我国健康保健类网站综合影响力评价研究

陈 玥 邓朝华

(华中科技大学同济医学院 武汉 430030)

[摘要] 采用直接访问网站和利用网站信息查询工具的方法调查我国 30 家健康保健类网站结构和内容、功能、交互性等, 选取网页数、链接数等指标, 借助链接分析法和灰色关联度分析法对网站影响力进行排名和评价。

[关键词] 健康保健类网站; 网络影响力; 灰色关联度分析法; 链接分析法

[中图分类号] R-056 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2019.06.011

Study on the Assessment of Comprehensive Influence of Health Care Websites in China CHEN Yue, DENG Zhaohua, Tongji Medical College of Huazhong University of Science and Technology, Wuhan 430030, China

[Abstract] The investigation on structure, content, function, interactivity and other aspects of 30 health care websites in China has been implemented through browsing the websites and utilizing website information inquiry tools. The paper chooses the number of website pages, the number of website links and so on as indicators to rank and assess the websites' influence by means of link analysis method and Grey Relation Analysis (GRA).

[Keywords] health care websites; network influence; Grey Relation Analysis (GRA); link analysis method

1 引言

互联网作为一种新兴的大众传播媒介使用率非常高。根据第 41 次中国互联网络发展情况统计报告^[1], 到 2017 年 12 月底互联网普及率为 55.8%, 网民规模 7.53 亿。健康网站将传统的健康知识普及与现代信息技术如在线咨询、就诊、挂号等相结合, 更好地满足人们对健康知识的需求, 成为当今公众最主要的健康信息来源^[2], 影响力逐渐增大。影响力是反映网站被利用程度的一个重要标志, 它越大表明越多的人知道这个网站, 其资源就越容易

被发现和利用^[3]。网站影响力评价是利用统计学、数据挖掘等手段对网络状态、特征以及内在规律进行定量描述和分析, 如通过对网站规模、用户访问量、搜索引擎评分、网络影响因子等指标分析能从一定程度上反映网站质量和网络影响力^[4]。

许多学者都从不同方面对健康网站进行评价。高琴^[5]利用层次分析法, 从内容和易用性等方面对网站网络信息资源质量进行评价。黄成^[6]建立我国医学健康网站可用性评价指标体系, 选择 10 个网站进行研究, 以了解我国医学健康网站的可用性现状。张会会^[7]等选取认证情况、信息来源等指标对 20 个健康网站的信息质量进行评估。目前对于网站影响力的研究较少, 在此背景下本文以此为切入点, 以期对网站的未来发展和功能完善有借鉴作用。

[修回日期] 2018-12-07

[作者简介] 陈玥, 硕士研究生; 通讯作者: 邓朝华, 副教授, 博士生导师。

2 研究内容及方法

2.1 研究内容

2.1.1 网站结构及内容 通过直接浏览的方法调查网站功能及内容、交互性、百度收录、隐私政策、广告铺设、友情链接和多平台服务 7 个指标并进行评价分析。

2.1.2 网站影响力 指网站的信息资源与服务通过网络实现传递、交流和利用,从而改变他人思想和行动的能力^[8]。链接分析法是网站影响力的客观

评价指标,其基本理论基础是一个网站对另一个网站的链接是对其的认可,网站的外部链接数越多说明其他网站认可程度越高,其内容传播的范围也越广,因此网页之间的相互链接能在一定程度体现网站影响力^[3]。本文从衡量网站链接数、应用、建设等方面的指标中选取网页数、总链接数、内链接数、外链接数、网络影响因子、内部网络因子、外部网络因子、Alexa 排名、谷歌网页分级 (Page Rank, PR) 和日均页面视图 (Page View, PV) 10 个分析指标,具体情况,见表 1。

