

陈述信

薛静老师：

您好！我是自动化系宋宸韬（学号 2021013525）。

首先我想感谢老师一个学期以来不辞辛劳的付出，让我在写作方面有了很大的突破，同时也感谢老师给予我的自信，让我这个曾经的“社恐”经过一节节研讨课的历练，也逐渐变得敢于跟大家分享自己的见解。

在长文面批过程中，经由老师的悉心指导，我意识到了长文结构上存在的不合理，并调整了内容的布局，并优化了语言表述。如果还有时间的话，我会希望添加更多的样本，来进一步改进研究。

再次诚挚地感谢薛老师！

宋宸韬

2022.01.02

对 bilibili 热门视频生命周期及其影响机制的研究

——基于模糊集定性比较分析(fsQCA)

宋宸韬 2021013525 sct21@mails.tsinghua.edu.cn

【摘要】 本文从微观视角出发,将热门视频个体作为研究对象,于 bilibili 视频平台上抽样得到热门视频样本,基于大量数据挖掘,对热门视频的生命周期进行了聚类,研究发现,热门视频的生命周期可分为两种类型:脉冲型与分布型。同时,运用定性比较分析的方法,发现了决定热门视频生命周期类型的微观影响机制。研究对网络视频的生产创作、视频平台的运营管理均有一定指导价值,对网络视频领域的学术研究也起到一定填补领域空白的作用。

【关键词】 网络热门视频 bilibili 生命周期 影响机制 定性比较分析(QCA)

近年来,我国网络视频行业呈高速增长态势,截至 2021 年 6 月,网络视频用户已占据网民整体的 93.4%^①。作为网络文化的重要组成部分,热门视频既形成了自身独具的文化场域,也有强大的文化生产与输出能力,可以打破文化社区与圈层的屏障,具有不可小觑的研究价值。然而,当前学界对网络视频的研究集中于宏观视角,对视频的传播机制、潮流更迭、商业模式、用户偏好等给予了较为充分的关注,但对单个视频的微观研究尚且欠缺。因此,本文将从微观视角出发,聚焦于每个独立热门视频个体热度的发展变化过程(即本文所探讨的生命周期),根据其特征进行聚类,并探寻背后存在的微观影响机制。

一、样本选取与相关理论

创建于 2009 年的 bilibili(或“B 站”)作为历史较长、近几年发展势头迅猛的视频平台,现已是弹幕视频行业的一家独大。bilibili 在游戏、影视、泛知识类、虚拟偶像等领域拓展纵深,形成了完善的内容生态;同时,其因为拥有同类型平台无可企及的“Z+世代”比例,有更大机会洞察到视频领域未来的发展前景。综上,本文将 bilibili 选取为本次研究热门视频样本的来源平台。

热门视频通常具有很高的播放量与评论量。基于对 B 站的观察,我们将播放量超过 1000 万的视频列入考察范围,在这些视频中抽样选取了知名度高、用户交互密集的 14 个典型样本,它们涵盖了游戏、生活、知识、科技、音乐、影视、动画、鬼畜等 B 站全部的主要分区,具有较高的代表性。如下表所示:

表 1 本文选取的热门视频样本

^① 中国互联网络信息中心:第 48 次《中国互联网络发展状况统计报告》, http://www.cnnic.net.cn/hlwfzyj/hlwzbg/hlwtjbg/202109/t20210915_71543.htm, 2021 年 11 月 21 日。

编号	视频内容	分区
1	何同学：5G 测评	科技
2	《黑神话：悟空》实机演示	游戏
3	改革春风吹满地	鬼畜
4	罗翔：谈谈网络喷子	知识
5	【猛男版】新宝岛	舞蹈
6	王冰冰 Vlog	生活
7	【古筝】千本樱	音乐
8	影流之主	生活
9	老番茄：Hitman 游戏实况	游戏
10	观察者网：中印为何不能开战	知识
11	雀巢：福尔摩斯动画配音	动画
12	二战电影镜头混剪	影视
13	史上最棒的炸鸡	美食
14	现代圣杯战争	生活

