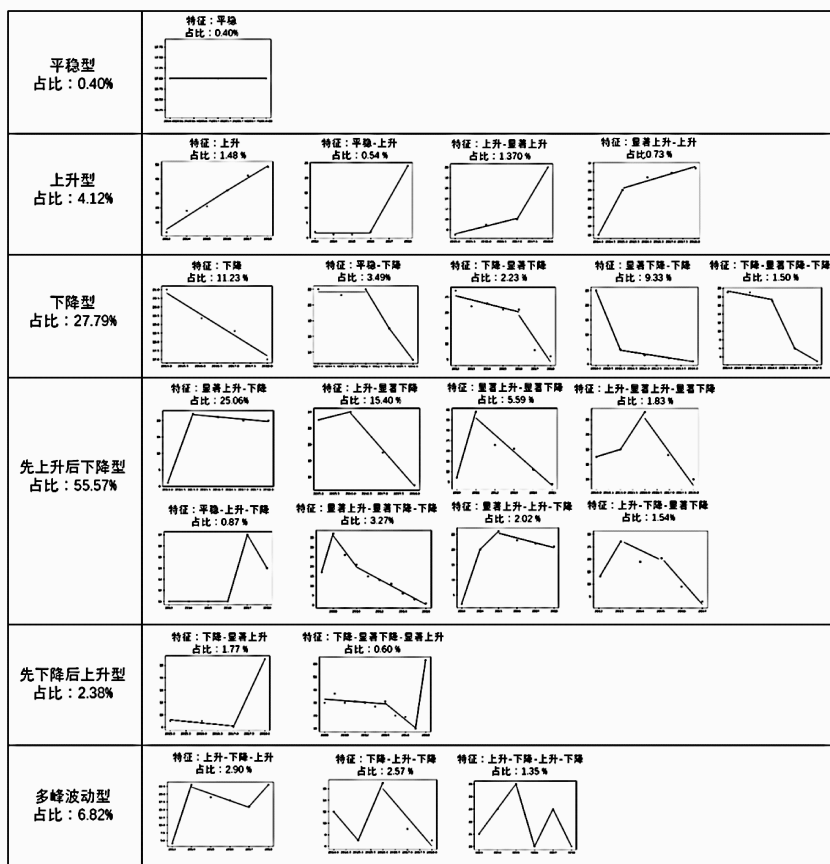


图 4 典型借阅曲线及数量占比



统计结果表明,高借阅型图书的借阅曲线以单峰型曲线为主,该现象符合生命周期理论的一般规律。根据生命周期理论,文献(图书)自其发表(上架)以来,随着其影响力的扩散,其单位时间的引用量(借阅量)逐渐增加,直至到达顶点;随后,因文献逐渐老化,其引用量(借阅量)逐渐衰减。在图4观测到的典型曲线中,出现次数最多的正是“先上升后下降型”借阅曲线;图5也印证了这一点,由于统计时间窗无法覆盖所有图书的完整生命周期,因此观测到了其生命周期的不同阶段:上架时间靠近2018年的图书,其借阅曲线大多表现为“上升型”,表明在统计窗内观测到了其生命周期的生长阶段;上架时间靠近2008年的图书,其借阅曲线大多表现为“下降型”,表明在统计窗内观测到了其生命周期的老化阶段;上架时间位于统计时间窗中心附近的图书,大多表现为“先上升后下降型”,表明在统计窗内既可观测到其生命周期的生长阶段,又可观测到其老化阶段。

