

# 互联网行为大数据应用于品牌测量的探索

兰颖<sup>1</sup> 北京数字新思科技有限公司

谭北平 北京数字新思科技有限公司

**摘要** 互联网改变了消费者与品牌的沟通方式，大数据技术的发展，使得这些沟通行为可以被测量。在互联网上，消费者与品牌的接触方式有很多种，但这种大数据与传统的品牌关键指标关联度如何，哪些指标更有效，这是大数据应用于品牌测量领域的关键问题。本研究基于大数据采集了近 200 个品牌的 30 多种消费者网络行为大数据，并且与对应的 200 个品牌的基于线下调查的品牌发展数据库进行比较。研究发现：消费者行为大数据可以分为四大类联结行为：内容、需求、行动与互动，并可以汇总为统一的联结指数。品类发展不均衡，各品类的联结指数差异较大。应用统计方法在消除品类差异后，互联网行为大数据的品牌指标与传统调查指标显著相关。这个研究提示我们，大数据指标确实能够测量品牌的表现，因为大数据更加敏捷、细节和准确，能够帮助营销实现基于数据驱动的革新。

**关键词** 品牌联结资产，互联网行为，多层次线性模型，大数据，品牌测量

## 1. 前言

我们在互联网上浏览信息，搜索内容，购买产品，分享见解.....互联网的发展，改变了我们的生活方式。

互联网也改变了消费者与品牌的沟通方式。消费者购买一个品牌，不再依赖于头脑中的封闭式的大量品牌的筛选，而是转化为开放性的、依靠更多信息的品牌评估<sup>1</sup>。这也让我们重新思考，在数字时代，品牌应该如何管理<sup>2</sup>。Chen<sup>3</sup>和Muller<sup>4</sup>的研究表明，网站对企业的品牌管理有着重要的作用，它是加强品牌属性（brand attribute）的重要工具。在 2002 年，Page & Lepkowska-White<sup>5</sup>首次提出

---

<sup>1</sup> 兰颖 就职于北京数字新思科技有限公司

Email: [lanying@xinsight.cn](mailto:lanying@xinsight.cn)

Tel: 18910280756

网站品牌资产的概念，以传统的方法构建了以网站认识和网站形象为中心的网络品牌资产的理论框架，帮助企业通过加强网站的管理进而增强品牌的管理。但 David Edelman<sup>1</sup> 认为，在数字时代用传统的指标、方式来管理品牌是不合时宜的，大部分企业在互联网上对品牌的花费用错了地方。

大数据的应用，给互联网上品牌的管理带来了新契机。可是当前，通过大数据从事最多的是关于社交媒体在大数据领域的研究。一些学者发现，社交媒体的应用能够增加品牌的销量<sup>67</sup>，然而，这些研究主要集中于特定的行业中，并不能广泛适用于所有品牌。虽然在数字化变革中，数字化触点的管理将是企业致胜的法宝<sup>8</sup>。可实际上，在数字时代，企业对如何全面的管理品牌，且消费者与品牌是如何相互影响和作用的，都缺乏清晰的了解和认识，也缺乏简便有效的方法。

## 2. 研究假设：

本研究将着重探索在数字时代，运用互联网行为大数据对品牌进行测量的方法，且以期此方法可以广泛应用于各行业中。

越来越多的企业已经意识到互联网对品牌的改变，也在努力寻找一些指标以测量品牌在互联网上的现状，如使用官网访问量、百度指数、微博、电商等指标数据。然而，这些指标多而庞杂，更重要的是，它们显示的仅仅是品牌在市场表现的一个个独立的点，无法系统地显示出品牌全貌。因此，在此基础上提出**假设 1：互联网行为大数据可以进行分类，并能够提炼出综合指标测量品牌的整体表现。**

互联网已经成为消费者与品牌发生关系的重要渠道，且不同品类和品牌对互联网的使用方式不同。如网上银行是网购过程中必不可少的支付环节；消费者习惯于在超市里挑选饮料，即时饮用；消费者在互联网上查看汽车的信息，分享经验，但倾向于在 4S 店里进行购买。基于此，本文提出**假设 2：不同品类在互联网上的发展阶段不同，且品类间的品牌表现具有显著性差异。**

品牌在互联网上的表现仅仅是品牌在市场中表现的一部分，品牌在线上的表现是否也是其在现实世界的反映，是企业非常关注的问题。基于此，本文将以传统的调查数据作为效标，验证互联网品牌大数据的现实意义。因此，提出**假设 3：互联网行为大数据的品牌指标与传统调查指标品牌力显著相关。**