网络视频生命周期的定义尚未给出，不过学界存在与之相近的产品生命周期。即“产品从准备进入市场开始到退出市场为止的全部运动过程”^①。对应到网络视频领域，我们给出定义：网络视频生命周期指视频从发布到互联网平台开始，到基本淡出用户讨论话语场为止的热度变化过程。

产品生命周期理论最早由雷蒙德·弗农在 1966 年提出^②。弗农认为，产品和生物一样，其寿命可划分为引入期，成长期，成熟期，衰退期四个时期。这就是经典的 S 型曲线，（图 1）。后世学者在其模型基础上补充了一些特殊曲线，但不外乎对四个时期的趋势进行一些调整。

^①<https://baike.baidu.com/item/%E4%BA%A7%E5%93%81%E7%94%9F%E5%91%BD%E5%91%A8%E6%9C%9F/2233724?fr=aladdin>, 2021 年 11 月 22 日。

^② Raymond Vernon, *International Investment and International Trade in the Product Cycle*, *The Quarterly Journal of Economics*, 1966(5), Vol. 80, No. 2, pp. 190-207.

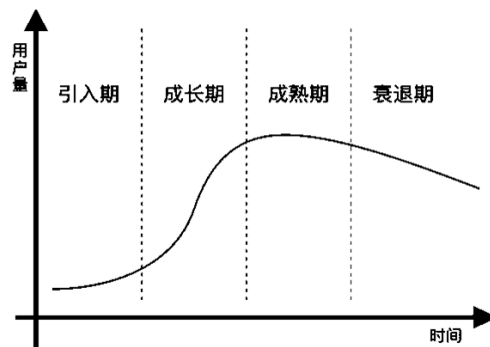


图 1

在流量巨大、形态多样的互联网视频领域，传统意义上的这四个时期是否还能契合热门视频的热度生命周期规律？如果不能，热门视频的生命周期又呈现怎样的曲线形态，可以分为何种类型呢？在这些类型背后，又是什么机制促成的呢？这些都是本文将要探究的问题。

二、研究设计与数据来源

在研究设计方面，首先需要对视频热度进行描述。根据既有研究，在互联网视频描述中，包含视频标签、上传者、上传时间、观看次数、评论次数等，这些称为视频的元信息(Metadata)^①。通过元信息的组合可以描述视频的热度。考虑到一个文化场域的生命力在于场域内个体之间的交互，热门视频生产、输出文化能力也依赖于用户互动的密度，本文主要考虑热门视频的用户交互热度。

因此，我们采用元信息中视频评论次数、各个评论获点赞总和的组合来描述视频的热度。即交互热度指数 H_i 可记为： $H_i = \text{Comment} + w * \sum \text{Like}$

其中 Comment 代表一段时期内发表的评论数， Like 代表这些评论所获的点赞数， w 为权重系数。根据 B 站评论平均点赞数的统计，将其取值为 0.024^②。

在此基础上，本研究对上文选取的 14 个视频样本进行了数据统计，时间跨度始于 2014 年 7 月 3 日，统一截止至 2021 年 11 月 20 日。研究应用 python 语言，对视频的公开评论信息进行数据挖掘，获得了历史上每个时间段内的评论数量与点赞总和，进而计算出各时间段内视频的热度指数。据此，便可绘制出视频样本热度随时间变化的散点图，也即视频的生命周期曲线。

三、主要研究发现

经过对上文获得的生命周期曲线进行观察与聚类，结果表明：在 14 个视频案例中，生命周期的类型较为鲜明地分为两类。根据其特征的不同，可以分别命名为脉冲型与分布型。

^① 王全杰：《互联网热点视频分析方法研究》，《广播电视信息》，2014 年第 8 期。

^② B 站评论平均点赞数据统计为 41.5，为使点赞因素与评论数因素所占比例相当，令 $w=1/41.5=0.024$

（一）脉冲型

此类视频的典型特征是发布之始具有很高的热度，但在很短的时间内热度骤降，在热潮之后立即进入持续的沉寂期，此时热度甚至低于非热门的普通视频。定量地来看，我们将这类视频刻画为：发布前三天内的热度占据总计热度的 50%以上，即归为脉冲型曲线。