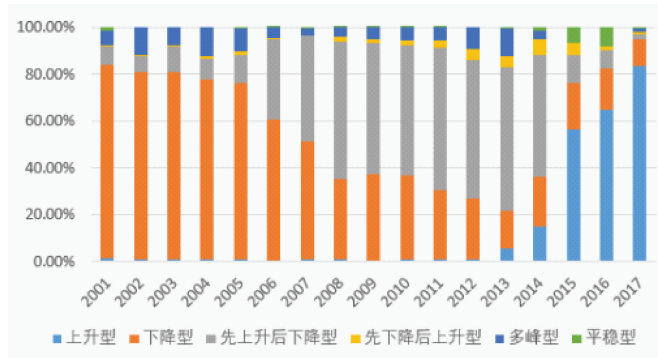


图5 借阅曲线随图书上架时间的分布图

实验发现,“下降型”借阅曲线的总量约为“上升型”借阅曲线总量的6倍。如果高借阅型馆藏的生命周期为对称型分布,则采样结果中的上升型和下降型曲线的总量应相差不大;因此,大多数图书的生命周期应为“下降”阶段长于“上升”阶段的偏态分布曲线。这意味着高借阅型馆藏往往以较短的时间达到其借阅顶峰,随后以较长的时间缓慢衰减。由此,构建如图6所示的单峰型偏态曲线,表征高借阅型图书的经典生命周期。

如图6所示,经典生命周期可分为沉睡期、萌芽期、成长期、成熟期、老化期、衰退期和衰亡期七个生命阶段。在每个生命阶段内,图书具有相似的借阅趋势;在不同生命阶段之间,图书的借阅趋势具有明显的区别。经过Fisher算法的分割,借阅曲线被自

适应地划分为段内差异不大、段间差异较大的数个子段;恰好可对应为其所经历的不同生命阶段。因此,根据借阅曲线的分段特征及相对位置关系,建立其生命阶段的判别规则,如表1所示。结合图4的实验结果,归纳高借阅型图书生命周期的发展规律。

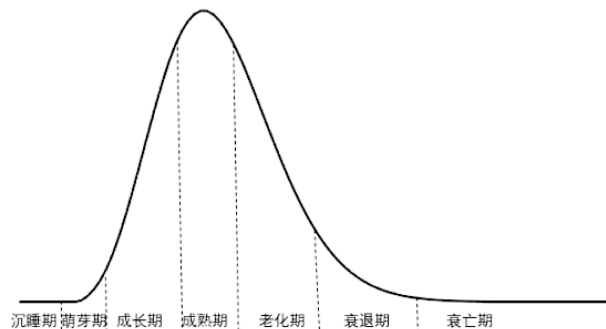


图6 高借阅型图书的经典生命周期及生命阶段划分图

表1 生命阶段的判别规则

生命阶段	趋势特征	位置特征
沉睡期	平稳	位于借阅曲线首段
萌芽期	上升	沉睡期后段或成长期前段
成长期	显著上升	/
成熟期	上升 平稳 下降	成长期后段 老化期前段
老化期	显著下降	/
衰退期	下降	老化期后段或衰亡期前段
衰亡期	平稳	位于借阅曲线尾段

注:表中“/”表示该生命阶段具有标志性的趋势特征,在判别过程中无需考虑位置特征,可根据趋势特征直接确定。

4.1 “上升型”借阅曲线

“上升型”借阅曲线表现为单调递增的形态。“上升型”曲线共有四种典型形态,其特征分别为“平稳—上升”“上升”“上升—显著上升”“显著上升—上升”,其对应的生命周期发展阶段分别为“沉睡期—萌芽期”“萌芽期”“萌芽期—成长期”“成长期—成熟期”。“上升型”借阅曲线主要反映了图书生命周期中的前半程,基本符合“沉睡—萌芽—成长—成熟”的规律,表明文献自入藏后,其借阅量并非立即到达顶峰,而是存在着一个传播、扩散的生长过程。这一现象印证了已有研究的观点^[24]。

同时,在实验中还发现一种特殊的借阅曲线——“睡美人”型曲线。这类曲线具有较长的沉睡期,以及一个明显的唤醒时间点,其特征表现为“平



稳一显著上升”,表明图书在上架后的较长时间内处于低借阅状态,而后其借阅量突然上升。分析“睡美人”型曲线的产生机理及影响因素,对于提升馆藏的借阅量具有非常重要的启示。

