|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 论文题目 | Research interests: their dynamics, structures and applications in unifying search and reasoning | | | |
| Paper URL | https://sci-hub.se/10.1007/s10844-010-0144-1 | | | |
| Project URL |  | | | |
| 综述/背景介绍 | 发展状况 | 原因 | 意义 | 关键词（速记词汇、信息索引词汇） |
| 基于 Web 的文献搜索系统（例如 Google Scholar、CiteSeerX）和研究人员在线网络（例如ResearchGATE），可以为研究人员提供各种服务，但是许多系统和平台缺乏从其动态和结构特征的角度对研究人员的兴趣进行更深入的分析。从这两个角度理解研究兴趣的性质和模型可能有助于为研究人员提供更好和更积极的服务。 | | |  |
| 假设 | 衡量研究兴趣可能有助于为研究人员获取更多背景信息，以支持他们在网络上的活动。然而，如果作者没有提供足够的信息（例如未在任何地方明确显示的兴趣），则并非所有这些都可以衡量。另一方面，作者以前的出版物可以被视为可以提取其研究兴趣的来源。在本文中，我们通过作者以前的出版物来衡量作者的研究兴趣。 | | |  |
| 方法描述(含图) | 研究兴趣的动态和结构特征的理解 | | |  |
| 实验设计 |  | | |  |
| 数据处理 | 输入 |  | | 关键词（速记词汇、信息索引词汇） |
| 语句 |  | |  |
| 结论 | 本文在统一搜索和推理 (ReaSearch) 框架的基础上，提出了基于兴趣的搜索和推理 (I-ReaSearch) 方法，旨在统一基于用户兴趣的搜索和推理来解决Web问题。然后，在科学文献搜索的背景下，我们提供了一些示例，说明了如何从不同角度（即保留兴趣、兴趣最长持续时间和兴趣累积持续时间的角度）结合时间和结构特征，使用获得的兴趣统一搜索和推理。为I-REASEARCH开发了两种具体的策略，即基于兴趣的查询精化和基于兴趣选择的查询。通过添加用户兴趣作为查询的上下文，基于兴趣的查询精化与未精化的查询相比产生更好的查询结果，但同时花费更多的处理时间。它不能很好地扩展，因为处理时间增加得太快。然而，这些缩放问题可以通过应用具有基于兴趣的选择的查询来解决。这两种策略的结果是等价的，但后者需要的查询处理时间要少得多。 | | | |
| 局限性分析 |  | | | |

**（论文名）：**研究兴趣：它们的动态、结构以及在统一搜索和推理中的应用

**（题目）：**《Research interests: their dynamics, structures and applications in unifying search and reasoning》

**（论文URL）：**https://sci-hub.se/10.1007/s10844-010-0144-1

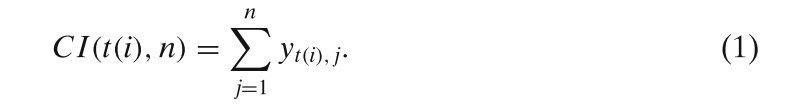
**（总结）：**本文集中研究兴趣，更具体地说是研究兴趣，以及它们在网络文献检索中的应用。提供了对研究兴趣的动态和结构特征的理解：对于动态的方面，本文讨论了一些跟踪研究兴趣动态变化过程的初步方法。在变化过程中，保留兴趣提供了研究人员当前兴趣的一些相关信息，因此，我们基于认知科学中的遗忘机制提出了保留兴趣模型，以便定量评估用户当前兴趣中剩下多少历史兴趣。在结构方面上，从网络理论的角度，考察了兴趣网络结构和演化过程的统计分布，对兴趣网络的演化特征提供了一些基本的认识。

本研究不仅旨在对研究兴趣的性质和模型提供初步认识，而且旨在将相关结果作为环境、上下文的基础，以便在网络文献检索过程中为研究人员提供更好的服务。本文在统一搜索和推理 (ReaSearch) 框架的基础上，提出了基于兴趣的搜索和推理 (I-ReaSearch) 方法，旨在统一基于用户兴趣的搜索和推理来解决Web问题。