依照此标准，14 个案例中有 8 个归为此种类型。其编号为 1、2、4、6、9、10、11、13，图 2 列出了其中三个样本的生命周期曲线。

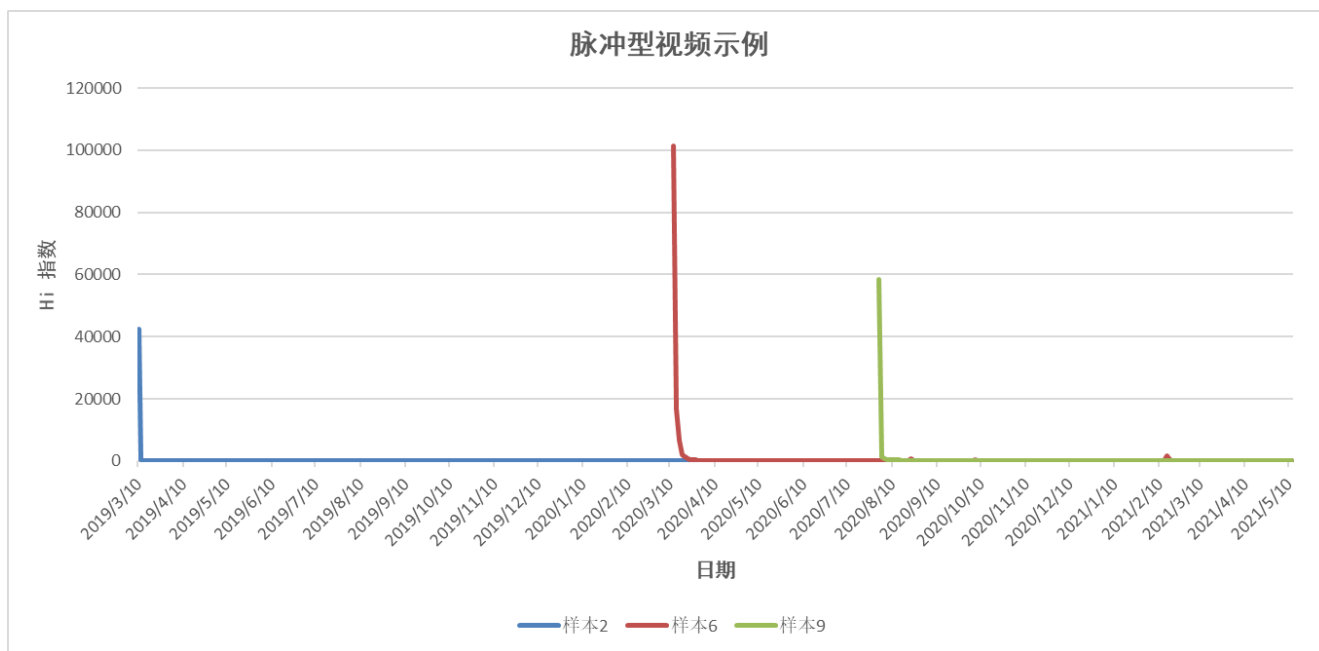


图 2

（二）分布型

此类视频的典型特征是热度存在多个高峰，而不仅只有发布时一次热潮，并且视频的热度较为分散地分布在各个时段，无固定分布规律。定量上，不属于脉冲型的视频，即热度并不集中于发布前几天内的视频，我们归为分布型曲线。

