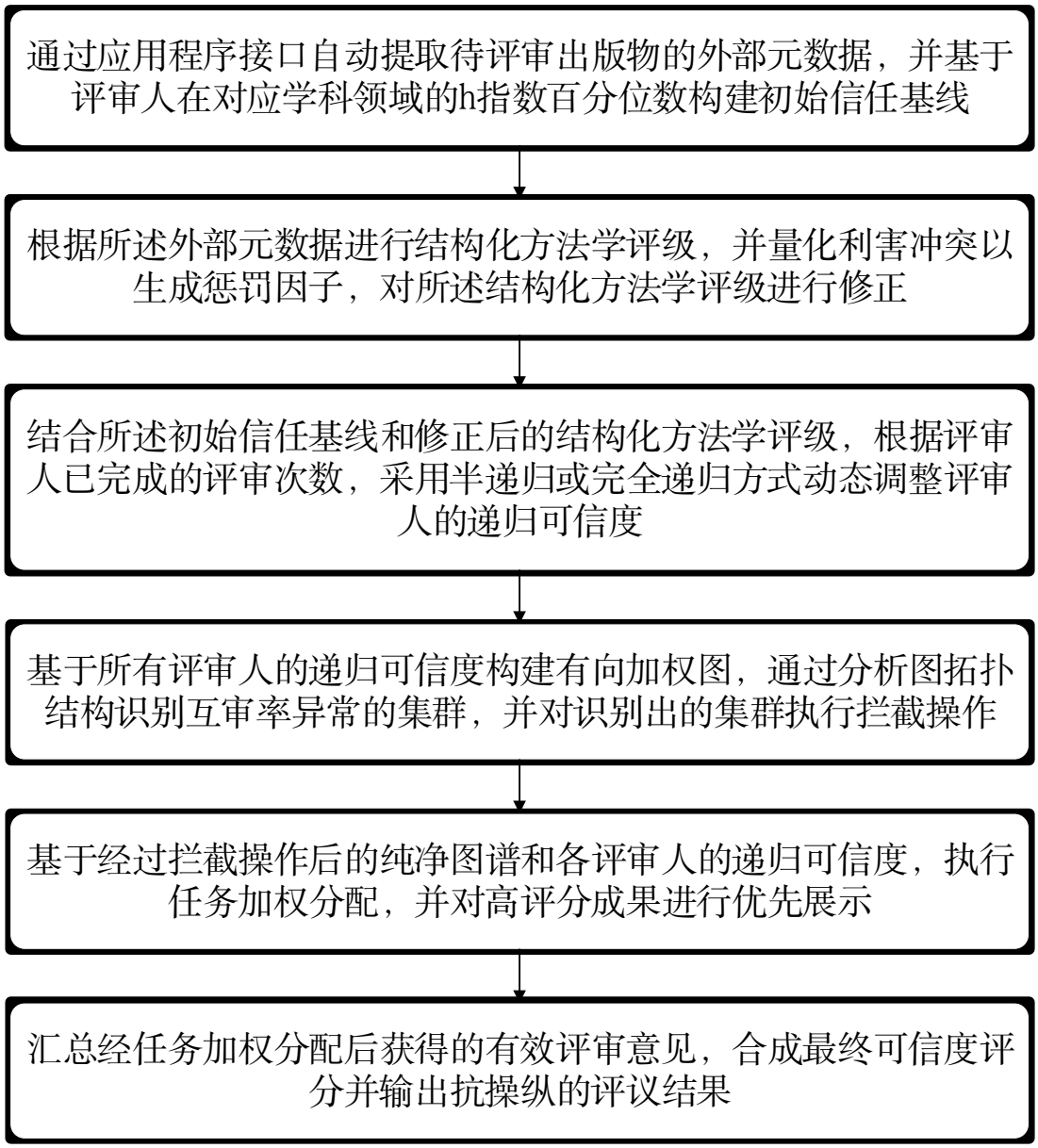
说明书摘要

本发明公开了一种出版物同行评议可信度评分方法及系统，通过利用应用程序接口自动提取外部元数据并结合 h 指数百分位数构建初始信任基线，通过量化利害冲突生成惩罚因子对结构化方法学评级进行动态修正，显著提升了利益冲突识别的精准度与评审意见的公正性；同时，采用基于评审次数的半递归与完全递归相结合的动态调整策略，使评审人的递归可信度能够随其历史表现自适应演进，基于纯净图谱执行的任务加权分配与高评分成果优先展示机制，形成了高信誉者获更多机会、优质成果获更高曝光的正向激励闭环，在消除人工录入误差的同时，全面增强了同行评议体系的抗操纵能力、透明度及整体公信力。

摘要附图


通过应用程序接口自动提取待评审出版物的外部元数据，并基于评审人在对应学科领域的h指数百分位数构建初始信任基线
根据所述外部元数据进行结构化方法学评级，并量化利害冲突以生成惩罚因子，对所述结构化方法学评级进行修正
结合所述初始信任基线和修正后的结构化方法学评级，根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度
基于所有评审人的递归可信度构建有向加权图，通过分析图拓扑结构识别互审率异常的集群，并对识别出的集群执行拦截操作
基于经过拦截操作后的纯净图谱和各评审人的递归可信度，执行任务加权分配，并对高评分成果进行优先展示
汇总经任务加权分配后获得的有效评审意见，合成最终可信度评分并输出抗操纵的评议结果

权利要求书

1.一种出版物同行评议可信度评分方法，其特征在于，包括：

通过应用程序接口自动提取待评审出版物的外部元数据，并基于评审人在对应学科领域的 h 指数百分位数构建初始信任基线；

根据所述外部元数据进行结构化方法学评级，并量化利害冲突以生成惩罚因子，对所述结构化方法学评级进行修正；

结合所述初始信任基线和修正后的结构化方法学评级，根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度；

基于所有评审人的递归可信度构建有向加权图，通过分析图拓扑结构识别互审率异常的集群，并对识别出的集群执行拦截操作；

基于经过拦截操作后的纯净图谱和各评审人的递归可信度，执行任务加权分配，并对高评分成果进行优先展示；

汇总经任务加权分配后获得的有效评审意见，合成最终可信度评分并输出抗操纵的评议结果。

2.根据权利要求 1 所述的一种出版物同行评议可信度评分方法，其特征在于，所述外部元数据至少包含作者机构分布、研究领域分类标签和基金资助信息，用于后续的利害冲突识别和评审人匹配。

