**河南大学2025届本科毕业设计**

**文献翻译**

**GroupLens：An Open Architecture for Collaborative Filtering of Netnews**

**GroupLens：一种用于网络新闻协同过滤的开放式架构**

|  |  |
| --- | --- |
| 学 生 姓 名： | 杜静霞 |
| 学 生 学 号： | 2114050011 |
| 所 在 学 院： | 计算机与信息工程学院 |
| 所 学 专 业： | 数据科学与大数据技术（明德计划） |
| 导 师 名 称： | 王雅娣 |
| 导 师 职 称： | 副教授 |

2025年04月21日

**摘要：**协同过滤帮助人们根据他人的意见做出选择。GroupLens是一个用于网络新闻协同过滤的系统，旨在帮助人们从海量的文章中找到自己喜欢的文章。新闻阅读器客户端会显示预测的评分，并让用户在阅读后轻松对文章进行评分。评分服务器被称为“更好的位局”，负责收集和传播评分。评分服务器基于这样一种启发式方法来预测评分：过去意见一致的人很可能再次意见一致。用户可以通过使用假名来保护自己的隐私，而不会降低评分预测的有效性。整个架构是开放的：可以独立开发替代的新闻客户端和“更好的位局”软件，并且可以与我们开发的组件进行互操作。

**关键词：协同过滤；信息过滤；电子公告板；社会过滤；Usenet；网络新闻；用户模型；信息选择性传播**

**引言**

计算机网络使人们能够组成跨越地理障碍的兴趣小组。公告板就是一个重要的机制。作者不是直接向已知的人群发表文章，而是将文章发布到新闻组中，任何对该主题感兴趣的人都可以在这个公共场所发表文章。Usenet网络新闻系统创造了一种世界上任何地方都可以使用的单一公告板的假象。它可以传播文章，使得在世界任何地方发布的内容都能在经过一定延迟后为其他所有人所获取。

最近的统计显示，新闻组数量超过8000个，平均每天流量超过100MB。新闻组包含公告、问题和讨论。在讨论（通常称为线程）中，一篇文章会引起其他几篇文章的回复，每篇文章也会引起回复。1994年1月24日对网络新闻参与情况的估计表明，在前两周有超过14万人发表了文章。还有更多的“潜伏者”阅读但不发表文章。显然，很多人都从这些公告板上获得了价值。

事实上，网络新闻的快速广播特性和广泛的读者群已经重塑了计算机领域的工作方式。系统管理员依靠网络新闻来了解最新的开发工作、安全漏洞和漏洞修复。研究人员依靠网络新闻来了解新的研究方向和会议间隔期间的重要成果。还有许多人使用网络新闻与世界各地的其他人保持联系，了解新书、新食谱、新音乐以及其他城市的生活情况。多年来，网络新闻已成为计算机用户之间共享信息的主要媒介。

即便如此，使用网络新闻组的体验并不尽如人意。几乎所有人都抱怨信噪比太低。除了少数发表回复的人，写作者很难判断自己的评论是否受到重视。有些人似乎不在乎读者的兴趣，只在乎自己的写作权。此外，每个人口味不同，因此没有一篇文章能吸引新闻组的所有读者。每个读者最终都要浏览大量新闻文章才能找到几篇有价值的文章。通常，读者会发现这个过程太令人沮丧，从而完全停止阅读网络新闻。

网络新闻提供了两种机制，帮助读者将注意力集中在可能令其感兴趣的条目上。首先，将公告板划分为新闻组，使读者能够专注于少数几个主题。当某个新闻组中的帖子数量过多时，通常会将其拆分为两个或多个具有可识别子主题的新闻组。其次，有些新闻组是经过审核的。向这些新闻组尝试发布的帖子会自动转发给审核员，由其决定这些帖子是否适合该新闻组。只有获得审核员批准的帖子才会在Usenet上传播。