4.2 “下降型”借阅曲线

“下降型”借阅曲线表现为单调递减的形态。“下降型”曲线共有五种典型形态,其特征分别为“平稳一下降”“下降”“下降—显著下降”“显著下降一下降”“下降—显著下降一下降”,其对应的生命周期发展阶段分别为“成熟期—衰退期”“衰退期”“成熟期—老化期”“老化期—衰退期”“成熟期—老化期—衰退期”。总体来看,“下降型”借阅曲线反映出,随着时间推移图书的利用价值逐渐降低直至衰亡的过程,基本符合“成熟期—老化期—衰退期”的规律。但实验发现,有一类图书的借阅曲线不符合这一规律,未经历逐步老化的进程,却突然以“显著下降”的特征,从较高的借阅量断崖式衰减至0,其借阅曲线呈现为“流星型”形态。如能找到导致其生命周期戛然而止的因素,则可针对性地采取预防措施,防止图书正常的生命周期受到影响。

4.3 “先上升后下降型”借阅曲线

“先上升后下降型”借阅曲线表现为先单调递增、后单调递减的形态。作为占比最多的曲线类型,“先上升后下降型”借阅曲线表现出很明显的生命周期特点:在统计窗内,既观测到了其生命周期中的生长过程,又观测到了其生命周期的老化过程。这类曲线共有8种典型曲线,总体来看,“先上升后下降型”曲线大体服从“沉睡—萌芽—成长—成熟—老化—衰退—衰亡”的生命周期发展历程,但部分曲线会跳过峰值附近的成熟期或老化期,直接步入下一阶段。

与理想中平滑的生命周期曲线不同,实际的借阅曲线峰值附近的斜率往往非常陡峭。经过计算发现,57.2%的高借阅型馆藏的成熟期不超过2年,91.63%的高借阅型馆藏的成熟期不超过4年。这表明大多数高借阅型馆藏成熟期较为短暂。

4.4 “平稳型”借阅曲线

“平稳型”借阅曲线没有明显的单调递增或递减规律,也没有明显的生命阶段特点。与大多数高借阅型馆藏不同,这类图书的借阅量始终保持在较高水平,表明其经久不衰地得到读者的喜爱。尽管这类曲线的数量最少,但对这类图书进行深入分析,有

助于掌握读者群体的阅读偏好,为文献资源建设和阅读推广提供决策依据。

4.5 “先下降后上升型”借阅曲线

“先下降后上升型”借阅曲线表现为先单调递减、后单调递增的形态。这类图书与一般的曲线不同之处在于,其具有两个连续的生命周期。在统计窗内观测到了前一个生命周期的老化阶段,和后一个生命周期的生长阶段。这种形态特点,表明其受到了某种外部因素影响,使其原本已经衰减式微的生命重新焕发出“第二春”。

4.6 “多峰波动型”借阅曲线

“多峰波动型”借阅曲线中包含了多个生命周期,其峰值不完全相同。在前一个生命周期进入尾声时,因某种因素影响产生了后一个生命周期,最终形成了多个生命周期排列组合的形态。总体来看,多峰波动型曲线表现出上下波动的态势,表明这类图书的借阅量较不稳定。

5 高借阅型馆藏生命阶段及借阅曲线的分布特点

5.1 高借阅型馆藏的生命阶段的历时分布

馆藏的生命周期由不同生命阶段构成。在不同的生命阶段中,馆藏的借阅价值和发展趋势都有所不同。如果将馆藏的生命阶段类比为人类的年龄阶段,则“沉睡期”“萌芽期”“成长期”“成熟期”“老化期”“衰退期”“衰亡期”可分别对应于“婴儿时期”“幼儿时期”“青年时期”“中年时期”“壮年时期”“老年时期”“晚年时期”。因此,不同生命阶段下馆藏的数量及占比情况构成了馆藏集合的“年龄结构”。对高借阅型馆藏生命阶段的历时分布进行分析,可以发现其“年龄结构”随时间推移的变迁过程。

为探究高借阅型馆藏在各年度下的“年龄结构”,依据表1所示的规则,以年度为单位,逐年计算其生命阶段的占比情况。如图7所示,高借阅型馆藏的生命阶段主要由“成长期”“成熟期”“老化期”“衰退期”构成,处于“成长期”的图书是高借阅型馆藏的主力军。然而,在2008年至2014年,“成长期”和“成熟期”的图书占比从80.1%下降至49.32%;而“老化期”“衰退期”“衰亡期”图书的占比逐年上升,并在2013年超越了总量的50%;与此同时,高借阅型馆藏的总量在2010年后呈现出逐年下降的态势。该现象说明,高借阅型馆藏的年龄结构在这一阶段逐渐趋于老龄化:大量曾经处于“成长