3.根据权利要求 1 所述的一种出版物同行评议可信度评分方法，其特征在于，所述结构化方法学评级将评审内容分解为研究设计严谨性、数据处理规范性、结论推导逻辑性和创新点实质性四个维度，并对每个维度进行标准化评分。

4.根据权利要求 1 所述的一种出版物同行评议可信度评分方法，其特征在于，所述根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度，包括：

当评审人已完成的评审次数小于或等于 5 次时，采用半递归方式计算新的

递归可信度，其中历史递归可信度占主导权重；

当评审次数超过 5 次后，切换至完全递归方式，其中历史递归可信度的权重进一步增大。

5. 根据权利要求 1 所述的一种出版物同行评议可信度评分方法，其特征在于，所述有向加权图的边权重由互评双方的递归可信度均值与评分向量余弦相似度共同决定。

6. 根据权利要求 1 所述的一种出版物同行评议可信度评分方法，其特征在于，所述通过分析图拓扑结构识别互审率异常的集群，包括：

针对所述有向加权图中的每个评审人节点，基于该节点的局部聚类系数和指向图谱外部社区的出度生成异常得分；

并对通过社区发现算法识别出的每个子图，计算该子图内所有节点的异常得分的平均值，将所述平均值超过预设阈值的子图判定为互审率异常的集群。

7. 根据权利要求 1 所述的一种出版物同行评议可信度评分方法，其特征在于，所述执行任务加权分配，包括：

评审人获得评审任务的分配概率与其当前递归可信度成正比，并受其当前待处理任务队列长度的反向制约。

8. 根据权利要求 1 所述的一种出版物同行评议可信度评分方法，其特征在于，所述对高评分成果进行优先展示，包括：

高评分成果的优先展示权重由修正后的结构化方法学评级、参与评审的评审人递归可信度平均值以及预测的潜在引用率共同确定。

9. 根据权利要求 1 所述的一种出版物同行评议可信度评分方法，其特征在于，所述最终可信度评分通过对所有有效评审意见进行基于评审人递归可信度的几何加权平均计算得出。

10.一种用于实现如权利要求1-9任意一项所述方法的出版物同行评议可信度评分系统，其特征在于，包括：

元数据提取单元，用于通过应用程序接口自动提取待评审出版物的外部元数据；

初始化单元，用于基于评审人在对应学科领域的h指数百分位数构建初始信任基线；

评级修正单元，用于根据所述外部元数据进行结构化方法学评级，并量化利害冲突以生成惩罚因子，对所述结构化方法学评级进行修正；

递归计算单元，用于结合所述初始信任基线和修正后的结构化方法学评级，根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度；

图分析拦截单元，用于基于所有评审人的递归可信度构建有向加权图，通过分析图拓扑结构识别互审率异常的集群，并对识别出的集群执行拦截操作；

分配展示单元，用于基于经过拦截操作后的纯净图谱和各评审人的递归可信度，执行任务加权分配，并对高评分成果进行优先展示；

结果合成单元，用于汇总经任务加权分配后获得的有效评审意见，合成最终可信度评分并输出抗操纵的评议结果。

说明书

一种出版物同行评议可信度评分方法及系统

技术领域

本发明涉及审评技术领域，具体是涉及一种出版物同行评议可信度评分方法及系统。

背景技术

现有的出版物同行评议系统主要依赖人工筛选评审专家并基于简单的算术平均法计算评分。在实施过程中，系统通常缺乏对评审人初始信誉的客观量化手段，导致新晋学者或无历史数据的评审人难以获得公平的评审任务分配；同时，现有方案多采用静态规则来识别利益冲突，无法动态适应复杂的学术关系网络。

目前，现有技术缺乏针对恶意行为的深度防御机制，无法有效识别和拦截通过伪造身份、构建虚假互评网络（即 Sybil 攻击）来操纵评分结果的异常集群，使得最终评议结果易受人为干扰，缺乏公信力。

发明内容

本发明旨在提供一种出版物同行评议可信度评分方法及系统，全面增强了同行评议体系的抗操纵能力、透明度及整体公信力。

为达到以上目的，本发明采用的技术方案为：一种出版物同行评议可信度评分方法，包括：

通过应用程序接口自动提取待评审出版物的外部元数据，并基于评审人在对应学科领域的 h 指数百分位数构建初始信任基线；

根据所述外部元数据进行结构化方法学评级，并量化利害冲突以生成惩罚因子，对所述结构化方法学评级进行修正；

结合所述初始信任基线和修正后的结构化方法学评级，根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度；

基于所有评审人的递归可信度构建有向加权图，通过分析图拓扑结构识别互审率异常的集群，并对识别出的集群执行拦截操作；

基于经过拦截操作后的纯净图谱和各评审人的递归可信度，执行任务加权分配，并对高评分成果进行优先展示；

汇总经任务加权分配后获得的有效评审意见，合成最终可信度评分并输出抗操纵的评议结果。

优选的，所述外部元数据至少包含作者机构分布、研究领域分类标签和基金资助信息，用于后续的利害冲突识别和评审人匹配。

优选的，所述结构化方法学评级将评审内容分解为研究设计严谨性、数据处理规范性、结论推导逻辑性和创新点实质性四个维度，并对每个维度进行标准化评分。

优选的，所述根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度，包括：

当评审人已完成的评审次数小于或等于 5 次时，采用半递归方式计算新的递归可信度，其中历史递归可信度占主导权重；

当评审次数超过 5 次后，切换至完全递归方式，其中历史递归可信度的权重进一步增大。

优选的，所述有向加权图的边权重由互评双方的递归可信度均值与评分向量余弦相似度共同决定。

优选的，所述通过分析图拓扑结构识别互审率异常的集群，包括：

针对所述有向加权图中的每个评审人节点，基于该节点的局部聚类系数和指向图谱外部社区的出度生成异常得分；

并对通过社区发现算法识别出的每个子图，计算该子图内所有节点的异常

得分的平均值，将所述平均值超过预设阈值的子图判定为互审率异常的集群。

优选的，所述执行任务加权分配，包括：

评审人获得评审任务的分配概率与其当前递归可信度成正比，并受其当前待处理任务队列长度的反向制约。

优选的，所述对高评分成果进行优先展示，包括：

高评分成果的优先展示权重由修正后的结构化方法学评级、参与评审的评审人递归可信度平均值以及预测的潜在引用率共同确定。

优选的，所述最终可信度评分通过对所有有效评审意见进行基于评审人递归可信度的几何加权平均计算得出。

另一方面，本发明提出一种出版物同行评议可信度评分系统，包括：

元数据提取单元，用于通过应用程序接口自动提取待评审出版物的外部元数据；

初始化单元，用于基于评审人在对应学科领域的 h 指数百分位数构建初始信任基线；

评级修正单元，用于根据所述外部元数据进行结构化方法学评级，并量化利害冲突以生成惩罚因子，对所述结构化方法学评级进行修正；

递归计算单元，用于结合所述初始信任基线和修正后的结构化方法学评级，根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度；