此外，阅读网络新闻的软件包（以下简称新闻客户端）还提供了其他机制来减轻读者的负担。首先，大多数新闻客户端都会显示新闻组中每条信息的作者和主题行摘要。然后，用户可以指出她想阅读哪些文章。其次，大多数新闻客户端会同时显示某个讨论主题中的所有文章。有些客户端最初只显示每个主题中的第一篇文章，让用户可以快速浏览当前的讨论主题。第三，有些新闻客户端提供 “杀文件”。“杀文件”可识别特定用户不感兴趣的文本字符串。如果用户将某篇文章的主题行放入 “杀文件”，就不会再显示该主题的其他文章。如果用户将作者姓名放入杀文件，则不会再显示该作者的其他文章。最后，有些新闻阅读器还提供字符串搜索功能。如果用户对提及 “协同过滤 ”的文章特别感兴趣，新闻客户端就能找到它们。

GroupLens提供了一种新的机制，帮助人们关注有趣的文章。它借鉴了一个看似简单的理念：对过去文章主观评价一致的人，很可能在未来再次达成一致。阅读文章后，用户会给文章打分。GroupLens通过两种方式使用这些评分。首先，它将评分关联起来，以确定哪些用户的评分彼此最为相似。其次，它会根据相似用户的评分来预测用户对新文章的喜爱程度。GroupLens的核心是一个开放式架构，包括用于输入评分和显示预测的新闻客户端，以及用于发布评分和提供预测的评分服务器。

**相关工作**

信息过载和信噪比低的普遍问题在研究文献中受到了广泛关注。我们使用“信息过滤”一词来泛指寻找所需的信息（过滤进来）和剔除不需要的信息（过滤出去），但相关工作也出现在信息检索和选择性信息传播的标签下[2]。此外，关于代理[12,13]、用户建模[1,9]、知识机器人[8]和中介[21]的研究也探讨了代表用户执行信息过滤的半自主计算机程序。

Malone等人[13]根据过滤技术在预测用户对文章的反应时所利用的信息源，将过滤技术分为认知、社会和经济三类。这三个类别为文献提供了一个有用的路线图。

认知或基于内容的过滤技术根据文档中的文本选择文档。例如，新闻客户端提供的杀文件和字符串搜索功能就是内容过滤技术。甚至将网络新闻划分为新闻组也是一个原始的例子，因为读者只关注“新闻组：”字段中包含特定文本字符串的文章。

也可以使用其他基于内容的过滤技术。要包含或删除哪些文本的配置文件可能比字符串集合更复杂。例如，字符串可以与布尔运算符AND、OR和NOT结合使用。另外，配置文件还可以由权重向量组成，权重表示一组术语中每个术语的相对重要性[4, 5, 16]。

有些内容过滤技术会根据用户是否喜欢当前配置文件所选文章的反馈信息自动更新配置文件。信息检索研究将这一过程称为相关性反馈[17]。更新技术可以借鉴贝叶斯概率[2]、遗传算法[18]或其他机器学习技术。

社会过滤技术根据人与人之间的关系及其主观判断来选择文章。将作者姓名放在杀文件中就是一个简单的例子。更复杂的技术还可以过滤掉以前曾与异议者合著过论文的人的文章。

基于其他读者主观评价的协同过滤是一种更有前途的社会过滤形式。在判断文本的相关性时，人类读者不会像计算机一样在同义词、多义词和上下文方面遇到困难。此外，人们还可以从质量、权威性或尊重程度等其他方面对文本进行判断。有版主的新闻组采用的是最原始的协同过滤形式，根据版主一个人的评价为所有潜在读者选择文章。

Tapestry系统[6]对主观评价的利用更为复杂。虽然它不是专门为网络新闻而设计的，但它可以过滤所有输入的信息流，包括网络新闻。许多人都可以发布评价，而不仅仅是一个主持人，读者可以选择关注哪些评价者。评价可以包含文本，而不仅仅是二进制的接受-拒绝建议。此外，过滤器可以将基于内容的标准和主观评价结合起来。例如，读者可以要求乔评价过的包含“CSCW”一词的文章，而且评价中包含“优秀”一词。