表 1 网站影响力评价指标及描述

指标	指标描述
网页数	指网页内网站数量,在一定程度上反映网站的规模和丰富程度
总链接数	指网站被链接的总数量,体现网站影响力与网络辐射力
内链接数	是来自网站自身内部链接情况的指标,集中反映网站内部层次结构的完备性
外部链接数	是对某网站范围外搜索到的与该网站存在链接的网页数,能更好地反映网站质量
网络影响因子	指总链接数/网页数,反映网站网页被链接的总平均水平
内部网络影响因子	指内链接数/网页数,反映网站网页自链接的总平均水平
外部网络影响因子	指外链接数/网页数,为网站网页被外部链接的总平均水平
Alexa 排名	指网站世界排名,反映网站影响力
谷歌 PR	PR 值从 1 级到 10 级,PR 值越高网页越重要
日均 PV	指页面浏览量或点击量,可用于衡量网站访问量情况

2.2 研究方法

目前有多种方法可用于评价模型的构建,本文采用灰色关联度分析法。目的是对信息不完全的系统作因子间的量化、序化。实质是整体比较,是有参考系、有测度的比较^[9]。具体步骤如下:首先选取 Alexa 排名的最小值以及其余 9 项指标的最大值作为参考数列,设为 X_0 ,各项指标的数值为比较数列,记为 $X_i(k)$ 。其次,分别计算各比较数列 $X_i(k)$ 与参考数列的绝对差值,从每项指标的差值中分别选出一个最大值和最小值,记作 $\max |X_0(k) - X_i(k)|$ 和 $\min |X_0(k) - X_i(k)|$,再从 $\max |X_0(k) - X_i(k)|$ 和 $\min |X_0(k) - X_i(k)|$ 中选出最大值和最小值,分别记作 $\max_i \max_k |X_0(k) - X_i(k)|$ 和 $\min_i \min_k |X_0(k) - X_i(k)|$ 。然后,比较数列 X_i 上各点与参考数列 X_0 的关联系数 ζ ,公式

为:

$$\zeta_i(k) = \frac{\min_i \min_k |X_0(k) - X_i(k)| + \zeta \max_i \max_k |X_0(k) - X_i(k)|}{|X_0(k) - X_i(k)| + \zeta \max_i \max_k |X_0(k) - X_i(k)|}$$

$\zeta \in [0, 1]$ 为分辨系数,一般按最少信息原理取为 0.5,即 $\zeta = 0.5$ 。最后,根据关联度公式计算关联度:

$$r_i = \frac{1}{N} \sum_{k=1}^N \zeta_i(k)$$

通常关联度 r_i 值越大关联程度越高。

3 数据收集

站长之家 (<http://www.chinaz.com/>) 是专门针对中文站点提供资讯、技术、资源、服务的网站,其网站排序是由页面基础得分(标题、描述、关键词等)、站内得分(内链等)、站外得分(外链

等) 和用户得分 (PV 量、回访量、搜索结果点击量等) 的总分进行排序。在健康保健类别选取综合统计排名前 30 的健康网站, 相关数据的收集和整理工作于 2017 年 12 月 30 - 31 日完成。目前网络计量学研究经常使用的搜索引擎主要有 Altavista、All the web、Google 与百度等。Altavista 具有较强的检索功能, 提供多种类型的限制检索, 同时具有较强的稳定性。All the web 拥有较大容量的数据库, 而且更新速度快, 搜索精度较高, 但两者对中文的支持并不是很好^[10]。本研究采用更新快、稳定性强以及检索功能强大的 Google 高级检索功能查询网页数和总链接数, 以 39 健康网为例的检索指令分别为 site: http://www. 39. net/和"http://www. 39. net/"。此外通过直接访问收集网站功能及内容、交互性、隐私政策、广告铺设等指标并进行分析, 同时利用站长工具网站中的搜索引擎优化 (Search Engine Optimzation, SEO) 综合查询功能对所有网站的内链接数、外链接数、友情链接数、PR 值和日均 PV 值进行查询。

4 数据分析

4.1 网站结构及内容 (表 2)

表 2 我国健康保健类网站相关数据统计

条目	健康网站数 (个)	百分比 (%)	
交互方式	问答咨询	19	63.33
	一对一咨询	6	20
	两者都无	11	36.67
百度快照	≤1 天	25	83.33
	1 ~ 10 天	5	16.67
	≤1 万	3	10
百度收录量	1 ~ 100 万	17	56.67
	100 ~ 1 000 万	7	23.33
	≥1 000 万	3	10
多平台服务	微信或微博	23	76.67
	客户端	18	60
	两者都无	5	16.67