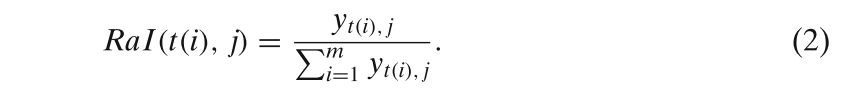
**（附图）：**

**提出了一些衡量研究兴趣的参数：**

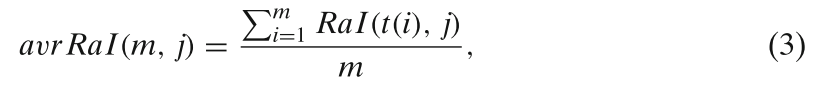
累积兴趣，表示为 CI(t(i), n)，用于计算在考虑的 n 个时间间隔内 t(i) 的累积出现次数。它可以表示为



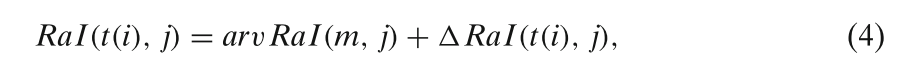
研究兴趣的比率，表示为 RaI(t(i), j)，是 t(i) 的兴趣与作者感兴趣的所有 m 个主题集合的兴趣之间的比率。



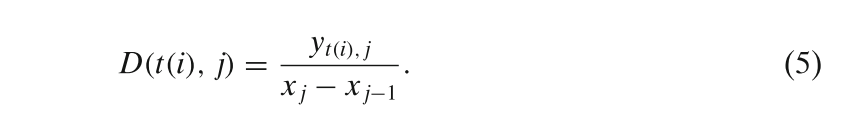
平均研究兴趣比率，表示为 avrRaI(m, j)，是时间间隔 j 内所有考虑的研究兴趣比率的平均值。



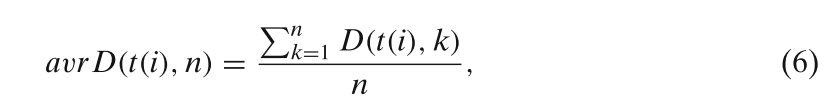
其中 m 是考虑的兴趣数。 RaI(t(i), j) 和 avrRaI(m, j) 之间的关系可以表示为：



研究兴趣程度，记为 D(t(i), j)，表示在时间间隔 j =[x j−1, x j] (x j -1 和 x j 表示时间间隔 j 的开始时间和结束时间）



平均研究兴趣度，表示为 avrD(t(i), n)，是主题 t(i) 在所有考虑的时间间隔内的研究兴趣度的平均值。

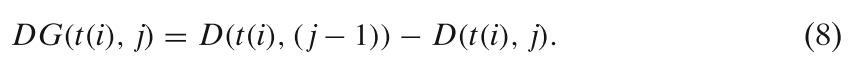


其中 D(t(i), k) 是主题 t(i) 的研究兴趣程度， k ∈[1, ..., n] 是一个特定的时间间隔。总共有 n 个考虑的时间间隔。

研究兴趣的相对程度，表示为 δD(t(i), k) 是 D(t(i), k) 和 avrD(t(i), n) 之间的差异。

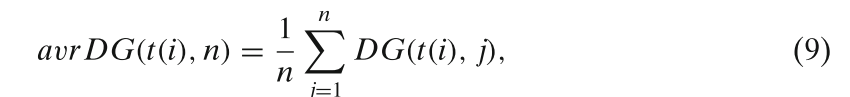


研究兴趣增长度（Degree ofresearch interest growth），记为DG(t(i),j)，是t(i)在连续两个时间间隔 (j-1) 和 j 的研究兴趣度增长：



可以通过 DG(t(i), j) 的值来比较不同主题的研究兴趣增长。假设作者的研究兴趣列表中有两个任意主题 t(i) 和 t(i')。如果 DG(t(i), j)> DG(t(i'), j)，那么我们说作者在时间间隔 j，t(i) 的研究兴趣增长高于 t(i') 。

