B题： **学术期刊相互关系密切程度的度量模型**

[Web of Science](http://tlink.lib.tsinghua.edu.cn/go?url=http://apps.webofknowledge.com/)是一个学术信息服务平台，支持自然科学、社会科学、艺术与人文学科的文献检索，数据来源于期刊、图书、专利、会议录、网络资源（包括免费开放资源）等。用户可以同时对该平台上已订购的所有数据库进行跨库检索或选择其中的某个数据库进行单库检索。

利用Web of Science 丰富而强大的检索功能，各学科领域的学者们可以方便快速地找到有价值的科研信息。同时，Web of Science提供了学术期刊评价与排名的多项指标，是学术研究与管理的重要辅助工具。

目前，清华大学用户可以访问平台上各文献数据库。

**期刊相似性的度量与刻画**是学术期刊评价、分类与聚类等研究基础。科学引文分析领域的重要学者，也是《科学引文索引》（SCI）的创办人Eugene Garfield在2002年与合作者发表文章（见附件）对这一问题进行了一些研究，考虑到学术期刊之间相互引用和期刊规模（即文章数）等因素，给出了简便易算的期刊相关度参考指标。但是，这篇文章给出的指标并未充分考虑到期刊的相互认可。例如，期刊A大量引用了期刊B，而期刊B对A的引用非常少，这个时候，Garfield的指标仍会给出一个比较高的关联系数。而此时，A与B之间的关系是单方面的，A看重B，但是B并不看重A，这种情况通常意味着期刊B的等级水平高于期刊A。我们以“谈笑有鸿儒，往来无白丁”的观点，设想期刊之间也是相互交朋友的关系；关系密切的期刊，相互的认可程度都比较高。基于这样观点的指标更适合于区分不同水平等级的期刊。此外，期刊规模（即每年出版的文章数）与引用、被引用数量等参数对期刊之间相互关系的影响是否为线性关系？期刊的规模是否要考虑引用与被引数？等等。Garfield原始模型中有很多方面的量化处理值得进一步探讨。

**请同学们建立适当的数学模型，刻画学术期刊之间的相互认可程度。**

请同学们利用Web of Science的数据，尤其是其中Journal Citation Reports数据库，自行收集和选取数据。例如，选择自己感兴趣的学科的一些期刊，考察它们之间的引用关系，利用你提出的期刊关联度量模型，建立每个期刊关系密切的期刊列表，进而获得一个学科领域中不同等级期刊的聚类结果（类似于中科院的1区和2区等学术期刊分级），或更多的有意义的相关分析。

**Garfield的原始文章信息如下，文章全文见附件。**

A.I. Pudovkin, E. Garfield, **Algorithmic Procedure for Finding Semantically Related Journals**, ***Journal of the American Society for Information Science and Technology***, 2002, 53(13):1113–1119.