4.1.1 网站功能及内容 调查显示健康保健类网站大都提供较为全面的健康资讯, 包括保健、饮食、美容、整形等, 来源可分为两类, 即网站自行编辑及其他网站内容。此外大部分网站都提供对医

院、药品和健康问题以及站内健康问答查询功能, 小部分网站有在线预约挂号功能, 对于网站使用者来说, 这些功能实用且方便。但是网站内容方面存在一些不足, 部分网站健康资讯的可信度不高, 如某网站对健康资讯的来源标注为互联网, 作者标注为佚名, 网站使用者无法了解信息来源。有的网站排版过于密集, 内容呈现比较杂乱, 使用者无法直观清晰地查看想了解的内容, 浏览体验较差。呈现方式比较单一, 大部分都只以图文为主, 而视频和音频信息相对较少。

4.1.2 网站交互方式 交互性可通过两个方式实现, 63.33% 网站都有问答方式, 用户提出疾病问题, 等待医生回复, 多为不严重的问题如消化不良、失眠等, 少数网站可供用户选择医生进行一对一沟通, 用图文或电话进行咨询, 这种方式更有针对性, 可能会收取用户一定服务费。

4.1.3 百度收录和快照 百度收录量即某网站被百度收录的网页数量, 对网站排名有重要作用。调查显示百度收录量在 100 万以上的网站占 33.33%, 收录量最多的是有问必答网站, 为 6 540 万以上, 最少的是医生在线网站, 仅为 554, 两者相差较大。百度快照更新较快, 由表 2 可知 30 家网站均在 10 天以内, 且 83.33% 的网站都在 1 天以内, 更新比较频繁。

4.1.4 网站隐私政策 通过浏览各网站的隐私保护政策或网站声明等, 查看网站是否向用户承诺除法律用途外不向第 3 方披露用户隐私。半数以下网站有完整的隐私保护条款, 其他网站存在服务声明未提及隐私内容相关声明或只有隐私免责声明的问题。

4.1.5 网站广告铺设 网站广告数量和位置对用户浏览体验有重要影响, 因此有必要从用户角度查看各网站广告情况。各网站广告投放主要在网页右侧和下部, 约半数网站都有投放广告, 一定程度上影响用户浏览体验, 各网站广告都没有遮挡资讯内容, 不存在需要用户关闭浮窗才能看到内容的广告设置。

4.1.6 友情链接 网站与网站间的一种简单合作, 是网站互相放置对方链接的推广方式, 对增加网站流量有很大帮助。调查显示 30 家网站都提供友情链接, 数量一般在 50 以下, 链接的有健康保健相关网站, 也有关于明星或星座等与健康保健无关的网站。

4.1.7 多平台服务 随着移动互联网的发展,仅使用网站已无法满足用户需求,是否开通微博、微信或手机 App 等功能可以用来衡量网站服务质量。调查显示有客户端或微博、微信平台的占半数以上,分别为 76.67% 和 60%,两者都没有的网站仅有 5 个,占 16.67%,可见大部分网站对移动端用户比较重视,开始利用其他平台吸引用户。

4.2 网站影响力

4.2.1 网站单向影响力分析 我国部分健康类网站影响力指标相关数据,见表 3。网站平均网页数为 325 293 个,说明各网站有一定规模和丰富程度,其中有问必答网的网页数最多。网站平均总链接数、外部链接数及内部链接数分别为 398 933 个、

4 535 个和 670 个,较为丰富,也有一定的关注度,但是各网站间的链接数差距较大,如好大夫在线的总链接数有 1 570 000 个,而中医中药秘方网仅有 10 300 个。网站 Alexa 平均排名为 54 844,整体排名情况良好,39 健康网排名为 224,比较受欢迎,同类网站中影响力较大。日均 PV 为 425 422,有较高的访问量,用户使用率高,具备对用户的吸引力,其中 39 健康网日均 PV 最高,健康第一线和求医网日均 PV 最低。PR 值为 5 及以上的网站共 19 个,占 63.33%,网站 PR 值为 7 及以上的网站共 4 个,占 13.33%。39 健康网、好大夫在线、家庭医生在线和爱康国宾体检 PR 值最高,乐哈健康网、百姓健康网和中医中药秘方网 PR 值最低。