期”的图书,在经历了生长与老化进程后,进入生命的尾声而逐渐消亡;与此同时,年轻的“成长期”图书未能补充进来,出现了“青黄不接”的局面。在2014年之后,尽管“成长期”和“成熟期”图书的数量略有回升,但由于大量“老龄化”图书逐渐退出历史舞台,导致高借阅型图书的总量仍处于下降趋势。

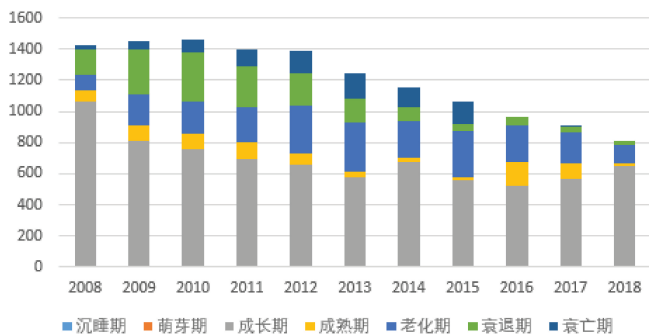


图7 高借阅型馆藏生命阶段的历时分布

5.2 高借阅型馆藏的借阅曲线在图书类别中的分布

根据《中国图书分类法》中的13个大类,对5种借阅曲线分别计算其在各类别下的分布。如图8所示,在各类别图书中,“先上升后下降型”型曲线是出现概率最大的借阅曲线;“平稳型”曲线是出现概率最小的借阅曲线。这是不同类别图书借阅曲线分布的共性特点。但从分布结构来看,不同图书类别之间存在着一些差异。

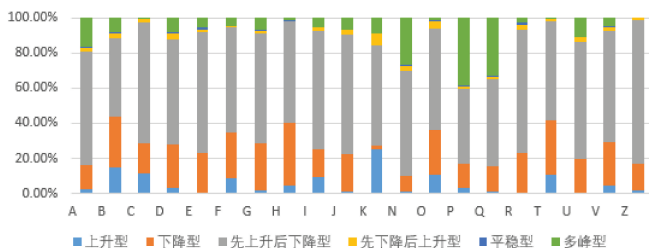


图8 借阅曲线在各类别图书中的分布

大多数图书类别中,“下降型”曲线数量总体高于“上升型”曲线,特别是E、G、J、N、Q、R、U、Z类图书,其“上升型”曲线的占比量极少,表明这类图书的高借阅型馆藏目前大多处于其生命周期的中、后半程,如果未来没有更多“年轻”的高借阅型图书补充进来,原有图书贡献的借阅量可能会逐渐衰减。

“上升型”曲线在各类图书中表现为不均匀分布。总体来看,“上升型”曲线在B、C、F、I、K、O、T类图书中出现较多。特别是对K类图书而言,其

“上升型”曲线甚至超过了“下降型”曲线的数量。其原因可能是,一方面,历史、地理类图书的老化速度较慢,因此下降型的借阅曲线占比较少;另一方面,一批近年上架的图书,如《习近平的七年知青岁月》《苦难辉煌》《丝绸之路》等“上升型”图书持续“走红”,受到读者的广泛喜爱,表现出强劲的增长势头。

“先下降后上升型”曲线多分布在D、K、O、R等类图书中。这类图书大多上架时间较早,原本处于其生命周期的后半程,但由于外界因素影响,其发展趋势从“下降”转变为“上升”。改变这类图书的借阅曲线形态的外界因素可能源于以下几个方面。

(1)书单推荐。出现在“习近平书单”中的《论党的精神》《社会契约论》等图书在推荐前已步入老化阶段;但在得到推荐后,受到读者的热烈欢迎,开启了生命周期的新阶段。

(2)竞赛活动。例如,以数学建模竞赛为代表的课外科技活动,近年来在全校范围得到了广泛参与。特别是在校内赛夏令营举办后,其参与人数逐年攀升。据统计,2011年的参赛人数约为2007年参赛人数的三倍。与此同时,与“数学建模”相关的图书也变得炙手可热,一些原本进入老化阶段的图书也重新焕发生机。