然而，不论在线下还是在线上，消费者的行为习惯会因品类存在较大差异。例如我们可能每隔一个月需要购买护发用品，而更换一部手机却要一年之久，因此，行业的差别是大数据指标现实意义验证中需要关注的因素。基于此，提出假设 4：品牌所属的品类，会显著性影响品牌在互联网上的表现与品牌力的关系。

### 3. 数据采集与分析

#### 3.1 大数据的采集

本研究所采用的互联网行为大数据，覆盖超过中国 1 亿网民用户。同时为保证研究的广泛性和适用性，在研究中，本文将涉及与消费者日常生活息息相关的 20 个品类，如服装、汽车、手机、快餐、护发等，包括两百余个国内外品牌。

其中，这些数字触点由两类数据组成。一种是基于消费者在一系列相关网站或页面的浏览数量。将每个网站或页面的 URL 作为统计依据，计算在一定时间范围内，每一个 URL 的浏览次数之和。另一种是基于消费者和品牌在互联网上发表或转发文章数量的统计。在一定的时间范围内，只要全文中包含特定关键词，则记为一篇。这些数据基本覆盖了消费者与品牌发生关系的各触点以及触及方式。

为便于统计整理，将这些触点按照消费者的决策路径划分为四个大维度：内容联结，需求联结，行动联结和互动联结，如表 1 所示。其中，内容联结为消费者对品牌官网、官方论坛等品牌相关内容的访问；需求联结为消费者主动搜索品牌；行动联结为消费者在电商搜索或者浏览品牌电商页；互动联结为消费者在社交媒体等互动平台上发布品牌的相关内容。

表 1 品牌联结指标结构

指标	指标说明
内容联结	官网：各品牌官网所有页面的浏览次数
	官方论坛：各品牌官方论坛所有页面的浏览次数
	垂直网站：某垂直网站上某品牌所有产品信息页面的浏览次数
	问答网站：搜索某品牌关键词的前十个页面的相关链接的浏览次数
	百科类：百科类网站中某品牌或公司的定义页面的浏览次数
需求联结	搜索引擎：在某搜索引擎上搜索某品牌关键词页面的浏览次数
	问答网站搜索：在问类答网站上搜索某品牌关键词页面的浏览次数
行动联结	垂直电商搜索：在某垂直电商首页搜索品牌关键词页面的浏览次数
	官方旗舰店：某品牌在天猫官方旗舰店的所有页面的浏览次数

---

	官方商城：某品牌在官方商城的所有页面的浏览次数
	综合电商搜索：某综合电商首页搜索品牌关键词页面的浏览次数
互动联结	社交类网站：在 SNS/博客/电子报/论坛/视频/微博/微信/问答/新闻等网站或页面上的帖量

---

### 3.2 线下品牌数据源

本文将全球知名传播服务集团 WPP 旗下市场调研机构华通明略 (MillwardBrown) 推出的 BrandZ™ 数据作校验，其每年推出的“最具价值品牌排行榜”在全球范围内具有广泛的影响。BrandZ™ 是通过传统市场调研中抽样访问的方法，访问数百万的消费者得到的品牌资产数据。本文所引用的数据是 2015 年全球品牌资产榜中品牌在中国的数据。

其中，品牌力是品牌资产重要的组成部分<sup>9</sup>，具有强劲品牌力的品牌，在市场上可以迅速成长，从而把品牌资产与品牌成长战略相联系起来<sup>10</sup>。另外，从 BrandZ™ 数据中很容易的分析出，品牌力与品牌排名具有较高的相关性 (0.62)。因此，在后文的分析中，选取品牌力作为衡量品牌线下表现的指标，探索互联网上品牌表现的指标与品牌力的关系。

### 3.3 分析方法

在前期的数据准备中，将互联网行为大数据所涉及的品类和品牌与 BrandZ™ 一一对应，剔除缺失值。最终，可用于研究的品类有 15 个，分别为：服饰类、银行类、啤酒类、汽车类、奶制品类、家电类、护肤品类、快餐类、食品类、护发类、医药类、婴儿食品类、保险类、手机类、软饮料类、以及运动品类，涵盖围绕消费者生活的品牌 109 个。