表 3 我国部分健康保健类网站影响力指标相关数据

网站名称	网页数	总链接数	内部链接数	外部链接数	网络影响力因子	内部网络影响力因子	外部网络影响力因子	Alexa 排名	PR	PV
39 健康网	42 100	585 000	2 534	28 830	13.895 4	0.060 2	0.684 7	224	7	11 336 025
有问必答网	3 970 000	4 290 000	1 465	5 860	1.080 6	0.000 4	0.001 5	17 995	6	65 250
好大夫在线	1 430 000	1 570 000	868	4 400	1.097 9	0.000 6	0.003 1	20 010	7	48 600
三九养生堂	422 000	608 000	387	23 822	1.440 7	0.000 9	0.056 4	22 367	5	57 000
A 医学百科	98 200	117 000	22	265	1.191 4	0.000 2	0.002 7	53 561	4	54 000
寻医问药	74 400	115 000	715	27 477	1.5457	0.009 6	0.369 3	15 139	6	103 275
99 健康网	104 000	298 000	1 687	5 058	2.865 4	0.016 2	0.048 6	20 563	6	58 800
飞华健康网	69 000	144 000	658	13 006	2.087	0.009 5	0.188 5	23 520	6	48 300
乐哈健康网	124 000	164 000	102	266	1.322 5	0.000 8	0.002 1	46 355	3	36 450
大众养生网	603 000	892 000	66	512	1.479 3	0.000 1	0.000 8	30 109	4	56 250
中华康网	167 000	247 000	631	408	1.479	0.003 8	0.002 4	25 500	5	34 200
家庭医生在线	32 300	43 700	1 873	6 510	1.352 9	0.058	0.201 5	1 078	7	27 000
丁香园	674 000	708 000	413	10 388	1.050 4	0.000 6	0.015 4	14 704	6	89 100
苹果绿养生网	96 700	165 000	1 763	472	1.706 3	0.018 2	0.004 9	103 050	5	9 000
养生之道网	227 000	344 000	585	895	1.515 4	0.002 5	0.003 9	47 938	4	9 750
医学百科	123 000	172 000	69	288	1.398 4	0.000 6	0.002 3	82 524	4	6 000
百姓健康网	502 000	530 000	559	839	1.055 8	0.001 1	0.001 6	36 420	3	20 625
中华养生网	107 000	202 000	144	544	1.887 9	0.001 3	0.005 1	105 657	4	30 000
医脉通	26 700	26 700	231	376	1	0.008 7	0.014 1	23 652	5	48 600
就医 160 挂号网	821	36 700	1 137	428	44.701 5	1.384 9	0.521 3	59 980	4	30 000
爱康国宾体检	2 620	20 100	44	265	7.671 8	0.016 8	0.101 1	27 726	7	133 650
天天营养网	25 500	66 900	212	234	2.623 5	0.008 3	0.009 2	54 886	5	42 750
医生在线	293	32 300	393	336	110.238 9	1.341 3	1.146 7	91 099	5	26 250
21 保健品网	4 530	29 100	502	265	6.423 8	0.110 8	0.058 4	33 286	4	240 000
华医网	3 670	17 600	200	188	4.695 6	0.054 5	0.051 2	82 393	5	99 000
健康第一线	225 000	248 000	467	104	1.102 2	0.002 1	0.000 5	95 635	6	1 500
360 常识网	3 480	142 000	706	115	40.840 6	0.202 9	0.033	89 624	4	2 250
求医网	32 200	8 3700	788	2 427	2.599 4	0.024 5	0.075 3	284 986	6	1 500
久久健康网	482 000	59 900	574	1 387	0.124 2	0.001 2	0.002 9	61 283	6	17 550
中医中药秘方网	86 300	10 300	304	89	0.119 4	0.003 5	0.001	74 082	3	30 000

4.2.2 网站综合影响力排序 利用灰色关联度分析法对 30 家健康保健类网站进行 10 项指标关联度排名, 见表 4。可以看出我国部分健康保健类网站灰色关联度在 0.84 ~ 0.94 之间, 关联度高于分辨系数