图分析拦截单元，用于基于所有评审人的递归可信度构建有向加权图，通过分析图拓扑结构识别互审率异常的集群，并对识别出的集群执行拦截操作；

分配展示单元，用于基于经过拦截操作后的纯净图谱和各评审人的递归可信度，执行任务加权分配，并对高评分成果进行优先展示；

结果合成单元，用于汇总经任务加权分配后获得的有效评审意见，合成最终可信度评分并输出抗操纵的评议结果。

与现有技术相比，本发明的有益效果在于:

本发明通过利用应用程序接口自动提取外部元数据并结合 h 指数百分位数构建初始信任基线，有效解决了新晋评审人因缺乏历史数据而面临的冷启动难题，确保了评审任务分配的公平性与起点的客观性；在此基础上，通过量化利害冲突生成惩罚因子对结构化方法学评级进行动态修正，显著提升了利益冲突识别的精准度与评审意见的公正性；同时，采用基于评审次数的半递归与完全递归相结合的动态调整策略，使评审人的递归可信度能够随其历史表现自适应演进，避免了静态评分的滞后性；进一步地，通过构建有向加权图并分析图拓扑结构，系统能够主动识别并拦截互审率异常的集群，从根本上阻断了恶意用户通过伪造身份和构建虚假互评网络操纵评分的漏洞；最终，基于纯净图谱执行的任务加权分配与高评分成果优先展示机制，形成了高信誉者获更多机会、优质成果获更高曝光的正向激励闭环，在消除人工录入误差的同时，全面增强了同行评议体系的抗操纵能力、透明度及整体公信力。

附图说明

图 1 为本发明出版物同行评议可信度评分方法的流程图;

图 2 为本发明出版物同行评议可信度评分系统的框图。

具体实施方式

下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明的一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动的前提下所获得的所有其他实施例，都属于本发明保护的范围。

需要说明，若本发明实施例中有涉及方向性指示（诸如上、下、左、右、前、后……），则该方向性指示仅用于解释在某一特定姿态下各部件之间的相对位置关系、运动情况等，如果该特定姿态发生改变时，则该方向性指示也相应地随之改变。

另外，若本发明实施例中有涉及“第一”、“第二”等的描述，则该“第一”、“第二”等的描述仅用于描述目的，而不能理解为指示或暗示其相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。另外，若全文中出现的“和/或”或者“及/或”，其含义包括三个并列的方案，以“A 和/或 B”为例，包括 A 方案、或 B 方案、或 A 和 B 同时满足的方案。另外，各个实施例之间的技术方案可以相互结合，但是必须是以本领域普通技术人员能够实现为基础，当技术方案的结合出现相互矛盾或无法实现时应当认为这种技术方案的结合不存在，也不在本发明要求的保护范围之内。

请参阅图 1，本发明提供一种出版物同行评议可信度评分方法，包括：

通过应用程序接口自动提取待评审出版物的外部元数据，并基于评审人在对应学科领域的 h 指数百分位数构建初始信任基线；外部元数据至少包含作者机构分布、研究领域分类标签和基金资助信息，用于后续的利害冲突识别和评审人匹配。

通过自动提取机构分布、领域标签及基金资助等外部元数据，实现了利害冲突的精准量化与评审人的智能匹配，有效规避了人工筛选的主观偏差；同时利用 h 指数百分位数构建初始信任基线，解决了新晋学者因缺乏历史评审记录而导致的冷启动难题，确保了评审任务分配起点的客观公平，为后续动态调整可信度奠定了可靠的数据基础。

根据所述外部元数据进行结构化方法学评级，并量化利害冲突以生成惩罚因子，对所述结构化方法学评级进行修正；通过外部元数据构建结构化的方法

学评级体系，并精准量化利害冲突以生成动态惩罚因子，有效消除了因利益关联导致的评分偏差；这种对评级的实时修正机制不仅显著提升了评审意见的客观性与公正性，还从源头上遏制了因潜在利益输送而引发的学术不端风险，确保最终评议结果真实反映研究质量。

结合所述初始信任基线和修正后的结构化方法学评级，根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度；具体包括：当评审人已完成的评审次数小于或等于 5 次时，采用半递归方式计算新的递归可信度，其中历史递归可信度占主导权重；当评审次数超过 5 次后，切换至完全递归方式，其中历史递归可信度的权重进一步增大。

通过区分评审次数采用半递归与完全递归的自适应调整策略，既在评审初期保护新晋学者免受单次异常表现的不利影响，又随着评审经验的积累逐步强化历史表现的权重，实现了信誉评估从保守起步到稳健演进的平滑过渡；这种机制有效避免了单一评审结果对可信度的过度干扰，确保了递归可信度能够真实、动态地反映评审人的长期专业水准与稳定性，从而提升了整个评议体系的抗噪能力。

结构化方法学评级将评审内容分解为研究设计严谨性、数据处理规范性、结论推导逻辑性和创新点实质性四个维度，并对每个维度进行标准化评分。

过将评审内容解构为研究设计、数据处理、逻辑推导及创新实质四个核心维度并实施标准化评分，将主观的定性评价转化为客观的定量分析，有效消除了传统评审中因评审人个人偏好或经验差异导致的评分模糊与不一致性；这种多维度的精细化评估不仅全面覆盖了学术成果的质量关键要素，还显著提升了评审结果的透明度与可解释性，确保最终评级能够精准、立体地反映研究成果的真实价值。

基于所有评审人的递归可信度构建有向加权图，有向加权图的边权重由互评双方的递归可信度均值与评分向量余弦相似度共同决定。通过分析图拓扑结构识别互审率异常的集群，并对识别出的集群执行拦截操作；具体包括：针对所述有向加权图中的每个评审人节点，基于该节点的局部聚类系数和指向图谱

外部社区的出度生成异常得分；并对通过社区发现算法识别出的每个子图，计算该子图内所有节点的异常得分的平均值，将所述平均值超过预设阈值的子图判定为互审率异常的集群。

利用递归可信度与评分相似度构建动态加权图谱，创新性地结合局部聚类系数与社区出度特征，实现了对隐蔽性互审利益链的精准画像与自动拦截；通过量化节点异常得分并基于子图聚合阈值判定集群，有效识别了传统单点检测难以发现的团伙式串通评审行为，显著提升了学术评议系统的抗攻击能力与生态净化水平。