本文研究过程中，将按照以下方法对假设进行验证。

#### 第一步：主成分分析

与传统调查方法中品牌影响力的指标一样，品牌大数据指标并不能简单地由各级指标加总而成，需要考虑各指标间的差异及相互作用，本文将采用主成分法确定品牌大数据指标各二级指标系数。在这一阶段，使用互联网行为大数据中采集到的 192 个品牌，超过 1,000,000 触点的数据。

#### 第二步：单因素方差分析

品牌所处的行业和品类的特性会使得品牌在互联网上的发展呈现差异性,进而影响品牌在互联网上的表现,因此,本文将通过单因素方差分析的方法来进行验证。在这一阶段,本文将采用上一步分析中,通过 192 个品牌总结出的品牌大数据指标。

### 第三步: HLM 多层线性模型

考虑到品牌嵌套于不同的品类和行业中,而这种嵌套关系的存在,难以满足品牌间随机误差独立性的假设条件,因此,传统的回归分析无法应用于此研究中。而传统的市场研究中,一般也只局限于同品类品牌的比较,无法找到合适的方法来解决这一问题。因此,本文将借用社会学研究的方法,采用多层线性模型(HLM)来研究影响品牌大数据指标与品牌力的关系。HLM 的优势在于充分考虑了数据分层的特点,通过建立多层回归方程组,将总误差分解为各层次的误差,解决了随机误差独立性假设违反的问题<sup>11</sup>(刘红云,孟庆茂,2002),以探讨不同层面自变量对因变量的影响。在这一阶段,将使用经过比对整理的 15 个品类,109 个品牌的大数据信息。

## 4. 分析结果

### 4.1 互联网的不同联结之间高度相关

在研究中发现,处于不同联结频次的品类的主要联结点有所不同,如图 1。高频品类对官方资源的控制力较强,善于利用官网、官方论坛、官方商城等与消费者联结。百科类网站对中低频品类有着重要的作用,是与消费者沟通的主要渠道之一。因此,通过了解品类在互联网上的发展状态,可以明确品牌自身的优劣势,采取差异化策略。

	低频品类 -	中频品类 +	高频品类 ++	
内容联结	<ul style="list-style-type: none"> <li>• 官网</li> <li>• 问答网坛</li> <li>• 百科类</li> </ul>	<ul style="list-style-type: none"> <li>• 官网</li> <li>• 官方论坛</li> <li>• 百科类</li> </ul>	<ul style="list-style-type: none"> <li>• 官网</li> <li>• 垂直网站</li> <li>• 官方论坛</li> </ul>	++
行动联结	<ul style="list-style-type: none"> <li>• 官方旗舰店</li> <li>• 综合电商</li> <li>• 垂直电商</li> </ul>	<ul style="list-style-type: none"> <li>• 官方旗舰店</li> <li>• 综合电商</li> <li>• 垂直电商</li> </ul>	<ul style="list-style-type: none"> <li>• 官网商城</li> <li>• 官方旗舰店</li> <li>• 综合电商</li> </ul>	+ -

图 1 联结频次与联结重要性呈现差异

同时，在分析中，我们也看到，品牌的各个指标并不是独立存在的，他们相互影响作用，构成了品牌在互联网上生存的生态圈。因此，需要找到一个指数，用于测量品牌在互联网上的整体表现。本文采用主成分法，得出以下分析结果：

表2 主成分法分析结果

成份	初始特征值			提取平方和载入		
	合计	方差的 %	累积 %	合计	方差的 %	累积 %
1	2.521	63.013	63.013	2.521	63.013	63.013
2	.965	24.127	87.141			
3	.318	7.959	95.100			
4	.196	4.900	100.000			

\*提取方法：主成份分析。

KMO值大于0.7，显著性为P=0.000，且提炼出的一个因子对指标都的解释率达到63%，说明可以通过主成分法确立指标系数。因此，互联网行为大数据可以进行分类，并可以提炼出一个综合指标，用于测量品牌通过各个触点与消费者之间的关系，假设1证明成立。由于品牌与消费者关系的本质在于联结，在后面的研究中，我们将这一指标定义为“品牌联结资产（BrandCQ, Brand Connection Quotient）”，表示每万人每周消费者与品牌的联结指数。下图为品牌联结资产指数与各指标的关系，品牌加强各触点与消费者的联结，有助于提高品牌在互联网上的表现，即增加品牌联结资产。