我们的工作在精神上与Tapestry类似，但在两个方面进行了扩展。首先，Tapestry是一个单体系统，旨在单个网站内共享评价。我们在网站之间共享评价，而且我们的架构是开放的，可以创建新的新闻客户端和评价服务器，以不同的方式使用评价。其次，Tapestry不包含任何汇总查询。我们实施的评级服务器会根据多个评估者过去评级的相关性，汇总他们的评级。读者不需要事先知道使用谁的评价，事实上甚至不需要知道实际使用的是谁的评价。在GroupLens中，用假名输入的评价与签名的评价一样有用。

Maltz开发了一种系统，可以汇总每篇网文的所有评价，为每篇网文确定一个分数[14]。相比之下，GroupLens可根据每个用户的情况定制分数预测，从而满足不同用户的兴趣和品味。由于评分服务器可以交换多个用户对一篇文章的评分摘要，而不是单个评分，因此Maltz的方案比我们的方案扩展性更好，但功能却有所减少。

协同过滤中使用的主观评价可能是隐性的，而不是显性的。Read Wear和Edit Wear[7]根据其他用户与某件作品的互动情况来引导用户。GroupLens新闻客户端监控用户阅读每篇文章的时间，但我们的评分服务器在预测分数时还没有使用该信息。

经济过滤技术根据制作和阅读文章的成本和收益来选择文章。例如，马龙（Malone）认为，群发邮件每个收件人的生产成本较低，因此应给予较低的优先级。将这一观点应用到网络新闻中，新闻客户端就可以过滤掉在多个新闻组中交叉发布的文章。更激进的方案可以向读者支付考虑文章的费用（真金白银或声誉积分），并根据读者对文章的喜爱程度向制作者支付费用。

斯托多尔斯基提出了一种结合社会和经济过滤技术的方案[19]。他提出了一种在线出版物，在这种出版物中，出版决定权最终掌握在作者手中。在初步出版期间，其他读者可以发布对文章的评分。然后，作者可以撤回文章，以避免因发表不受欢迎的文章而付出声誉损失的代价。

**概述**

本文的GROUPLENS部分介绍了GroupLens架构及其演变。正在进行的实验部分介绍了正在准备的更大规模的架构测试。社会影响部分探讨了GroupLens可能引发的网络新闻使用方面的社会变化。

**GROUPLENS**

GroupLens是一个分布式系统，用于收集、传播和使用一些用户的评分来预测其他用户对文章的兴趣。它包括Macintosh和Unix计算机的新闻阅读客户端，以及“更好位局”，即收集评分并进行预测的服务器。整体架构和特定组件都是通过反复设计和试验测试逐步形成的，以实现以下目标：

开放性：目前常用的新闻客户端有几十种，每种都在其用户群中拥有很高的人气。这些客户端中的任何一个或全部都可以改编成GroupLens。GroupLens还允许创建其他更好的位局，以不同的方式使用评级来预测用户对新闻文章的兴趣。

易用性：评分易于形成和交流，预测易于识别和解释。这最大程度地减轻了协同过滤给用户带来的额外负担。

兼容性：该架构与现有的新闻机制兼容。兼容性降低了用户利用新工具的开销，简化了将其引入网络新闻的过程。

可扩展性：随着用户数量的增加，预测的质量应该会提高，速度也不会降低。如果GroupLens的规模变得非常大，一个潜在的增长限制将是评级的运输和存储。

隐私性：有些用户不愿意让别人知道他们阅读了哪些类型的文章，喜欢哪些类型的文章。GroupLens中的更好位局可以有效利用评分，即使评分是以化名提供的。

**概述**

Usenet包括互联网站点和UUCP站点。通常情况下，一个站点会声明一台机器作为其新闻服务器。每个站点的用户在自己的计算机上调用新闻客户端，并连接到新闻服务器，以检索新闻文章。用户还可以撰写新文章，并通过新闻客户端将其发布到新闻服务器上。