基于经过拦截操作后的纯净图谱和各评审人的递归可信度，执行任务加权分配，并对高评分成果进行优先展示；具体包括：高评分成果的优先展示权重由修正后的结构化方法学评级、参与评审的评审人递归可信度平均值以及预测的潜在引用率共同确定。

其中，执行任务加权分配包括：评审人获得评审任务的分配概率与其当前递归可信度成正比，并受其当前待处理任务队列长度的反向制约。

通过融合修正后的方法学评级、评审人历史可信度及潜在引用率构建多维展示权重，实现了优质学术成果的智能筛选与优先曝光，有效提升了高价值研究的可见度与传播效率；同时建立基于递归可信度正相关与任务队列负相关的动态分配机制，不仅优化了评审资源的配置效率，还激励高信誉评审人承担更多任务，从系统层面保障了评议流程的公平性、响应速度与整体运行效能。

汇总经任务加权分配后获得的有效评审意见，合成最终可信度评分并输出抗操纵的评议结果，最终可信度评分通过对所有有效评审意见进行基于评审人递归可信度的几何加权平均计算得出。

利用基于递归可信度的几何加权平均算法合成最终评分，通过指数级放大高信誉评审人的话语权并抑制低信誉或异常数据的干扰，从数学机制上构建了抵御恶意刷分与串通操纵的坚固防线；这种聚合方式显著增强了评议结果的鲁棒性与抗噪能力，确保最终输出能够真实、客观地反映学术成果的核心价值，从而维护了学术评价体系的公正性与公信力。

另一方面，本发明提出一种出版物同行评议可信度评分系统，如图 2 所示，包括：

元数据提取单元，用于通过应用程序接口自动提取待评审出版物的外部元数据；

初始化单元，用于基于评审人在对应学科领域的 h 指数百分位数构建初始信任基线；

评级修正单元，用于根据所述外部元数据进行结构化方法学评级，并量化利害冲突以生成惩罚因子，对所述结构化方法学评级进行修正；

递归计算单元，用于结合所述初始信任基线和修正后的结构化方法学评级，根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度；

图分析拦截单元，用于基于所有评审人的递归可信度构建有向加权图，通过分析图拓扑结构识别互审率异常的集群，并对识别出的集群执行拦截操作；

分配展示单元，用于基于经过拦截操作后的纯净图谱和各评审人的递归可信度，执行任务加权分配，并对高评分成果进行优先展示；

结果合成单元，用于汇总经任务加权分配后获得的有效评审意见，合成最终可信度评分并输出抗操纵的评议结果。

进一步的，上述各单元在执行时还用于实现上述一种出版物同行评议可信度评分方法的其他步骤，如下：

第一步：外部元数据自动提取与新评审员初始信任基线构建；

本实施例的首要环节在于解决数据录入繁琐及新评审员身份模糊的问题。采用自动化接口读取技术，直接从学术数据库及投稿系统中获取论文的核心属性，并据此为新加入的评审员建立初始的可信度基准。该步骤是整个评分体系的起点，其输出结果将直接作为后续所有计算的基础输入量。

具体而言，执行过程首先通过预设的应用程序接口，自动抓取待评审论文的元数据。这些元数据包含论文标题、摘要、关键词、作者机构分布、发表期刊的历史引用率以及研究领域的分类标签。系统利用自然语言处理中的语义匹

配逻辑，将提取出的关键词与领域知识图谱进行比对，确定论文所属的细分学科方向。随后，系统根据该学科方向检索当前注册在该领域的潜在评审员名单。对于新注册的评审员，系统不会赋予其默认的平均分值，而是依据其学术影响力指标进行差异化初始化。这里采用的核心指标是 h 指数百分位数。h 指数反映了评审员在特定领域内的持续产出能力，将其转化为百分位数可以消除不同学科间 h 指数绝对数值差异带来的不公平性。

为了将这一指标转化为可计算的初始信任值，系统执行如下映射操作。设某新评审员在对应学科领域的 h 指数为h_i，该学科内所有活跃评审员的 h 指数集合为$\{h_1, h_2, \ldots, h_N\}$。首先对该集合进行升序排列，得到有序序列。然后计算该评审员 h 指数在集合中的相对位置，即其百分位数p_i。计算公式表达为:

$$p_i = \frac{\mathrm{Rank}(h_i)}{N};$$

在此公式中，$\mathrm{Rank}(h_i)$表示h_i在集合中的排名（从小到大），N 为该学科内参与过有效评审的活跃评审员总数。该比值p_i的取值范围限定在 0 到 1 之间，代表了该评审员相对于同行的相对水平。若p_i接近 1，说明其处于顶尖梯队；若接近 0，则处于边缘。

基于计算得到的百分位数p_i，系统进一步构建初始信任基线值T_{init}。该值的设定并非线性比例，而是采用了对数压缩函数以平滑极端值的影响，避免个别超高 h 指数者获得不合理的初始优势。计算公式定义为:

$$T_{init} = T_{min} + (T_{max} - T_{min}) \cdot \frac{\ln(1 + \lambda \cdot p_i)}{\ln(1 + \lambda)};$$

式中，T_{min}代表初始信任基线的下限阈值，设定为 0.4，代表最低限度的可信度；T_{max}代表上限阈值，设定为 0.9，保留给极少数公认的权威专家；λ为调节系数，用于控制增长曲线的陡峭程度，此处取值为 5。该公式确保了即使 h 指数百分位数较低的新评审员也能获得一个非零的起始权重，从而具备参与初步评审的资格，同时高百分位数的评审员能获得更高的初始权重，体现了对学

术资历的认可。

在获取了论文的元数据特征向量以及新评审员的初始信任基线T_{init}后，系统将这些数据封装为结构化数据包，传递至下一处理环节。由于初始信任基线T_{init}是基于客观的 h 指数统计得出的，它避免了主观判断的偏差，为后续的递归计算提供了公平且稳定的起点。此外，提取的论文元数据中包含的领域分类信息，将在后续步骤中用于筛选匹配的评审任务，确保评审人仅在其专业范围内接受指派，从而保证评价的专业性。

