**专家领域影响力分析**

（一）国内外现状

著名的Google PageRank[1]算法是一种根据网页之间相互的超链接数量来进行网页排名技术。PageRank算法最初仅是应用在搜索引擎中，用来计算网页排名，但随后，部分学者将PageRank算法引用到社交网络中，认为PageRank算法是社交网络用户个体影响力分析的基础算法[2][3]。

在学术领域，经典的科研专家的学术影响力评价指标包括文章引用次数、H指数、I10指数、G指数等。随着Web2.0技术及社交网络的发展，这些评价体系的权威性受到了冲击，Priem等提出一种多元化科研人员评价体系Altmetrics[4]，使用更多的社交网络参数来进行学术声望评判。Altmetric评分根据不同社会化媒体引用文献的次数，一般包括新闻报纸、博客、微博，论坛等。

基于社交网络的用户影响力分析模型，国内外各大学术社交网络中针对其科研用户也推出各类影响力排名算法，由清华大学唐杰等人开发的Aminer[5]研究人员社会网络，通过分别统计科研人员的文章数量、引用数量、H 指数、A 指数、G 指数等信息，生成专家统计信息雷达图，并可分别通过上述指数进行专家排名。

（二）研究意义

识别给定领域的主要专家及其研究内容，并构建专家影响力评价指标，对专家在领域内影响力进行评价。结合专家学术文献影响和社交网络影响，给出专家影响力分析的计算模型。

（三）研究内容

基于社交网络的专家影响力的具体计算模型如下：

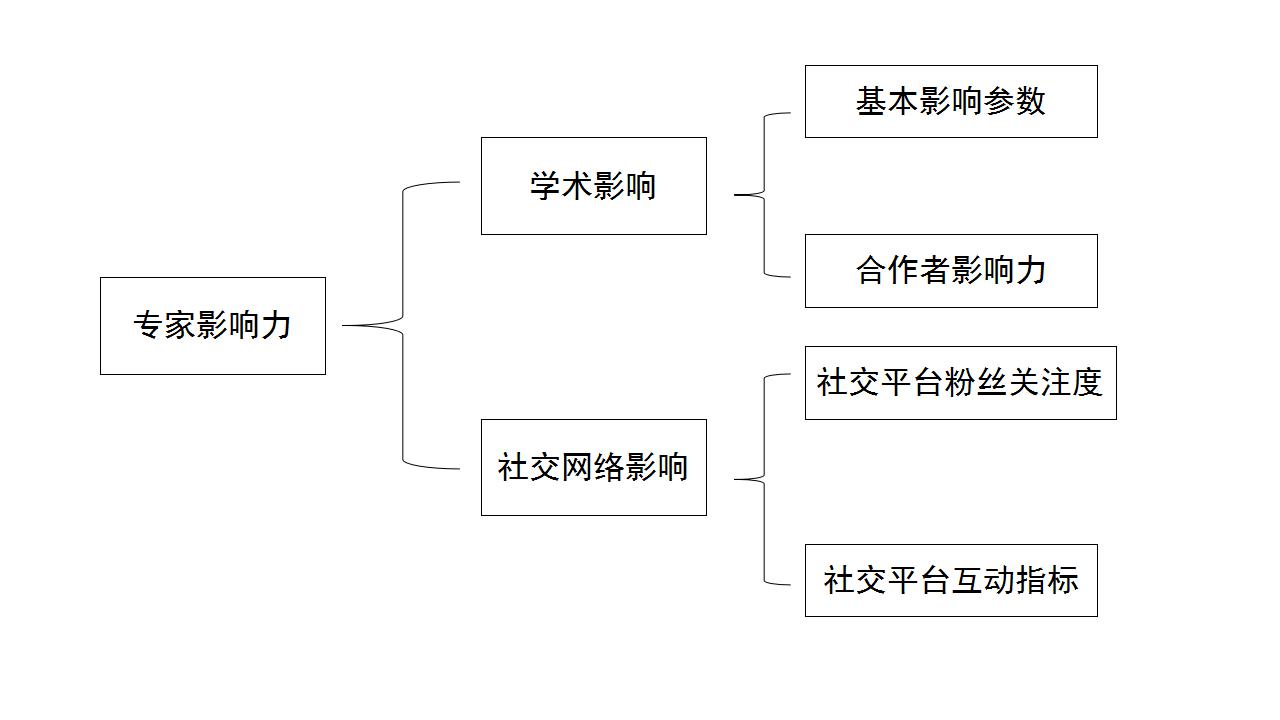


图1 专家影响力评价模型

（1）将专家影响力的影响因素分为学术影响参数A与社交网络影响参数S两部分。

（2）将专家学术声誉计算参数定义为A。科研领域学术影响因子确定为基本影响参数与合作者影响参数两部分。基本影响参数引入专家的H指数、总影响因子和所发文章总数三个参数。基本影响参数的具体公式定义为：



其中H为H指数，I表示影响因子总和，P表示作者所发文章数量。该公式表示，如果在发的论文总数相同的情况下，作者的H指数和总影响因子越高，说明该专家的科研声望值越高。

同时将合作者影响力参数引入到专家影响力指标中，认为文章合作者的影响力会对专家的影响力有较大的影响。最终确定公式为：



式中U表示合作专家的影响力值，L表示该专家在文章中的署名位置，第一作者为1，第二作者为2，以此类推。

（3）引入PageRank 算法进行社交网络专家影响力分析。

根据PageRank 公式，模型中的S 参数定义如下：

对于不同社交平台（如微博、公众号、领英等），首先计算专家的粉丝关注情况值



R表示要计算的专家粉丝关注情况值，B为该专家的粉丝数（被关注数），F表示关注数，c为常数。该模型认为拥有越多高影响力粉丝的专家，那么该专家的学术影响力值也就越高。

考虑到专家在社交平台发布消息的互动情况，计算



T 表示某专家的用户评论情况值，G 表示好评数，N 表示差评数，C 表示所有评论数。该公式指出，用户对某专家的好评越多，该专家的影响力越高；差评越多，该专家的影响力越小。

确定社交网络影响参数为专家关注情况与评论情况之和：



（4）定义专家影响力模型AS公式为：



AS表示专家影响力评价值，S 表示专家社交网络影响参数值，A表示学术影响参数值。a与b表示两类影响因子权值，且a+b=1。

参考文献

[1]Page L, Brin S, Motwani R, et al. The PageRank citation ranking: Bringing order to the web[R]. Stanford InfoLab, 1999.

[2]Tunkelang D. A twitter analog to pagerank[J]. The Noisy Channel, 2009.

[3]Haveliwala T, Kamvar S, Jeh G. An analytical comparison of approaches to personalizing pagerank[R]. Stanford, 2003.

[4]Bhue S, Singh K, Biswal S K. Altmetrics: Article Level Metrics Makes Easy for User Community[J]. 2016.

[5]唐杰. AMiner[EB/OL]. (2006-09-06)http://aminer.org.