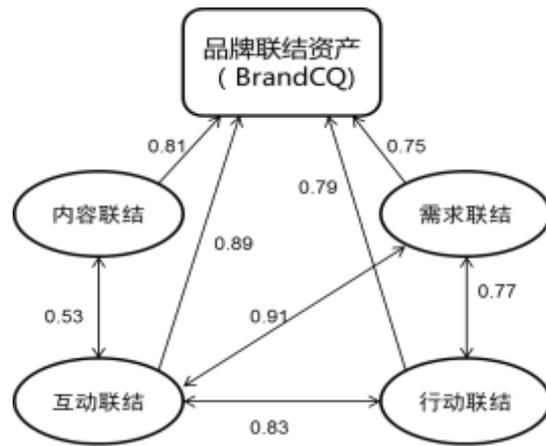


图2 品牌联结资产与各指标关系

## 4.2 互联网的数字联结品类间差异很大

虽然在互联网上，各品牌与消费者关系的本质都是联结，都可以用统一的指标进行测量，但由于品类的物理特性，功能特性等差异，各品类对互联网的利用方式和程度不同，其所属品牌在互联网上的表现也会不同。本文通过单因素方差检验，也证实了这一点，如表3：

表3 单因素方差分析检验

	平方和	Df	均方	F	显著性
组间	1.32 x 10 <sup>8</sup>	26	5.08 x 10 <sup>6</sup>	17.128	.000
组内	4.90 x 10 <sup>7</sup>	165	2.96 x 10 <sup>5</sup>		
总数	1.81 x 10 <sup>8</sup>	191			

单因素方差检验的结果显示，品类间存在着显著性差异（P=0.000），这说明，假设2成立，不同品类在互联网上的发展阶段不同，且品类间的品牌表现具有显著性差异。

在进一步的分析中，我们可以看到（如图3），银行、手机、汽车等与人们每天生活都密不可分的品类的数字化程度越高，消费者与这些品类每个月的联结占到所有联结的95%左右；而面部护理，婴儿纸尿裤，奶制品，计生用品等品类处于中等数字化水平，每个月消费者与这些品类的联结占到所有联结的4%左右；而婴儿奶粉、软饮料等品类，在互联网上的数字化程度较低，消费者与这些品类的联结仅有1%左右。可见，消费者与品牌的亲密关系直接体现在联结上。



固定效应	回归系数	标准误	t 检验	自由度	P 值
截距	316	151	2.09	14	0.055
品牌力斜率	3.00	1.48	2.03	105	0.045

表 6 模型 2 固定部分参数估计结果

随机效应	标准差	方差	自由度	卡方	P 值
截距	583	340,341	14	148	0.000
第一层	542	294,044			

由固定部分的参数可见，品牌力与品牌联结资产呈现显著上升的趋势（回归系数 3.00，p 值<0.05），即品牌力越强的品牌，其品牌联结资产也越高，因此，假设 3 成立，品牌在互联网上的表现指标——品牌联结资产对品牌力具有显著性影响。但从随机部分的参数估计结果可以看出，品牌个体之间存在显著性差异，于是加入品类品牌联结资产这一变量，得到完全模型，输出结果如下：

$$\text{层 1: 品牌联结资产} = B0 + B1 * (\text{品牌力}) + R$$

$$\text{层 2: } B0 = G00 + G01 * (\text{品类联结资产}) + U0$$

$$B1 = G10$$

表 7 完全模型 固定部分参数估计结果

固定效应	回归系数	标准误	t 检验	自由度	P 值
第一层截距, B0					
截距, G00	316	50	6.26	13	0.000
品类联结资产, G01	1.01	0.08	12.51	13	0.000
第一层品牌力斜率, B1					
截距, G10	2.59	0.60	4.30	104	0.000

表 8 完全模型 随机部分参数估计结果

随机效应	标准差	方差	自由度	卡方	P 值
截距	8.01	64.22	13	2.93	>.500
第一层	518	267,843			

如以上结果，交互作用显著，品类作为第二层变量，对品牌联结资产和品牌力有显著性影响( $P=0.000$ )，同时，作为第二层的品类很好的解释了品牌间的差异( $P>.5$ )。这说明在互联网上发展较成熟的品类，其所属的品牌在互联网上的品牌联结资产与品牌力的相互作用就越明显。**假设 4 成立，品牌所属的品类，会对品牌联结资产与品牌牌力的关系产生显著性影响。**