第二步：结构化方法学评级与利害冲突量化惩罚；

本步骤聚焦于对评审内容的质量进行结构化拆解，并对可能存在的利益关联进行量化扣除。此环节是确保评分客观性的关键，它要求将抽象的评审意见转化为具体的数值指标，并剔除因利益冲突导致的主观偏差。

在执行过程中，系统首先调用预置的结构化评价模板。该模板将评审内容分解为四个核心维度：研究设计的严谨性、数据处理的规范性、结论推导的逻辑性以及创新点的实质性。每个维度下设若干细项指标，例如在“数据处理的规范性”下包含样本量充足性、统计方法适用性、异常值处理合理性等子项。评审员在提交评审意见时，需针对每个子项填写具体的评分等级或数值。系统自动解析这些文本或结构化输入，将其映射为标准化分数。

设第 j 个维度的原始评分为$S_{raw,j}$，经过归一化处理后的标准分为$S_{std,j}$。归一化过程采用极差法，即减去该维度的历史最低分并除以历史最高分与最低分之差，确保所有维度的分数落在 0 到 1 区间内。由此得到该评审员对当前论文的方法学综合评分M_{score}，其计算方式为各维度加权求和：

$$M_{score} = \sum_{j=1}^{4} w_j \cdot S_{std,j}$$；

式中，w_j为第 j 个维度的预设权重，满足$\sum w_j = 1$。通常情况下，研究设计严谨性和数据规范性被赋予较高权重，因为它们直接关系到结果的可靠性。

该公式将评审员的定性评价转化为定量的方法学得分，使得不同评审员之间的评价具有可比性。

然而，仅有高质量的评价并不足以保证公正，必须剔除利益冲突带来的干扰。系统利用第一步中提取的作者机构、合作网络及基金资助信息，结合评审员的个人档案数据进行交叉比对。比对规则包括：同一机构任职、过去五年内有共同发表论文、存在直系亲属关系、或受同一基金资助等情况。一旦检测到上述任一情形，即判定存在利害冲突。

为了量化这种冲突对评分的负面影响，系统引入利害冲突惩罚因子 $C_{penalty}$。该因子的大小取决于冲突的严重程度。若仅为同机构无直接合作，设为轻度冲突；若有共同发表或资金往来，设为重度冲突。惩罚因子的计算逻辑如下：

$$C_{penalty} = \begin{cases} 0 & \text{若无冲突} \\ \alpha \cdot D_{conflict} & \text{若存在冲突} \end{cases};$$

其中，$D_{conflict}$ 为冲突强度系数，轻度冲突取 0.1，重度冲突取 0.3；α 为调节参数，取值为 1.0。该惩罚因子将从方法学评分中扣除，但为了保证最终评分不为负，系统设定了一个最小截断值。修正后的方法学评分 M'_{score} 计算为：

$$M'_{score} = \max(0, M_{score} - C_{penalty});$$

此步骤输出的 M'_{score} 即为去除了利益干扰后的纯净方法学得分。该得分直接依赖于第一步提供的元数据（用于识别冲突）以及评审员填写的结构化反馈。通过将复杂的伦理判断转化为可计算的数值惩罚，系统有效地降低了人情分和恶意打压的可能性。

值得注意的是，此步骤产生的 M'_{score} 不仅是一个独立的评分结果，更是后续计算评审人自身可信度的重要输入变量。如果一个评审员长期给出高分但缺乏方法学支撑，或者其给出的评分频繁受到利害冲突惩罚，这些数据将被记录并用于调整其在第三步中的递归权重。因此，第二步不仅是对外部评审质量的校验，也是对内部评审人行为模式的初步画像，为构建动态的递归模型奠定了

坚实的数据基础。

第三步：研究设计指标加权与评审人递归可信度动态调整；

基于前两步所获得的论文元数据、新评审员初始信任基线 T_{init} 以及经利害冲突修正后的方法学评分 M'_{score}，本步骤进入动态评分构建阶段。此阶段旨在整合多维度指标，计算出一个能够反映评审人真实水平的递归可信度值，并根据评审次数动态调整计算策略。这是整个方案的枢纽，它将静态的初始值和单次评价结果融合为随时间演变的动态信任值。

首先，系统引入研究设计指标作为加权因子。不同的研究类型（如随机对照试验、观察性研究、理论推导等）对方法学的要求不同。系统根据第一步提取的论文类型标签，匹配相应的研究设计权重系数 k_{design}。例如，随机对照试验通常要求更严格的数据验证，其权重系数较高；而理论推导则侧重于逻辑自洽，权重系数适中。该系数将作用于方法学评分 M'_{score}，形成经过研究类型校正的评分分量 V_{design}：

$$V_{design} = k_{design} \cdot M'_{score};$$

其中，k_{design} 的取值范围在 0.8 至 1.2 之间，具体数值由领域专家委员会预先设定并存储在数据库中。这一步骤确保了不同性质的研究在评价体系中享有公平的基准，避免了用同一把尺子衡量所有研究类型的偏差。

系统计算当前的单轮评审贡献值 $R_{current}$。该值不仅包含校正后的方法学评分，还融合了评审人的初始信任基线 T_{init}。初始信任基线代表了评审人的历史积淀，而方法学评分代表了当下的表现。两者的结合方式采用几何平均与算术平均的混合形式，以平衡两者的重要性：

$$R_{current} = \beta \cdot T_{init} + (1 - \beta) \cdot V_{design};$$

式中，β 为平衡系数，取值范围为 0 到 1。在本实施例中，考虑到初始信誉的重要性，β 设定为 0.4，意味着初始信任占 40%的比重，而本次评审表现占 60%

的比重。这一设定鼓励评审员在保持良好历史记录的同时，必须在每次评审中付出实质性的努力。

递归可信度的更新，包括：系统根据评审人已经完成的评审次数 N，决定采用何种计算逻辑来生成新的递归可信度T_{new}。当评审次数 N 小于或等于 5 次时，系统采用半递归模式，即新值部分依赖于旧值，部分依赖于当前表现，以防止新评审员因偶然失误而被过早否定，或因运气好而虚高。当评审次数 N 超过 5 次后，系统切换至完全递归模式，此时新值几乎完全由历史累积表现决定，体现久经考验的原则。

具体的递归计算公式如下：$T_{new} = \gamma_n \cdot T_{old} + (1 - \gamma_n) \cdot R_{current}$；