(3)热点信息。2016年4月8日是西安交通大学建校120周年、迁校60周年的校庆纪念日。全校范围内开展了“风云两甲子,弦歌三世纪”为主题的校庆盛典活动;而图书馆中校史相关馆藏也成为借阅热门。随着“西迁精神”学习热潮的迅速掀起,校史类图书的借阅热度也持续上升。以《交通大学西迁回忆录》为代表的图书,尽管书龄已超过十年,但在2016年后又步入了借阅量上升的生命阶段。

6 特殊借阅曲线的特征分析

6.1 “睡美人”型曲线

在高借阅型馆藏集合中,筛选出130种“睡美人”图书。经过统计发现,其沉睡期从4年至9年不等,平均沉睡期为6.35年。从唤醒时间点的分布来看,2018年是“睡美人”图书唤醒最多的一年,有74.61%的“睡美人”图书在该年被唤醒。经过分析发现,该现象与2018年开展的“百本经典”阅读推广活动具有明显的相关性。在2018年被唤醒的97种“睡美人”图书中,近40%的图书位于“百本经典”活动的阅读推荐目录之中:如《如何阅读一本书》《爱的



艺术》《逻辑的力量》等。这些图书曾经长期处于低借阅状态,但在2018年开展“百本经典”阅读推广后,其当年借阅量相较早年平均增加4.93倍。这表明,阅读推广活动能够将原本沉睡的图书唤醒,赋予其蓬勃的生命力。

6.2 “经久不衰”型曲线

与大多数高借阅型馆藏不同,有少数馆藏的成熟期持续了相当长的时间,表明其经久不衰,持续得到读者的喜爱。对成熟期大于4年的447种高借阅型图书进行主题内容归纳分析,发现其主要集中于以下方面:(1)经典党政文献,如《毛泽东选集》《邓小平文选》《资本论》等;(2)经典小说类书目:如《平凡的世界》《白鹿原》《傲慢与偏见》等;(3)学科奠基著作,如《国富论》《数字图像处理》《微积分》等;(4)课程辅导类书目,如《复变函数与积分变换学习辅导与习题选解》《偏微分方程习题集》《概率统计复习与习题全解》等;(5)程序设计相关书目,如《MATLAB宝典》《Visual C++开发入门与编程实践》《由浅入深学Java》等;(6)英语水平测试用书,如《TOEFL核心词汇21天突破》《大学英语六级阅读特训》《考研英语阅读理解精读100篇》等。

值得注意的是,在近三年发布的“Top10年度借阅榜单”中,有90%的书目属于该类“经久不衰型”图书,其平均成熟期为7.24年,其中86.66%的书目至今仍处于成熟期。这表明,位于借阅排行榜首的高借阅型图书既具有极高的借阅量,又具有相当长的成熟期。比较历年发布的“Top10年度借阅榜单”,发现有36.67%的高借阅型书目重复出现。其原因可能是,一方面由于这部分高借阅型书目本身仍处于成熟期,其借阅量在未来一段时间仍将保持于峰值附近;另一方面,由于借阅榜单的发布,进一步加深了读者的认知度,使这类图书变得更为抢手。

6.3 “流星”型曲线

少数馆藏的生命周期出现“戛然而止”的现象,其借阅曲线呈现出“流星”型。根据统计,有63种图书的借阅曲线呈现为“流星型”,在其借阅量下降为0的前一年,平均借阅量为12次。调查发现,“流星型”图书均经历了从开架书库变化为提存书库等闭架书库的库位调整,造成了其生命周期的突然衰落。如果这类藏书保持原馆藏地不变,其生命周期可能在未来得到延续。