## 5. 讨论

### 5.1 需要用全面的联结行为作为互联网时代品牌管理的工具

品牌已经存在于消费者大脑和互联网外脑。在数字时代，品牌不仅存在于消费者的头脑中，还存在于互联网这个大空间中，互联网上的各类信息已经成为我们可以任意读取、存储、重复利用的廉价的电子记忆形态(E-Memory)<sup>12</sup>。消费者与品牌在互联网上的联结行为纷繁复杂，多种多样。

从文中分析可知，单一的搜索或者社交数据只是品牌表现的一部分。品牌在互联网上与消费者联结的各触点并不是孤立存在的，它们对品牌有不同的作用，且触点间相互影响，共同形成了品牌在互联网上的数字化状态。因此，企业需要一个综合全面的数据指标对品牌进行测量。

品牌联结资产的提出，正是基于对品牌各触点的整合，让企业不仅可以直观了解品牌的整体表现，还可以结构化地对品牌表现进行分解，有目的对品牌进行管理，以确保品牌在互联网上全面、均衡、健康发展。

### 5.2 品牌联结资产为品牌的数字化管理提供了新方法

大数据技术的应用，使得消费者与品牌各触点的联结可以被追踪和记录，并最终转化为数字，让品牌可以被测量。品牌的数字化测量方法相对传统的品牌调查，有诸多优势。

一、**速度快**。品牌联结资产按照统一标准，实时采集消费者行为大数据，可以程序化的应用于品牌监测中，使之可以敏捷、动态地反馈品牌在市场中的表现，加快品牌管理的迭代更新。

二、**数据客观**。与传统的特定情境下的访问式调查不同，品牌联结资产作为一个观察者，真实记录消费者与品牌在互联网上的每一次接触，剔除人为干扰及方法上的系统误差，保证测量结果客观准确。

三、**结果精细**。相较于粗略的抽样访问，品牌联结资产能够覆盖全国所有省市和各级别市场，洞察所有网民与品牌的各个触点，全面而立体地测量品牌在市场中的表现。

### 5.3 数字化发展阶段不同的品类应采用不同的品牌数字化策略

数字化程度较低的品类应加强品牌与消费者在官方类网站上的联结。由前文分析可知，数字化发展程度较高的品类在官方渠道，如内容联结中的官网、官方论坛，行动联结中的官方商城等，与消费者的联结量明显高于低品类的品牌。这说明消费者更倾向于从官方渠道上了解品牌和产品，并进行购买，这也说明品牌的官方类网站更易于与消费者建立联系。因此，企业可以通过加强官方类网站的管理，建立良好的品牌形象，加强与消费者的互动沟通，以增加消费者的联结量。

而数字化程度较高的品类应着重探索互联网应用的新方向。如前文分析，银行、手机、汽车处于数字化发展较高阶段，银行通过官网的支付功能与消费者联结，汽车正在试图让消费者在互联网上预约购买，而手机涌现出如小米的互联网品牌。尤此可见，互联网不仅是品牌推广的媒介，销售的渠道，也是服务本身。

### 5.4 探索行为数据与态度数据的结合，以更好洞察品牌

品牌联结资产的行为数据与问卷调查的态度数据相结合，能更完整洞察品牌。品牌联结资产由消费者行为大数据构成，是消费者对品牌主动行为的体现；而问卷调查在测量消费者的感情态度方面具有优势，用来了解消费者行为背后的原因。如文中分析，品牌联结资产与品牌力高度相关，因此，品牌联结资产可以和调查所得数据相交互，全面深度解读品牌表现。

在辅以问卷调查时，建议采用简短快速的 Online 调查。设计有针对性的问卷，可以定时、定地点、连续地投放在相关网站上，快速、周期性收集消费者反馈。这样，调查结果不仅能反映出消费者态度，更具有时间上的连续性，与品牌联结资产相结合时，研究结果更准确，更能适应品牌市场的瞬息万变，推动品牌管理的更新迭代。

#### 5.4. 本文贡献与研究局限

本文的理论贡献在于：探索性地提出了“品牌联结资产”这一测量品牌的新指标，并对其进行了验证。不论是在我国还是国外，大数据都尚属新兴领域，将大数据应用在市场营销领域以及品牌管理领域的研究更是少之又少，本研究既可丰富此领域的理论基础又对将来的研究起到抛砖引玉的作用。