在此公式中，T_{old} 为上一轮次结束后的可信度值（对于第一次评审，T_{old} 即为T_{init}）；γ_n为衰减系数，它是评审次数 n 的函数。当$n \leq 5$时，γ_n设定为 0.7，表示历史成绩仍占主导地位，新成绩仅占 30%；当 n>5 时，γ_n逐渐增大至 0.9，甚至更高，表示历史成绩的权重极大，新成绩仅起微调作用。这种动态调整机制既保护了新人的成长空间，又确保了资深评审员评分的稳定性。

通过上述公式计算得到的T_{new}，即为该评审人在完成当前任务后的最新递归可信度。该值将立即更新到评审员的状态库中，并作为下一次评审任务的初始输入。此外，系统还会同步计算该评审人的累计贡献度，用于后续的任务分配。

本步骤输出最新的递归可信度T_{new}，将直接传递给第四步，作为图分析模块的节点属性输入。因为只有在确定了每个评审节点的当前可信度状态后，才能有效地分析节点间的连接关系，识别潜在的异常集群。如果忽略这一步的动态更新，后续的图分析将失去准确的节点权重基础，无法准确识别出那些通过伪造高信誉来发起攻击的 Sybil 节点。

此外，本步骤中引入的 5 次评审切换机制，解决了冷启动与稳态运行的矛盾。在前 5 次评审中，系统给予较高的容错率和成长性，允许新评审员通过多次合格的表现逐步积累信誉；而在 5 次之后，系统则强调经验的沉淀，防止评

分波动过大。

第四步：基于图拓扑的互审异常检测与 Sybil 攻击集群拦截；

承接第三步生成的动态递归可信度T_{new}，本步骤利用图论分析技术，将评审人及其评审行为抽象为复杂的网络拓扑结构。通过深度挖掘节点间的连接关系，系统能够自动识别出那些试图通过伪造身份、构建虚假互评网络来操纵评分结果的 Sybil 攻击集群。这一步骤是保障系统安全性的核心防线，其输入完全依赖于前一步骤计算出的每个节点的实时信誉状态。

在实施过程中，系统首先构建一个有向加权图 G=(V,E)。其中，集合 v 代表所有参与过评审的评审员节点，集合 E 代表评审关系边。若评审员 i 对论文 j 进行了评审，且该论文由评审员 k 撰写或由其关联作者提交，则建立一条从 i 指向 k 的潜在依赖边；更关键的是，若评审员 i 长期对评审员 k 的论文给予高分，而评审员 k 也频繁对评审员 i 的论文给予高分，则两者之间形成高强度的双向交互边。边的权重W_{ij}定义为两人在过去一段时间内的互评一致性系数，结合双方当前的递归可信度T_i和T_j进行加权计算：

$$W_{ij} = \frac{1}{2}\left(\frac{T_i + T_j}{2}\right) \cdot \cos(\theta_{ij})$$；

式中，$\cos(\theta_{ij})$表示两人评分向量之间的余弦相似度，用于量化评分倾向的一致性程度；T_i和T_j分别为第三步输出的最新递归可信度值。该公式确保了高信誉节点之间的互评具有更高的权重，同时也利用了评分的一致性特征来过滤随机噪声。

构建完成图 G 后，系统执行社区发现算法，识别图中紧密连接的子图结构。正常的学术网络通常呈现稀疏分布，评审关系多为单向或弱双向，且社区边界清晰。然而，Sybil 攻击往往表现为高密度的互评小团体，即一群节点内部相互打分极高，而对外部节点保持冷漠或恶意打压。为了捕捉这种异常，系统计算每个节点的中心度指标及局部聚类系数。对于疑似攻击集群中的节点 v，其异常

得分$S_{abnormal}$定义为:

$$S_{abnormal}(v) = \alpha \cdot C_{local}(v) - \beta \cdot D_{out}(v)$$;

其中，$C_{local}(v)$为该节点在子图中的局部聚类系数，衡量其邻居节点之间相互连接的程度；$D_{out}(v)$为该节点指向外部正常社区的出度（即给非圈子内人员打分的次数）；α和β为调节参数，分别赋予内部连接紧密度和外部隔离度的权重。在本实施例中，α取值为 0.6，β取值为 0.4，以突出内部闭环的特征。

当某个子图内所有节点的$S_{abnormal}$平均值超过预设阈值$\Theta_{threshold}$时，系统判定该子图为异常互审集群。此时，系统立即触发拦截机制。对于被标记为异常集群的成员，其当前及未来一段时间内产生的所有评分将被暂时冻结，不再计入最终的可信度更新中。同时，这些节点的递归可信度T_{new}将被强制重置为初始基准值，切断其通过历史积累获得的信用优势。这一操作直接依赖于第三步建立的动态信任体系，因为如果缺乏实时的T_{new}作为节点属性，图分析将无法区分正常的高频互评（如长期合作团队）与恶意的攻击集群。

此外，系统还会记录被拦截节点的原始数据特征，包括其注册 IP 地址段、设备指纹以及关联的机构信息，用于生成防御报告。这一步骤不仅清除了当前的攻击威胁，还通过阻断异常路径，防止了错误评分向后续流程扩散。经过此步骤处理后的图谱G'，剔除了噪点和恶意节点，保留了真实可靠的评审关系网络。

该步骤输出清洗后的图谱G'以及被剔除的异常节点列表，将直接传递给第五步。第五步需要基于这个纯净的网络环境，进行任务分配和成果展示策略的制定。如果在存在大量 Sybil 攻击的情况下进行任务分配，可能会导致优质论文被恶意压低分数，或者劣质论文因互保而获得高曝光。因此，第四步的拦截效果直接决定了后续激励闭环的有效性和公正性。通过图拓扑分析，系统将隐形的操纵行为显性化，用数学逻辑构建了坚实的防火墙，确保了同行评议环境的纯净度。

第五步：任务加权分配与高评分成果优先展示策略；

基于第四步清理后的纯净图谱G'以及各节点的最新递归可信度T_{new}，本步骤专注于资源的优化配置与成果的差异化展示。其核心逻辑在于将评审任务精准地指派给高信誉且无利益冲突的节点，同时将经过验证的高质量研究成果优先推向公众视野，从而形成正向的价值导向。这一步骤是整个系统的输出端，直接面向用户和应用场景，其输入完全依赖于前序步骤积累的动态信誉数据和安全的网络结构。

在执行任务分配时，系统首先筛选出待评审论文的候选评审员池。该池中的成员必须满足两个硬性条件：一是其所属的学科领域与论文标签匹配；二是其在图谱G'中未被标记为异常节点，且其递归可信度T_{new}高于设定门槛。在此基础上，系统引入任务分配概率模型。对于符合条件的评审员 i，其获得某篇论文评审任务的概率$P_{assign,i}$与其当前的递归可信度 $T_{new,i}$成正比，同时受到其近期工作负荷的反向制约：

$$P_{assign,i} = \frac{T_{new,i}^{\mu}}{\sum_{k \in \text{Pool}} T_{new,k}^{\mu}} \cdot \frac{1}{1 + \lambda_L \cdot L_i};$$