7 数字资源对纸质馆藏借阅趋势的影响分析

近年来,随着数字阅读的普及,传统的纸质阅读受到了明显的冲击。相较于纸质载体,数字设备存储容量大,能与网络互联互通,在信息检索、信息获取和信息存储方面都更具优势;随着技术发展,数字设备的价格逐渐降低、体积更加轻便、体验不断优化^[25],数字阅读的国民接触率也逐年上升^[26]。已有研究认为,数字资源的广泛使用,使人们的阅读方式和阅读习惯在潜移默化中发生转变,降低了对纸质文献的需求,因此,数字资源是造成纸质馆藏借阅量下降的主要因素^[27]。

然而,数字资源是否只会给纸质馆藏带来负面影响?在研究阅读推广专题网站的书目借阅趋势时,笔者却观测到一种特别的现象。

西安交通大学图书馆的阅读推广专题网站中,提供了推荐书目的纸本馆藏信息;对有数字资源的书目,在版权许可范围内,同时提供在线阅读链接。从使用统计来看,一方面,数字资源的总体使用量的确比纸质馆藏高:推荐书目的数字资源年均访问量为9398次,而纸质馆藏资源年均借阅量为3615册次;另一方面,在数字资源与纸质馆藏同步推荐的模式下,数字资源却能够促进纸本馆藏借阅量的增长:将专题网站的推荐书目分为“仅提供纸质馆藏”和“同时提供纸质馆藏与数字资源”两类,分别计算其借阅曲线在推荐后的变化趋势,发现“提供纸质馆藏”的每种书目在推荐后年借阅量平均增加27.63次,“同时提供纸质馆藏与数字资源”的每种书目在推荐后年借阅量平均增加32.37次;表明,“同时推荐纸质和电子资源”的馆藏借阅量增长幅度比“仅推荐纸质馆藏资源”的借阅量增长幅度更高;在书目推荐场景下,电子资源的同步推荐不仅没有降低纸本馆藏的借阅量,反而能够起到提升作用。

究其原因,一是由于多途径的数字资源推荐模式使读者更容易发现被推荐的书目;二是由于数字资源为读者提供了详细了解图书内容的便捷入口;在线阅读后,如果读者被其内容吸引,则会选择借阅对应的纸本馆藏以便进行深层次的阅读、学习和研究。毕竟,数字阅读虽然在时效性上更胜一筹,但其更适合于浅层次的碎片化阅读;而在阅读体验方面,纸本馆藏所特有的质感和触感具有不可替代的优势,特别适合于深层次、集中性、沉浸式阅读。而当纸本馆藏复本全部借出时,数字资源则充分发挥了



对纸本馆藏重要的补充作用,让更多的读者能够发现有价值的纸质馆藏。从这个角度来看,纸本资源和数字资源之间相互促进、优势互补、相得益彰。

8 高借阅型馆藏生命周期研究的启示

高借阅型馆藏生命周期的研究,为馆藏价值的历时性评估提供了崭新的视角,同时也为图书馆的业务实施策略带来启示。

8.1 实施生命周期监测,评估馆藏“年龄结构”

图书馆可定期进行生命周期监测,评估当前馆藏生命阶段的分布情况。如果处于“老化期”“衰退期”“衰亡期”的馆藏数量大于“萌芽期”“成长期”“成熟期”,则表明当前读者借阅的对象多为处于生命周期后半程的图书,提示当前馆藏的“年龄结构”趋于老龄化,未来借阅量下降的风险将持续增加。图书馆应着力采取应对措施,“培育”更多“上升型”馆藏,以避免“青黄不接”的问题出现。

8.2 充分感知读者需求,调整馆藏建设布局策略

在知识快速更新迭代的过程中,读者的需求也在动态变化,处于上升阶段的馆藏蕴含着读者需求的变化方向。在馆藏建设时,可对“上升型”“先下降后上升型”“睡美人型”图书进行聚类分析,将其中新兴的热门借阅主题词、热门作者信息,纳入馆藏采访的考虑因素中,充分保障读者科研与学习的需要。

对于“经久不衰”型馆藏,表明读者对其的需求持续处于较高水平,在馆藏建设方面,应对这类图书进行侧重配置,并对其复本进行适当增订,增加其借阅量进一步提升的空间。

在进行馆藏库位调整时,也应充分考虑到读者的历史需求,避免造成馆藏生命周期中断的情况。为了减少“流星型”馆藏的数量,应当在清点和剔除操作中更加谨慎,不仅要考虑最后一次借出时间,还应将历史流通数据纳入考虑范畴,减少对高借阅型馆藏的库位调整。