当用户发布文章时，文章会从撰写文章的新闻客户端传送到本地新闻服务器，再从本地新闻服务器传送到附近站点的新闻服务器。文章离开发起站点后，会在整个Usenet上传播，从一个站点跳到另一个站点。由于没有对传播过程进行集中协调，一篇文章可能会通过多条路径到达一个站点。不过，由于文章具有全球唯一的标识符，而且一旦发布就不会更改，因此任何网站都能识别出文章的重复副本，避免将其传递出去。Lotus Notes也使用类似的发布程序[10]。图1概述了网络新闻架构。

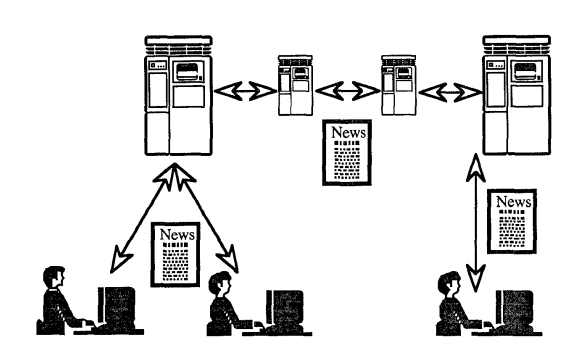


图1 网络新闻架构。新闻文章在新闻服务器之间跳转。新闻客户端在其所在站点连接到新闻服务器，并向用户展示文章。

如图2所示，GroupLens为网络新闻架构添加了一种新型实体-更好位局。更好位局提供预测用户喜欢文章程度的分数，并在用户阅读文章后从新闻客户端收集评分。更好位局还使用特殊的新闻组相互分享评分，以便在不同网站的用户之间进行协同过滤。本节的其余部分将追溯评级的创建、发布和使用过程，并描述它们如何实现开放、易用、兼容、可扩展性和隐私等设计目标。

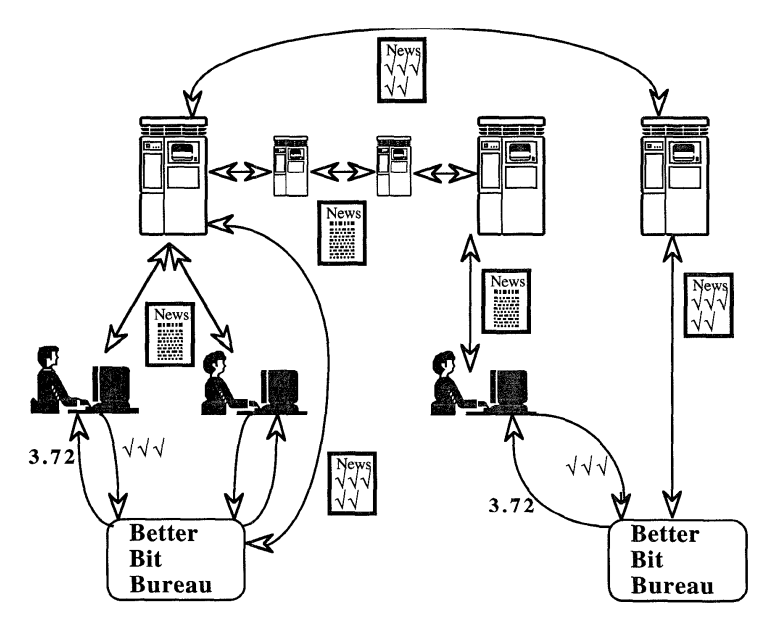


图2 GroupLens架构。“更好位局”从客户端收集评分，通过新闻服务器进行信息交互，并用这些评分生成数值评分预测，再发送给客户端。客户端连接到本地新闻服务器，并且可以连接到使用相同或不同新闻服务器的“更好位局”。