式中，μ为信誉放大系数，取值大于 1（例如 1.5），用于强化高信誉者的优先权；L_i为评审员 i 当前的待处理任务队列长度；λ_L为负荷调节因子，用于平衡工作量。该公式确保高信誉者获得更多机会，但不会导致其过度疲劳，从而维持评审质量。这种分配策略直接利用了第三步生成的动态T_{new}值，使得信誉越高者承担的任务越重，责任越大，体现了权责对等的原则。

在完成评审任务并收集结果后，系统进入成果展示环节。对于最终评分较高的论文，系统将其标记为高可信度成果，并在平台首页、推荐流及相关学科分类页中进行优先展示。展示权重的计算综合考虑了论文本身的引用潜力、方法学评分M'_{score}以及评审团的整体置信度。设某论文的展示指数$I_{display}$为：

$$I_{display} = \omega_1 \cdot M'_{score} + \omega_2 \cdot \bar{T}_{reviewers} + \omega_3 \cdot C_{potential};$$

其中，$\bar{T}_{reviewers}$ 为参与该论文评审的所有有效节点的递归可信度平均值；$C_{potential}$ 为基于摘要和关键词预测的潜在引用率；$\omega_1, \omega_2, \omega_3$ 为归一化权重系数，且三者之和为 1。该公式表明，一篇论文能否获得高曝光，不仅取决于其自身质量，还取决于评审它的人是否值得信赖。这直接将评审人的信誉与成果的传播力挂钩，倒逼评审人珍惜自己的声誉资本。

为了进一步增强激励效果，系统建立了信誉-曝光联动机制。对于连续多次给出高质量评审意见（即其评审结果与最终录用决策高度一致，且未受利害冲突影响）的评审员，其个人主页将获得特殊的标识，其历史评审记录将公开可见，供学术界监督。这种公开透明的做法，利用了社会心理学中的声誉效应，促使评审员自发地维护自身形象。该机制的运行基础正是第四步所确立的安全网络，只有在排除了作弊干扰后，公开的信誉记录才具有公信力。

本步骤的输出包括两份关键清单：一份是下一轮次的具体任务分配表，另一份是高可信度成果展示列表。任务分配表将直接驱动第六步的执行，决定谁去评审哪篇文章；而展示列表则是整个评议过程的最终产出，直接影响学术界的关注流向。通过这种精细化的分配和展示策略，系统将抽象的信誉数值转化为具体的行动指令和流量资源，实现了从评价到应用的跨越。

值得注意的是，此步骤是直接基于当前的静态快照（即清洗后的图谱和最新的信誉值）进行一次性决策。这种确定性执行避免了因反复调整带来的不确定性，保证了系统在每一轮循环中的可预测性和稳定性。任务分配和成果展示的逻辑链条清晰：高信誉带来高概率分配，高分配带来高质量评审，高质量评审带来高可信度成果，高可信度成果带来高曝光。

第六步：最终可信度评分合成与抗操纵结果输出；

承接第五步生成的任务分配结果和高可信度成果展示列表，本步骤作为整个流程的终点，负责将所有分散的评估数据进行最终合成，输出具有法律效力的同行评议结论。该步骤不引入新的学习或优化过程，而是严格依据前五个步骤积累的所有数据，通过确定的数学规则计算出最终的出版物可信度评分，并

生成防篡改的评估报告。

在执行过程中，系统首先汇总所有针对该出版物的有效评审意见。这些意见来自第五步分配的任务，且已经通过了第四步的异常节点清洗。对于每篇论文，系统计算其综合可信度评分S_{final}。该评分不是简单的算术平均，而是采用基于信誉权重的几何加权平均，以突显高信誉评审员意见的主导地位：

$$S_{final} = \prod_{i=1}^{m} (S_i)^{w_i}$$；

式中，S_i为第 i 位评审员给出的标准化评分；w_i为该评审员的归一化权重，其值正比于其当前的递归可信度$T_{new,i}$，即$w_i = \frac{T_{new,i}}{\sum_{j=1}^{m} T_{new,j}}$；m 为有效评审员的总数。该公式利用几何平均的特性，使得任何一个低信誉评审员的极端低分都能显著拉低总分，从而增强了系统对恶意差评的抵抗力；同时，高信誉评审员的高分也能更有效地推高总分，体现了对权威意见的尊重。

随后，系统生成最终的评估报告。报告中不仅包含S_{final}数值，还详细列出了支撑该分数的各项指标来源：包括初始信任基线T_{init}的来源、方法学评分M'_{score}的构成、利害冲突惩罚的具体情况、以及图谱分析中对该论文相关评审网络的定性描述。报告的生成过程完全自动化，数据均源自前序步骤的计算结果，杜绝了人工干预的可能性。为了确保数据的不可篡改性，系统会对报告内容生成数字摘要，并存储在分布式账本中，供第三方审计。

最后，系统根据S_{final}的大小，执行分级处理策略。对于高分论文，直接推送至高可信度成果展示列表并进入发表流程；对于中等分数的论文，启动复审程序，指派更高信誉的专家进行二次评估；对于低分论文，则生成详细的拒稿理由报告，并归档至数据库以备查证。这一系列动作完全由S_{final}触发，逻辑简单直接，不涉及任何复杂的判断或策略调整。