8.3 强化阅读推广服务效能,激发馆藏的生命力

对“睡美人型”借阅曲线的分析表明,沉睡的低借阅型馆藏可以被唤醒;对“先下降后上升型”借阅曲线的分析表明,进入老化阶段的图书也可重新焕发生机。因此,馆藏的生命周期可以被外界因素改变,而阅读推广、书单推荐、借阅榜单等活动都是重要的影响因素。图书馆应当充分利用阅读推广的各种形式,加大阅读推广实施的广度和力度,让更多的

低借阅型馆藏焕发生机。

8.4 从馆藏的生命阶段出发,针对性开展宣传

对处于不同生命阶段的馆藏,应当采用不同的宣传策略。处于上升阶段的馆藏,其影响力尚处于扩散阶段,图书馆应优先加大对它的宣传力度和频率,采取多平台、全方位的主动推送方式,使其能为更多读者所知。对处于“成熟期”的馆藏,表明其已经得到受众群体的认可,宜采用图书借阅榜单、读者书评分享等方式彰显其价值,使其知名度得到进一步提升;对于处于下降阶段的馆藏,其借阅量虽衰减式微,但可采用经典回顾式的宣传,反映其曾在巅峰期发挥过的作用,从而唤起读者的兴趣。

8.5 深入教学实践环节,拓展科研学习支持服务

研究发现,“经久不衰型”图书多集中于学科经典著作、课程辅导教材、复习备考用书、程序设计等方面,表明学生在其课内外的学习、备考、竞赛、研究和实践过程中,对文献信息获取始终具有强烈的需求。高校图书馆应当深入教学实践环节,全面拓展科研学习支持服务。例如,图书馆可针对热门研究主题,如人工智能、数学建模、机器学习等,遴选、整合、汇集优质的馆藏资源,向相应的读者群体进行个性化推介;从读者的周期性需求出发,图书馆可在“开学季”“考试季”“竞赛季”“就业季”等不同时段,根据读者的年级、专业、研究方向,针对性地推送相应主题的馆藏资源。在进行推送时,应适当增加低借阅型馆藏的曝光度,避免“高借阅型图书一本难求,低借阅型图书无人问津”的“马太效应”进一步加剧。

8.6 建设专题阅读网站,以数字资源促进纸质资源借阅量提升

研究发现,在阅读推广专题网站中,数字资源的同步推荐能够促进纸质馆藏借阅量的提升。因此,应当充分发挥数字资源在资源推荐中的优势,结合文化建设的需求,加强阅读推广专题网站建设。在版权许可的范围内,将推荐对象的数字资源和纸质馆藏进行同步推荐,例如,清华大学的“读在清华”专题网站、西安交通大学图书馆的“百本经典”专题网站等。同时,应在微信、微博和各类社交媒体中,进行多途径的专题宣传,增强资源推荐的曝光度,让读者能够更方便地进行资源访问。



9 结语

基于借阅曲线的馆藏生命周期研究为馆藏评估提供了一种崭新的视角,有助于发现馆藏生命周期的影响因素、感知读者需求的变化方向、揭示当前馆藏建设布局的存在问题。图书馆应加大对馆藏生命周期的监测工作,从而有针对性地调整当前馆藏建设与布局的实施策略;同时,在新形势下,应优化馆藏推荐模式,采用多途径的宣传策略,以数字资源促进馆藏资源的揭示,从而实现馆藏利用率的进一步提升。