**输入评分**

在GroupLens中，评分是一个从1到5的数字，用户还可以选择补充阅读文章的秒数。我们鼓励用户根据自己对文章的喜爱程度进行评分，5分最高，1分最低。用户可以选择一个假名来与她的评分联系起来，这个假名可能与她发布新闻文章时使用的名字不同。这样既能检测到两个评分来自同一个人，又能防止检测到这个人的确切身份。

GroupLens对评分形式和含义的选择只是丰富设计空间中的一种可能性。可以从很多方面对文章进行评分：对主题的兴趣、写作质量、作者的权威性等。可以向读者征集多个维度的单独评分，而不是单一的综合评分。可以输入自由文本评分而不是数字。可以要求读者预测他们认为其他读者会有多喜欢某篇文章，而不是报告他们自己有多喜欢这篇文章。评分可以只限于正面评价，也可以只限于负面评价。隐私程度也可以是多种多样的，可以是完全匿名的，也可以是经过认证的签名。

事实上，早期的Macintosh新闻客户端[20]采用的评分形式与当前的GroupLens架构截然不同。由于网络新闻的信噪比非常低，因此只需指出好的文章即可，用户只需输入认可和正面评价。读者认可的文章是他们认为已知小团体中的其他人会喜欢的文章。最后，读者用自己的真实姓名签署认可，这样其他人就可以选择某个朋友认可的所有文章。

在斯伦贝谢研究实验室对早期的认可机制进行的试点测试表明，七人小组的规模可能不足以充分发挥协同过滤的作用。当我们考虑扩大群组规模时，我们认为有些用户会不太愿意在自己的评分上签名，而且用户会越来越难知道群组中其他人会喜欢什么文章。

试点测试还强化了尽可能方便用户输入批注的重要性。要进行批注，用户必须从下拉菜单中进行选择，等待窗口打开，选择性地在窗口中输入文字，然后关闭窗口。如果用户不输入任何文字，整个过程只需要几秒钟，但仍然比正常情况下进入下一篇文章所需的时间长很多。

在GroupLens系统中，我们尽可能简化了评分的输入。我们修改了三个新闻客户端：UNIX机器上的Emacs Gnus和NN，以及Macintoshes机器上的NewsWatcher。在每种情况下，评分的输入都符合新闻客户端的整体范式。例如，在修改后的NewsWatcher中，数字1至5在用户阅读文章时作为可选按钮出现（图3），用户还可以输入数字作为这些按钮的快捷键。在Gnus中，没有按钮显示，但读者仍可直接键入评分。在NN中，读者首先键入字母“v”（进入“评分模式”），然后键入评分。