按照此标准，样本中有 6 个对象属于此类，分别为 3、5、7、8、12、14 号。因篇幅原因，下面同样展示其中三个的生命周期曲线（图 3-图 5）。

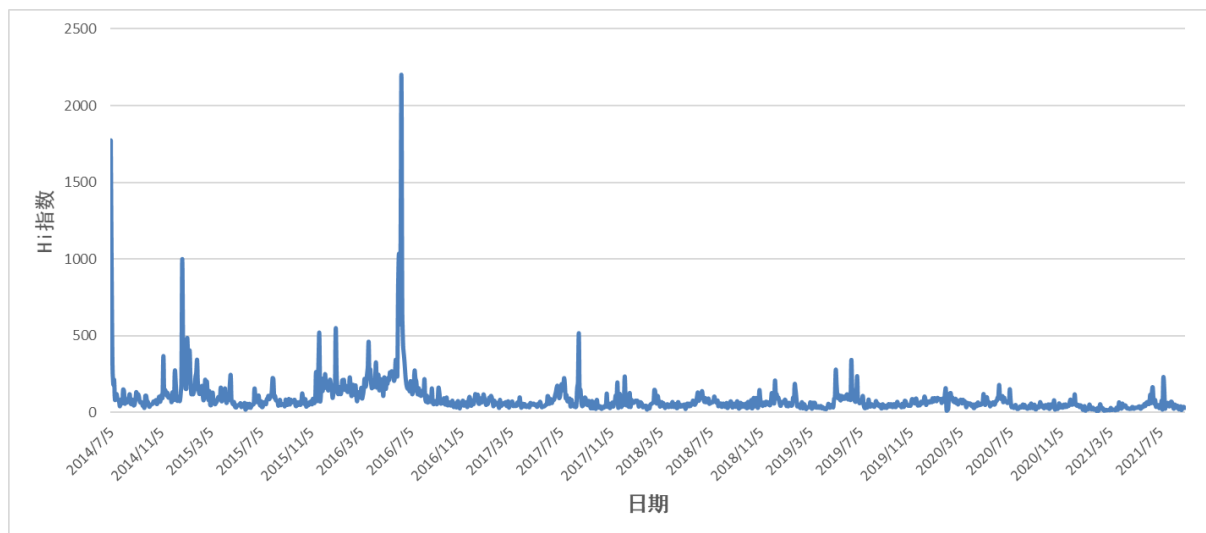


图3 【古筝】千本櫻

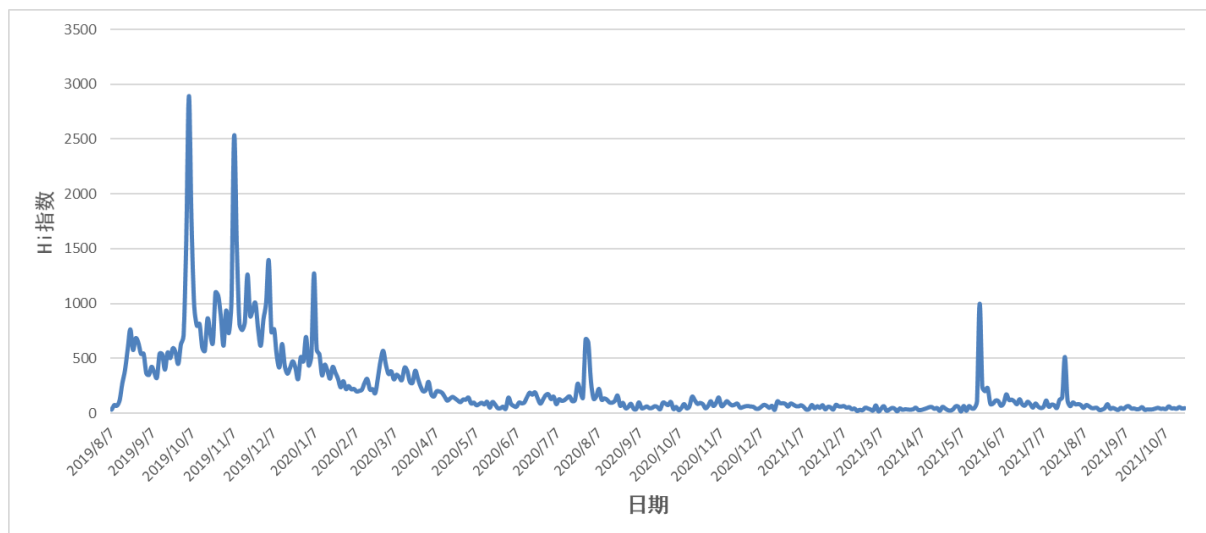


图4 影流之主

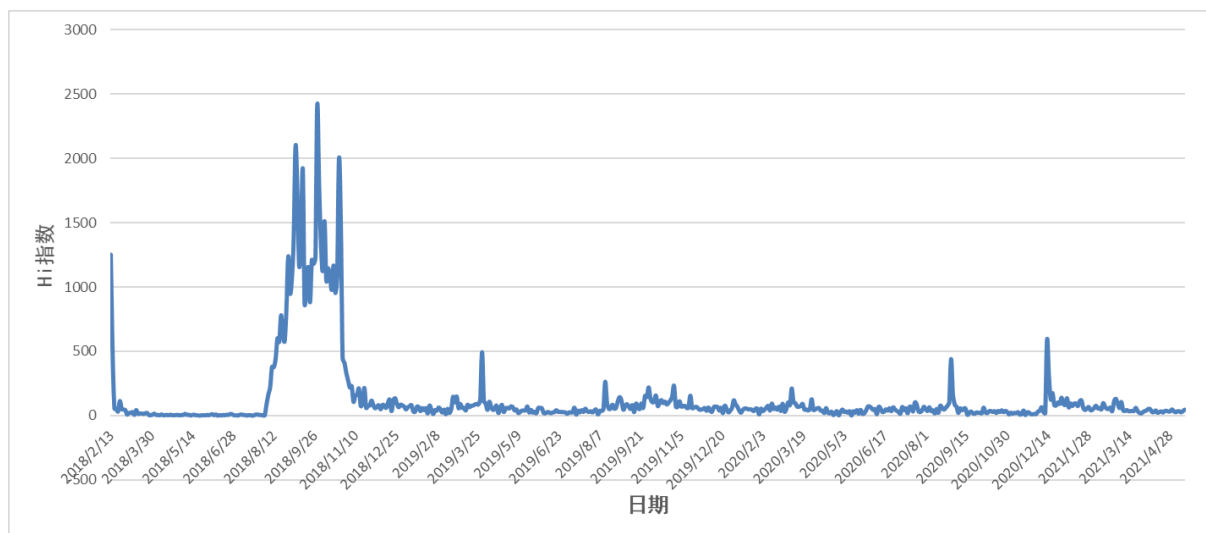


图5 改革春风吹满地（赵本山鬼畜）

从上面的分类与曲线分析中，我们可以总结出：传统的生命周期理论已经不能解释现今

互联网视频的热度变化规律。

一方面,脉冲型的视频不再具备引入期、成长期、成熟期与平滑的衰退期,而是瞬间便失去热度,因此生命周期可划分为骤降期与沉寂期。“发布即巅峰”的原因可能是由于现今互联网视频平台的强大推送效率及巨大的用户流量,新视频在发布伊始即可能获得大量用户的关注;同时,由于视频平台的用户兴趣点迁移迅速,且大量的新视频持续性产出,该视频也可能迅速失去热度。

另一方面,分布型的视频亦是如此,传统的四个周期不复存在,视频热度在高峰期与相对的低潮期之间随机波动。这可能是因为视频平台存在多种因素不可预知的相互作用,视频热度因此随机变化。例如,在用户就视频内容进行讨论时,可能会对视频要素进行重新整合,或引入外来元素,从而创造出被广泛接受的新的观点、议题或“热梗”,进而赋予视频新的生命力,引发新一轮的热度高峰。