参考文献

- 冯向春.对高校图书馆纸质文献借阅现状的思考[J].图书馆研究,2015,45(2):49-53.
- 李伟.数字信息时代高校图书馆纸质馆藏的价值与服务对策[J].图书馆情报工作,2017,61(3):79-85.
- 张婧,杜智涛.基于生命周期的电子信息资源评价研究[J].图书馆学研究,2011(21):68-73.
- 王爱芳.文献信息学的基本原理[J].现代情报,2006,26(6):25-26.
- Marchand D A, Horton Jr F W. Infotrends: profiting from your information resources[M]. John Wiley & Sons, Inc., 1986.
- 肖珑,张宇红.电子资源评价指标体系的建立初探[J].大学图书馆学报,2002(3):35-42,91.
- 史敏鸽.利用信息生命周期理论管理图书馆视音频数据[J].情报探索,2006(3):93-95.
- 索传军.基于信息生命周期的数字馆藏管理研究[J].大学图书馆学报,2005(1):26-29.
- 杨晓秋.我国数字图书馆信息资源生命周期管理研究综述[J].图书馆情报工作,2011,55(1):26-30.
- Gosnell C F. Obsolescence of books in college libraries[J]. College & Research Libraries, 1944, 5(2):115-125.
- Brookes B C. Obsolescence of special library periodicals: sampling errors and utility contours[J]. Journal of the American Society for Information Science & Technology, 1970, 21(5):320-329.

- 王素丽.基于学科引文的文献生命周期模型探析[J].情报理论与实践,2012,35(3):37-41.
- 邓奇强,熊熊,郭锋.文献老化在图书馆书目挖掘中的应用研究[J].图书馆情报工作,2013,57(6):91-96.
- 彭跃峰.高校图书馆中文藏书老化的半衰期分析——以华南农业大学图书馆为例[J].农业图书情报学刊,2018,30(7):91-99.
- 段峰,曹南灵.馆藏图书老化实证分析与应用——以山东大学图书馆为例[J].农业图书情报学刊,2015,27(4):67-69.
- 张惠玲.基于书龄分析的图书剔旧实证研究[J].图书馆论坛,2013,33(4):126-129.
- 李江,姜明利,李玥婷.引文曲线的分析框架研究——以诺贝尔奖得主的引文曲线为例[J].中国图书馆学报,2014,40(2):41-49.
- 屈文建,胡志伟,周小渝.面向图情学科热点高被引论文引文曲线特征分析[J].情报杂志,2017,36(8):138-143.
- 龚李莉,李敏,蔡梅,等.基于 Fisher 最优分割法的太湖调度线时段划分研究[J].人民长江,2019,50(4):106-110,172.
- 郭江龙,张晗,李晓光,等.基于 Fisher 最优分割法的汽轮机稳态运行数据提取方法研究及应用[J].汽轮机技术,2015,57(5):393-395.
- 杨永国.数学地质[M].中国矿业大学出版社,2010.
- 孙虹,俞会新.从发展阶段划分看中国经济与发达国家的差距[J].学术交流,2019(9):126-136.
- 边肇祺,张学工.模式识别(第二版)[M].2000.
- 刘文,王庆林,周东华,等.科学文献老化规律的分析与建模[J].情报理论与实践,1996(4):17-20.
- 邹鑫,杨方铭.电子书产品因素与读者需求的关系[J].图书馆论坛,2020,40(7):78-86.
- 余力杨.高校图书馆中文电子书采购策略探究[J].图书馆工作与研究,2020(7):91-95.
- 李伟.数字信息时代高校图书馆纸质馆藏的价值与服务对策[J].图书馆情报工作,2017,61(3):79-85.

作者单位:西安交通大学图书馆,陕西西安,710049

收稿日期:2020年7月7日

修回日期:2021年1月22日

(责任编辑:支娟)

Research on Life Cycle of Highly-circulated Collections Based on Circulation Curve Analysis

Chen Yadi Li Juan Liang Dong Han Meng

Abstract: Highly-circulated collections signify readers' interests thus studying on their features helps university libraries optimizing collection management strategy. Based on the circulation data of Xi'an Jiao-tong University Library from 2008 to 2018, the paper proposes a feature extraction method using Fisher's optimal segmentation to obtain the distribution characteristics of borrowing curves and common features of life cycle. Then it analyzes the mechanism of several special borrowing curves and finds out the influencing factors of their life cycles. Finally, suggestions have been given on reading promotion activities and collection adjustment.

Keywords: Borrowing Curve; Time Series Characteristics; Fisher's Optimal Segmentation; Usage Analysis; Life Cycle