以上所述仅为本发明的示例性的实施方式，并非因此限制本发明的专利范围，凡是在本发明的技术构思下，利用本发明说明书及附图内容所作的等效结

构变换，或直接/间接运用在其他相关的技术领域均包括在本发明的专利保护范围内。

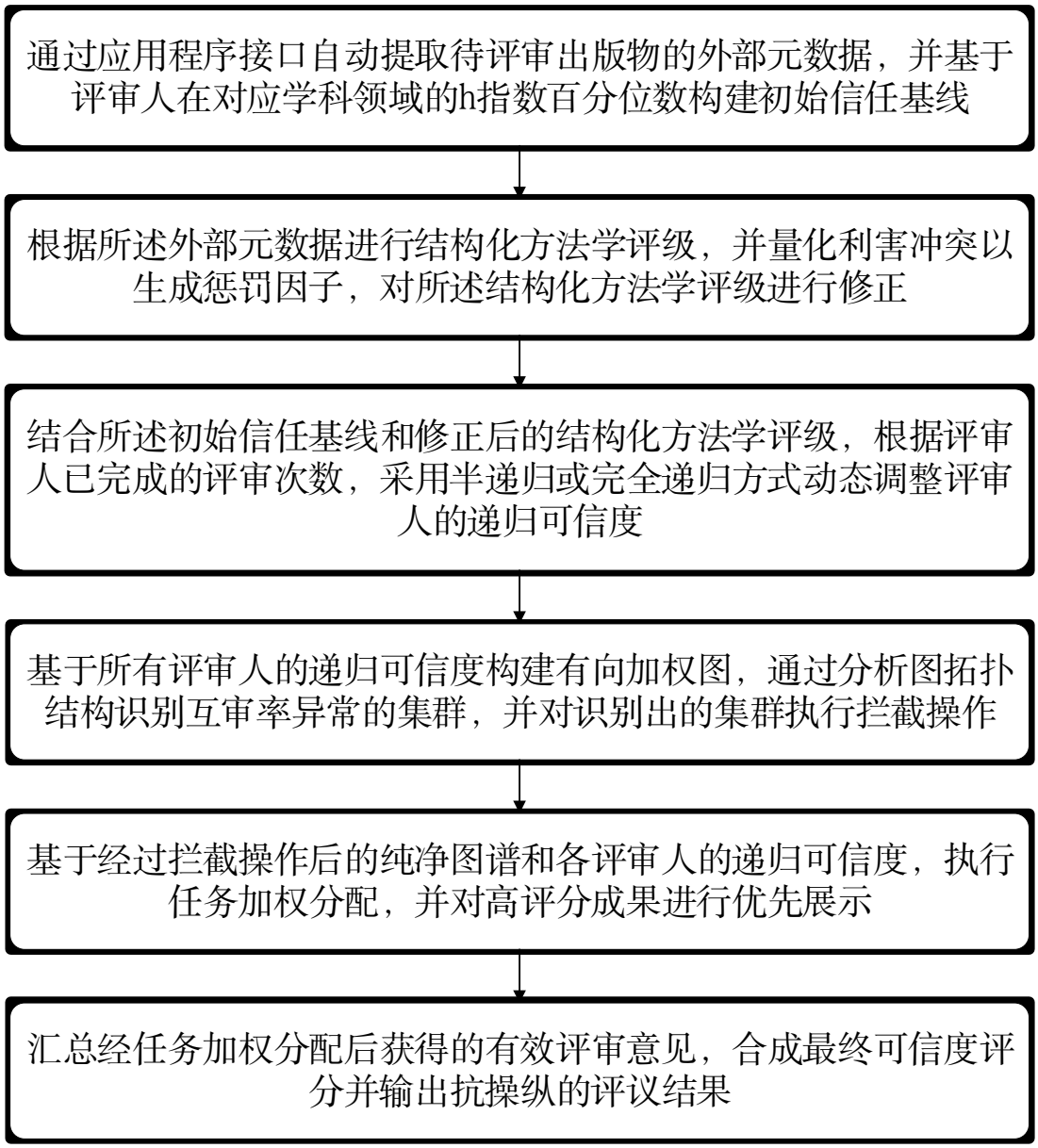

通过应用程序接口自动提取待评审出版物的外部元数据，并基于评审人在对应学科领域的h指数百分位数构建初始信任基线
根据所述外部元数据进行结构化方法学评级，并量化利害冲突以生成惩罚因子，对所述结构化方法学评级进行修正
结合所述初始信任基线和修正后的结构化方法学评级，根据评审人已完成的评审次数，采用半递归或完全递归方式动态调整评审人的递归可信度
基于所有评审人的递归可信度构建有向加权图，通过分析图拓扑结构识别互审率异常的集群，并对识别出的集群执行拦截操作
基于经过拦截操作后的纯净图谱和各评审人的递归可信度，执行任务加权分配，并对高评分成果进行优先展示
汇总经任务加权分配后获得的有效评审意见，合成最终可信度评分并输出抗操纵的评议结果

图 1

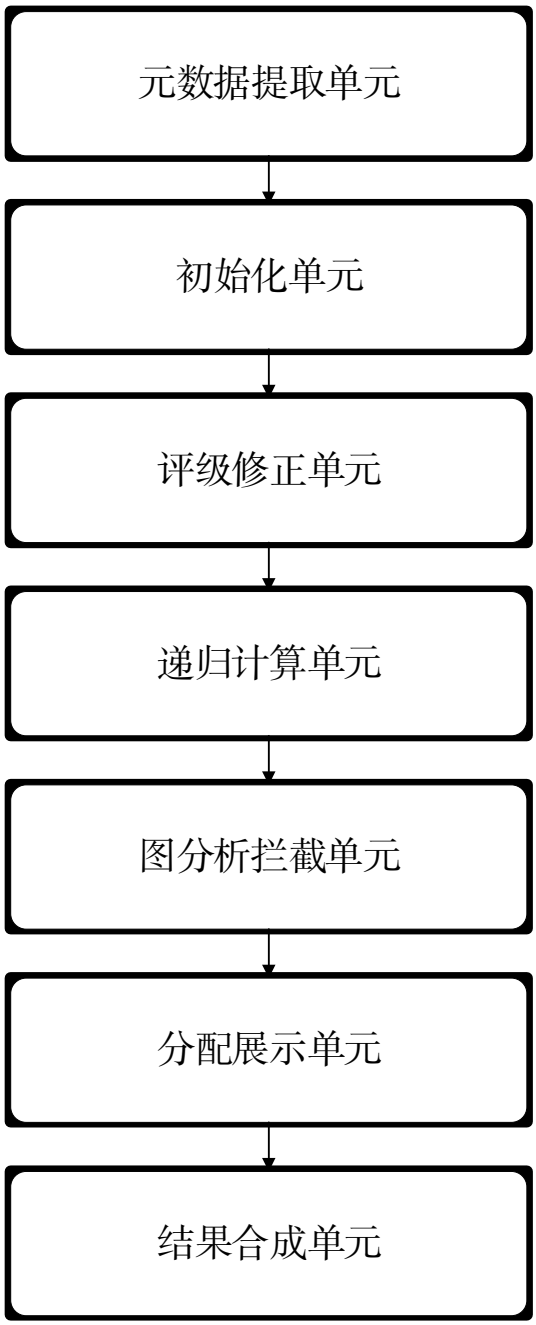


图 2