另外，分层线性模型分析的引入，丰富了市场研究的方法，更为大数据在市场营销领域的应用提出了新的思路。

实践意义在于：在互联网的影响下，企业传统的对品牌的管理已经无法满足快速变化市场的需求，基于大数据的品牌联结资产，可以成为企业快速管理品牌的工具，其具有的敏锐、细节、准确的特性，帮助企业进行品牌市场决策。

尽管本研究取得了一定的成果，但仍存在一些不足。首先，由于传统调查得到的是截面数据，基于大数据的品牌联结资产在时间上的优势没有得到体现。其次，缺少适当的销量数据与品牌联结资产相互验证，品牌联结资产对销量的影响还无法有效证明。在未来，将引入线下的销量数据，对指标的销售预测能力进行探索。

## 参考文献

---

- 1 Edelman, D. (2010) 'Branding in the digital age: You're spending your money in all the wrong places', *Harvard Business Review*, December, No. 65
- 2 Hatch, M.-J., & Schultz, M. (2010). Toward a theory of brand co-creation with implications for brand governance. *Brand Management*, 17(8), 590–604.
- 3 Chen, S. (2001) 'Assessing the impact of the Internet on brands'. *Journal of Brand Management*, Vol. 8, No. 4, pp. 288-302.
- 4 Muller, B. and Chandon, J. B. (2004) 'The impact of a World Wide Web site visit on brand image in the motor vehicle and mobile telephone industries', *Journal of Marketing Communications*, Vol. 10, No. 2, pp. 153-165.
- 5 Page, C., & Lepkowska-Whiet, E. L. Web equity: A framework for building consumer value in online companies[J]. *Journal of Consumer Marketing*, 2002, 19(3), 231-248
- 6 Stephen, T. and Galak, J. (2012) 'The effects of traditional and social earned media on sales: A study of a microlending marketplace', *Journal of Marketing Research*, Vol. 49, No. 5, pp. 624-639.
- 7 Duana, W., Gub, B. and Whinston, A. (2008) 'The dynamics of online word-of-mouth and product sales — An empirical investigation of the movie industry', *Journal of Retailing*, Vol. 84, No. 2, pp. 233-242.
- 8 Jacques Bughin (2014) 'Brand success in an era of digital Darwinism', *JOURNAL OF BRAND STRATEGY VOL. 2. NO. 4, 355-365 WINTER 2013-2014*.
- 9 卢泰宏, 周志民. 基于品牌关系的品牌理论: 研究模型及展望 [J]. *商业经济与管理*, 2003(2):4–9.
- 10 Pitta D & Katsanis L. Understanding brand equity for successful extension [J]. *Journal of Consumer Marketing*, 12(4), 51~64 .1995.
- 11 刘红云, 孟庆茂. 教育和心理研究中的多层线性模型[J]. *心理科学进展*, 2002, 10(2)
- 12 Robert W. Clowes (2013) . 'The Cognitive Integration of E-Memory', *Rev.Phil.Psych.* DOI 10.1007/s13164-013-0130-y

---

## The exploratory study of using behaviors' big data in internet to measure brand performance

Ying. Lan      Beijing Xinsight technology Co. Limited.  
Beiping. Tan      Beijing Xinsight technology Co. Limited.

**Abstract** The Internet has been changed the relationship between the consumers and brands. As the meanwhile, the development of big data technology makes the relationships can be tracked. There are many different touch points between consumers and brands which all can be tracked. But the biggest difficult points are whether and which touch points from big data have high relationship with the traditional indices. Only solve this problem, the big data can be used in brand management. In this article, we will explore the relationship and apply it to brand measurement in digital era. We collected the big data about nearly 200 brands information. Each brand has more than 30 touch points information which gathered to four big parts, Content connection, Desire connection, Action connection and Interaction connection. We found that, the behaviors' big data in internet can be gathered and calculated to a single one index, which named Brand Connection Quotient, to measure the brand performance. We also found that different category has different stages of development in the internet. Some categories have very high frequency connection with consumers but some are very low. The other very important result is the index of behaviors' big data shows significant impact upon the traditional brand power index and the category has the same significant impact to the relationship between the index of behaviors and brand power index. With the research, We learn that big data is truly a very useful technology to measure brand performance in the internet and the real world. It makes the measurement more rapid, detailed and accurate. The companies need to care about the changing relationship about the brand and consumer and use the big data technology to measure the brands and their consumers in marketing.

**Key Word** Brand Connection Quotient, Internet behavior, HLM, Big data, Brand tracking