0.5, 以关联度 r_i 值越大、关联度越高的标准来看, 30 家网站都有较高影响力, 而排名前 3 的网站有问必答网、39 健康网和好大夫的关联度比较高, 说明网站综合影响力较高, 网站有一定的关注度和被认可度。

表 4 我国部分健康保健类网站灰色关联度排名

序号	网站名称	关联度 r_i	序号	网站名称	关联度 r_i
1	有问必答网	0.932 7	16	家庭医生在线	0.849 1
2	39 健康网	0.919 5	17	A 医学百科	0.849 0
3	好大夫在线	0.869 2	18	医学百科	0.848 8
4	大众养生网	0.857 6	19	中华养生网	0.848 6
5	丁香园	0.857 4	20	21 保健品网	0.848 6
6	三九养生堂	0.855 0	21	医脉通	0.848 5
7	百姓健康网	0.854 4	22	爱国康宾体检	0.848 4
8	养生之道网	0.851 2	23	苹果绿养生网	0.848 3
9	久久健康网	0.850 9	24	天天营养网	0.848 2
10	99 健康网	0.850 8	25	中医中药秘方网	0.847 9
11	中华康网	0.850 6	26	就医 160 挂号网	0.847 8
12	寻医问药	0.850 1	27	360 常识网	0.847 8
13	健康第一线	0.849 7	28	华医网	0.847 4
14	飞华健康网	0.849 7	29	医生在线	0.847 2
15	乐哈健康网	0.849 5	30	求医网	0.844 6

5 结论

通过对健康保健类网站结构及内容以及影响力调查分析了解到各网站都提供健康资讯、健康问答、药品和医院查询等基础功能; 有不同交互方式供用户选择, 满足不同需求。较高的网页数、链接数和日均 PV 表明网站有一定规模和影响力, 被关注情况较好。部分网站页面、广告、排版等仍需要努力改进, 存在信息形式较为单一、内容不够严谨独特等问题, 这些差距会对用户的浏览体验产生影响, 可能会流失潜在用户, 上述问题在 Alexa 排名偏低的网站中有所体现。多数网站对于移动端有一定程度的重视, 充分利用移动端平台可以为网站吸引更多潜在用户。针对上述问题提出以下两点建议: 一是丰富网站内容。内容和质量对网站有显著影响。用户浏览此类网站一般是因为有疾病问诊或对健康信息有了解的需求。若想丰富网站内容, 应从健康信息资源入手, 提高健康栏目信息内容和质

量, 及时更新有价值信息, 删除失效或过时的信息, 以便更好地满足大众健康信息需求, 同时扩充专家团队力量, 增加原创、独创、新颖的信息内容, 加强特色建设, 以增强健康信息的可信度及权威性。二是利用信息技术手段优化和完善网站链接。从各项链接指标结果看, 不少网站仍需要优化和完善网络结构, 目前链接方式相对单一。而完善丰富的链接手段可以为网站带来相对稳定的用户访问量, 也能提高百度收录和网站权重等。同时增加其他网站链接, 使访问者在链接到其他网站时提高本网页在检索结果中的排名以及在搜索引擎中的可见度^[10]。网站需优化和完善链接, 使网络结构更加合理, 进而提升影响力。

参考文献

- 1 中国互联网络信息中心. 第 41 次中国互联网络发展状况统计报告 [EB/OL]. [2018-06-08]. <http://cn-nic.cn/hlwfzyj/hlwzbg/hlwtjbg/201803/P020180305409870339136.pdf>.

(下转第 72 页)