正是因为视频平台多种因素的交织,想要探寻背后是何种机制使得视频热度呈现出不同的类型,便成为较困难的工作。为了寻找生命周期的影响机制,我们需要引入科学的研究方法,即模糊集定性比较分析。

四、研究方法简述与变量设计

模糊集定性比较分析(fuzzy-set Qualitative Comparative Analysis)是一种案例导向的研究方法。它以集合论与布尔代数为支撑,被用来考察社会现象是“如何”发生的,从样本中得出有解释力的一般性结论。^①在样本较小和因果关系复杂的情况下,模糊集定性比较分析尤其适用,因此符合本研究的应用场景。

由于模糊集定性比较分析运用布尔代数进行逻辑分析,需要将研究涉及到的因素设计为二分取值的变量^②。具体地,定性比较分析的流程为:选取样本——设置变量——变量编码——单变量必要性分析——多变量组态分析。其中样本选取已在上文完成,下面进行变量的设置,共五个解释变量与一个结果变量。

(一) 解释变量的设计

1. 聚焦内容变量

该变量代表视频聚焦的主要内容,每个视频都具有不同的内容指向,在本研究中,我们主要关注视频的主要内容是否明显指向具体的社会热点事件。如样本1:“何同学:5G测评”

^① 李蔚,何海兵:《定性比较分析方法的研究逻辑及其应用》,《上海行政学院学报》,2015年第16卷第5期。

^② 即取值为0或1的变量

具有明显的内容指向，即对公众关心的中国 5G 建设热潮进行了回应与解读。而样本 9：“老番茄：Hitman 游戏实况”则不具有此种属性。本研究将视频聚焦内容设为有影响力的解释变量之一。

2. 整体风格变量

该变量代表视频呈现出的整体风格，在本研究中，主要关注视频的整体风格是否戏谑夸张，带有网络狂欢的属性。从狂欢理论的视域来看，近年来很多引发受众共鸣的潮流文化，都带有网络狂欢的色彩，包括各种鬼畜剪辑、二次创作、流行热梗等，例如，对刘华强、马保国等形象的鬼畜，原神、明日方舟等的二创、“雨女无瓜”等热梗。B 站视频正是这些狂欢文化的重要源流之一。因此，我们把视频整体风格是否带有网络狂欢的属性纳入解释变量范围。

3. 发布主体变量

该变量代表视频发布者的身份属性，本研究中，我们主要关注发布主体是否带有明显的偶像身份属性。例如在此次样本中，影响力大、有明显的偶像身份的主体有王冰冰、何同学、罗翔等。

在 B 站，视频的发布者被称作 UP 主。而 B 站 89% 的播放量^①来自 UP 主创作的 PUGV^②内容。这也是很多视频平台的共同特点：视频的发布主体在视频的传播中扮演着重要角色。因为视频平台的关注功能与推荐机制，影响力大的发布主体创作的视频可能会引发偶像效应，为视频吸引到更多的流量，因此发布者的身份属性也是需要考虑的因素。

4. 用户情绪变量

该变量代表视频的讨论话语场内用户的整体情绪偏向。研究运用 SnowNLP 语言处理库，对总计 819,741 条评论的文本内容进行了情感分析并按积极/消极程度进行评分。用户参与讨论的情绪对视频热度是一个非常重要的影响因素，积极的情绪氛围会引发更多用户的情感共鸣与积极反馈，但同时消极的情绪氛围也能通过制造更多争议话题来增进视频的讨论热度。因此研究将用户的情绪偏向纳入解释变量的范围，希望探寻其对视频热度的具体影响效果。

5. 音乐元素变量

该变量代表视频是否以音乐为主体创作元素，此变量的设置在视频研究领域尚属创新之

^① 动点科技：《PUGV、OGV 两条腿走路，看十周年后的 B 站如何破圈不破功》，<https://cn.technode.com/post/newsnow/allbebil/>，2021 年 11 月 22 日。