平均研究兴趣增长程度，记为avrDG(t(i),n)，是主题t(i)在不同时间间隔内的DG(t(i),j)的平均值。



其中 n 是考虑的时间间隔的总数。

研究兴趣增长的相对程度，记为δDG(t(i),k)，是研究兴趣增长DG(t(i),k)与平均研究兴趣增长程度avrDG(t(i),n)的差。

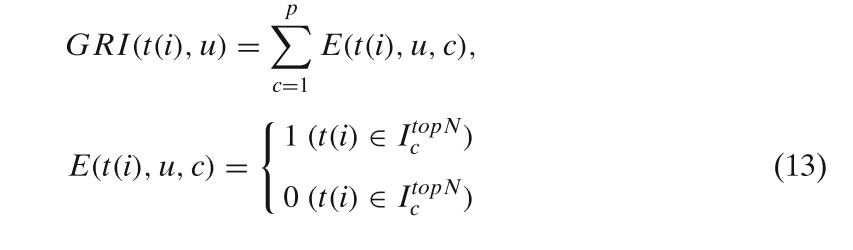


单个研究兴趣的权重，记为 w(t(i), j)，是在指定时间间隔 j =[x j−1, x j] 内出现的所有主题中主题 t(i) 的权重。

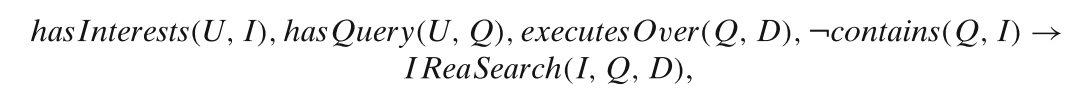


我们在此对个人留存兴趣与其群体留存兴趣的关系进行初步研究。

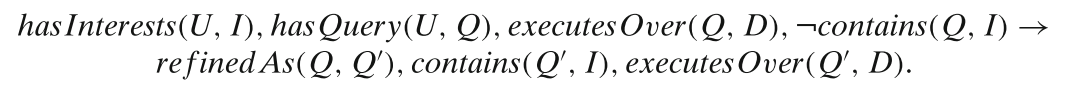
对于特定兴趣“t(i)”，对特定作者“u”群体留存兴趣，即“GRI(t(i), u)”可以定量定义为：

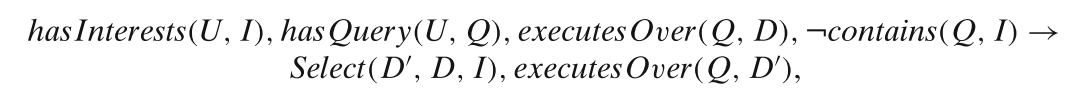


**提出了I-SeaReach 框架：**



以及两种实现方法





当查询模糊/不完整时，研究兴趣可以作为约束条件，用于优化这些查询。它们可以从不同的角度进行评估，每个角度都反映了研究兴趣的一个独特特征。当用户对某一角度不满意时，可以更改用于基于兴趣的 ReaSearch 的兴趣列表。基于兴趣的 ReaSearch 的过程可以用以下规则来描述：

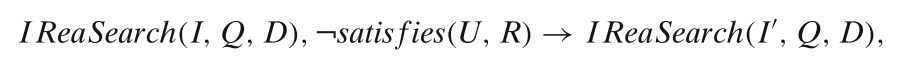


表3显示了使用模糊查询“Intelligent”对搜索结果的部分比较研究，并将来自各种兴趣列表的隐式约束添加到原始查询中。基于这三个角度，选择不同的搜索结果并提供给用户，以满足其不同的需求（在这部分结果列表中，选择了包含查询关键字和研究兴趣约束的文献并将其排在首位。作为示例，在每个列表中，我们的系统显示了根据每个研究兴趣的约束获得的第一个搜索结果）。