- 学报 (哲学社会科学版), 2015, 68 (2): 80-85.
- 2 郝莎, 董芳, 胡林萍, 等. 造血干细胞生物学及临床应用研究概况 [EB/OL]. [2019-02-03]. <http://kns.cnki.net/kcms/detail/31.2035.Q.20190121.1011.006.html>.
 - 3 中国科学院. 中国学科发展战略, 再生医学 [M]. 北京: 科学出版社, 2015.
 - 4 NLM. PubMed [EB/OL]. [2019-03-05]. <https://www.ncbi.nlm.nih.gov/pubmed/>.
 - 5 Clarivate. DII [EB/OL]. [2019-03-05]. <https://www.isiknowledge.com/>.
 - 6 Clarivate. Cortellis [EB/OL]. [2019-03-08]. <https://clarivate.com.cn/>.
 - 7 Clarivate. Derwent Data Analyzer [EB/OL]. [2019-03-05]. <https://clarivate.com/products/derwent-data-analyzer/>.
 - 8 Van N J, Waltman L. Text Mining and Visualization Using VOSViewer [J]. ISSI Newsletter, 2011, 7 (3), 50-54.
 - 9 吴益宏. 基于文献词篇共现的浮针研究可视化分析 [J]. 医学信息学杂志, 2018, 39 (12): 66-72.
 - 10 蔡晓珍, 朱东屏, 陈贞春, 等. 神经干细胞发展热点可视化分析 [J]. 医学信息学杂志, 2017, 38 (11): 67-72.
 - 11 澎湃新闻. 科学家发现造血干细胞“归巢”秘密: 有望破解肿瘤转移 [EB/OL]. [2019-03-12]. <https://tech.sina.com.cn/d/f/2018-11-20/doc-ihnyuqhi4599322.shtml>.
 - 12 Dantong L, Wenzhi X, Weijun P, et al. T VCAM-1 + Macrophages Guide the Homing of HSPCs to a Vascular Niche [J]. Nature, 2018, 564 (11): 119-124.
 - 13 Zhang C, Chen Y, Sun B, et al. m6A modulates haematopoietic stem and progenitor cell specification [J]. Nature, 2017, 549 (9): 273-276.
 - 14 Sugimura R, Jha D K, Han A, et al. Haematopoietic stem and progenitor cells from human pluripotent stem cells [J]. Nature, 2017, 545 (5): 432-438.
 - 15 张梦然. 美制造出具有造血干细胞功能的细胞. [N]. 科技日报, 2017-05-18 (01).
 - 16 Javier C, Yousef R, Raif S, et al. 2016 年基础及临床免疫学进展 [J]. 中华临床免疫和变态反应杂志, 2018, 12 (2): 215-229.
 - 17 张健萍, 程涛, 张孝兵. 造血干细胞基因治疗: 进展与挑战 [J]. 中国科学: 生命科学, 2017, 47 (12): 1323-1335.
 - 18 Andrew R, Gennery M, Mary A, et al. Transplantation of Hematopoietic Stem Cells and Long-term Survival for Primary Immunodeficiencies in Europe: entering a new century, do we do better? [J]. Journal of Allergy and Clinical Immunology, 2010, 126 (3): 602-610.
 - 19 Lisowski L, Dane A P, Chu K, et al. Selection and Evaluation of Clinically Relevant AAV Variants in a Xenograft Liver Model [J]. Nature, 2014, 506 (2): 382-386.

(上接第 53 页)

- 2 俞文敏, 王杰, 周宏宇, 等. 健康需求者对健康类网站的知识需求调查 [J]. 护理学杂志, 2009, 24 (5): 86-87.
- 3 姜吉栋, 彭洁, 赵辉. 网站影响力评价研究现状综述 [J]. 情报科学, 2015, 33 (9): 157-161.
- 4 赵乃瑄, 张若冉. 跨系统区域图书馆联盟网络影响力评价研究 [J]. 图书情报工作, 2017, 61 (7): 28-33.
- 5 高琴. 中文健康信息网站的评价 [J]. 中华医学图书情报杂志, 2010, 19 (2): 40-44.
- 6 黄成. 基于非医学专业信息用户需求的我国医学健康网站可用性评价研究 [D]. 重庆: 西南大学, 2008.
- 7 张会会, 马敬东, 蒋春红, 等. 健康类网站信息质量的评估研究 [J]. 医学信息学杂志, 2013, 34 (7): 2-6.
- 8 曾荷. 电子政务信息资源的网络影响力评价研究 [D]. 上海: 华东师范大学, 2007.
- 9 邓聚龙. 灰色系统基础方法 [M]. 武汉: 华中科技大学出版社, 2005: 74-103.
- 10 张向先, 袁小姗. 基于链接分析法的我国省级科技信息研究所网站影响力评价研究 [J]. 情报理论与实践, 2011, 34 (3): 91-94.