^② Professional User Generated Video，专业用户生成视频

举。在选取视频样本时，笔者注意到，许多以音乐为信息载体，或创作对象即为音乐的视频获得了极高热度，且很多此类视频属于分布型生命周期。如样本中“改革春风吹满地”、电影配乐混剪等。出于此种考虑，我们将视频是否以音乐为主体创作元素列入解释变量范畴。

（二）结果变量的设计

根据前文研究得到的结果与本研究目的，我们将结果变量设置为视频生命周期的类型，即脉冲型或分布型。

五、定性比较分析与结论

（一）变量编码

表 2 变量编码表

变量 类型	变量	解释说明	变量表示	赋值
解释 变量	聚焦内容	内容指向具体热点事件	Content	1
		内容无具体热点事件指向	~Content	0
	整体风格	整体风格为戏谑夸张	Style	1
		整体风格非戏谑夸张	~Style	0
	发布主体	发布者具有明显偶像属性	Individual	1
		发布者不具有明显偶像属性	~Individual	0
	用户情绪	积极情绪主导	Sentiments	1
		消极情绪主导	~Sentiments	0
	音乐元素	以音乐为主体创作元素	Music	1
		不以音乐为主体创作元素	~Music	0
结果 变量	周期类型	脉冲型	Type	1
		分布型	~Type	0

注：具体赋值时，数值表示隶属度，越接近 1，隶属度越高，反之越低。最终对数据进行校准以保证分析准确性。^①

（二）单变量必要性分析

^① 单晓红, 崔风艳, 刘晓燕, 王春稳, 戴莹:《组态视角下网络舆情衍生话题形成路径研究——基于微博数据的模糊集定性比较分析》,《情报理论与实践》录用定稿, <http://kns.cnki.net/kcms/detail/11.1762.G3.20211103.1443.003.html>, 2021 年 11 月 26 日。

在进行多变量分析之前，需要进行单变量必要性分析，以确定变量中是否含有导致结果变量的必要条件。一致性（Consistency）和覆盖率（Coverage）可视为变量对结果变量的解释程度。使用 fsQCA3.0 软件运行结果如下：

	Consistency	Coverage
Content	0.375000	1.000000
Style	0.250000	0.333333
Individual	0.687500	0.846154
Sentiments	0.610000	0.561565
Music	0.000000	0.000000

图 8 各变量对【脉冲型】的解释程度

	Consistency	Coverage
Content	0.000000	0.000000
Style	0.666667	0.666667
Individual	0.166667	0.153846
Sentiments	0.635000	0.438435
Music	0.666667	1.000000

图 6 各变量对【分布型】的解释程度

根据既有研究^①，一致性大于 0.9 的变量，可看作结果的必要条件。由图可知，变量一致性均小于 0.9，这说明单一变量无法解释生命周期类型的形成，生命周期类型是多因素共同作用的结果。

（三）多变量组态分析

使用软件对样本进行多变量分析，可得到复杂解、中间解和简单解三种原因组合。本文借鉴前人研究基础，重点关注中间解。^②软件运行结果如下，其中 raw coverage 为覆盖率，即该组合能解释的案例占总体的比例，反映组合的解释程度强弱。覆盖率最高的组合即为最典型的组合。

	raw coverage	unique coverage	consistency
Content*~Style*Sentiments*~Music	0.2325	0.0775	1
~Style*Individual*Sentiments*~Music	0.35875	0.20375	1
solution coverage: 0.43625			
solution consistency: 1			

图 7 脉冲型结果的多变量分析^③

^① 周俊,王敏:《网络流行语传播的微观影响机制研究——基于 12 例公共事件的清晰集定性比较分析》,《国际新闻界》2016 年第 38 卷第 4 期。
^② 单晓红,崔凤艳,刘晓燕,王春稳,戴莹:《组态视角下网络舆情衍生活题形成路径研究——基于微博数据的模糊集定性比较分析》,《情报理论与实践》录用定稿, <http://kns.cnki.net/kcms/detail/11.1762.G3.20211103.1443.003.html>, 2021 年 11 月 26 日。
^③ 结果中*代表“与”符号,即条件同时成立

	raw coverage	unique coverage	consistency
~Content*~Individual*Sentiments*Music	0.468333	0.468333	1
solution coverage: 0.468333			
solution consistency: 1			

图 8 分布型结果的多变量分析

（四）结论及解读

通过对多变量分析的结果进行总结，我们得到以下结论。

① 对脉冲型结果而言，覆盖率最高的典型组合为：

~Style*Individual*Sentiments*~Music，含义为：视频风格非戏谑夸张，发布者带有偶像属性，用户情绪积极，且不以音乐为主体元素。

② 对分布型结果而言，覆盖率最高的典型组合为：

~Content*~Individual*Sentiments*Music，含义为：视频不聚焦于具体热点事件，发布者不具偶像属性，用户情绪积极，且以音乐为主体元素。

基于此结论，可以发现两种类型成因的共同之处为：用户情绪反馈积极。这对大众平时所见的各类靠引发争议、挑战底线等消极元素来博人眼球的一些“网红”案例提出了挑战。该现象说明，靠消极内容来获得热度只能是个例，而非长久之计。

而相比较来看，脉冲型视频独具的成因为：视频风格非戏谑夸张，且发布者带有偶像属性。

视频风格戏谑夸张更易获得持续热度，这可以从勒庞《乌合之众》中“群体情感的非理性”得到解释，传播学理论中的集合行为对此也有所论述：当个体处于某个具有情绪和观念感染力的群体中时，其理性就会淡化，“无意识”地参与进集合中信息的传播。此类“玩梗文化”在 B 站更是较为显著。正如平台营销部总经理王旭言，“这套基于认可的共创机制，是突破圈层的抓手”。经由用户群体的狂欢行为，视频的影响力可得到极大扩散。而发布者的偶像效应可能会使视频偏离内容导向，导致热度与视频实际质量不相符，这会对视频后续的热度持续性维持产生负面影响。

分布型视频独具的成因则为：视频不聚焦于具体热点事件，且以音乐为主体元素。

聚焦热点事件的视频在初始阶段即具有一定用户关注度的基础，因此较易获得很高的脉冲热度；而由于热点事件较强的时效性，在公众关注消退后视频热度也会随之回落。因此不利于持续性热度的形成。而对于音乐元素，尼尔·波兹曼在著作《娱乐至死》中认为，任何通过电视和网络媒介传播的信息都需要以娱乐的面貌出现，而应避开完整且清晰的文字阐

述。“不能存在任何需要记忆、学习、运用甚至忍受的东西……因为最重要的是受众的满意程度，而不是受众的成长。”^①这也与现代传播学中的“使用与满足”理论观点相合。在视频平台生产过剩，选择繁多的情况下，受众的浏览行为是基于个人需求进行的，形式简单，且能够让用户感官上得到满足的音乐形式自然能够拥有更多的流量，并且更能保持旺盛的生命力。正是因为如此，很多广告才设计为音乐的形式，以期获得受众持续性的兴趣。

六、结语

区别于以往视频领域宏观层面研究，本文深入地对热门视频个体进行了分析，选取了较为新颖流行的材料，发现热门视频的生命周期类型可分为脉冲型与分布型。运用科学研究方法，总结出决定热门视频生命周期类型最为典型的机制。

须要指出的是，受笔者能力、资源所限，研究或存在潜在的不甚周全的因素。解释变量的设计受视野局限性，可能存在其他影响变量未纳入。此外，样本选取也受到一定时间、精力限制。以上未解决的问题，都有待在后续的相关研究中进一步深化。^②

^① Neil Postman, *Amusing Ourselves to Death*, Penguin Books, 1986, pp. 170-186.

^② 周俊, 王敏:《网络流行语传播的微观影响机制研究——基于 12 例公共事件的清晰集定性比较分析》,《国际新闻界》2016 年第 38 卷第 4 期。