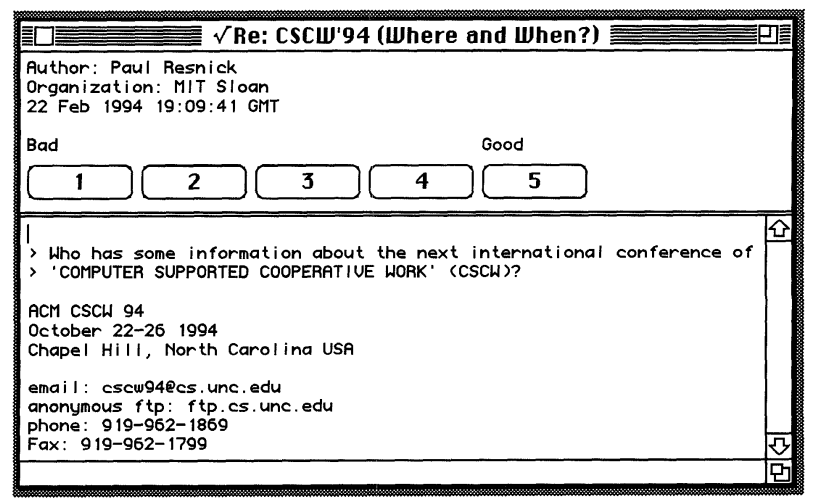


图3 使用改良后的 NewsWatcher 客户端阅读文章。用户可以用鼠标点击五个评分按钮中的一个，或者在键盘上输入 1 到 5 之间的数字。

GroupLens的架构只要求按1到5的比例报告评分，而不要求新闻客户端按此比例显示评分。为了便于学生理解评分标准，NN和Gnus客户端接受字母等级而不是数字。在向更好位局报告评级时，它们将“a”转换为5，“b”转换为4，以此类推。其他新闻客户端可以允许更多的评级等级（如1到100），并将其报告为1到5之间的分数。

**分发评分**

GroupLens完全不会干扰Usenet传播方案。恰恰相反，它在很大程度上依赖于此。更好位局（Bette rBi tBureau，BBB）按照图4中的格式，将一个或多个评级打包成一篇新闻文章，并发布到新闻服务器上。这样，GroupLens就可以利用Usenet传播计划。多年来，尽管新闻服务器的数量急剧增加，但Usenet仍能将文章传播到其他每个Usenet网站。评级服务器可以通过互联网或UUCP链接直接交换评级，但它们必须重新实现Usenet中已有的许多传播功能。

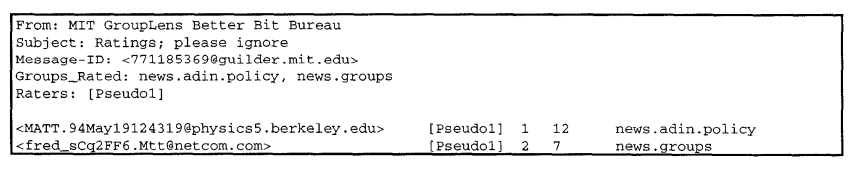


图4 一篇评分文章示例。文章主体部分的每一行都包含一个人对一篇文章的评分。每行的五个字段分别是文章编号、评分者的化名、评分、读者在评分前浏览文章所花的秒数，以及文章所在的新闻组。时间计数是可选的。行尾还可以包含其他已识别的关键字字段。

我们定义的信息格式允许在一篇文章中批量发送多个评级。每个评分都只有一行文字，而每篇Usenet netnews文章都需要几行标题。因此，将多个评级打包到一篇文章中可以节省大量开销。我们的BBB在会话级别上进行批处理（即用户在阅读会话期间输入的所有评分都会合并到一篇评分文章中）。也可以采用其他批处理策略，例如一个网站在过去一小时内的所有评分。

评分会发布在专门用于评分文章的新闻组中。一种自然的配置是为每个“普通”Usenet组建立一个并行的“评级传送”新闻组。这种方法的不足之处在于，如果一篇评级文章包含多个评级，则可能需要交叉发布到多个评级新闻组。另一个不足是，它要求新闻服务器携带大量专门用于评级的新新闻组，这可能会增加管理开销。目前，我们的BBB在一个新闻组中发布所有评级。

为了促进GroupLens的初期传播，即使用户的本地新闻服务器不携带评级新闻组，即使他们的本地网站管理员没有建立更好的位局，用户也可以参与。GroupLens的架构允许用户连接到远程BBB。图2左侧显示的是一个本地BBB，它将评级文章发布到客户端连接的同一新闻服务器上。图2右侧显示的是连接到远程BBB的客户端，该BBB通过不同的新闻服务器传播评级文章。

**预测评分**

BBB会预测读者对文章的喜爱程度。内容过滤器会根据文章中是否存在单词进行预测，而GroupLens中的BBB则会使用已经对文章进行过评分的其他人的意见。如果没有人读过某篇文章，BBB就无法对其进行预测。

当对某篇文章的评分可用时，由于评分者的观点和目标不同，他们的评分不可能是统一的。BBB将不同的评分结合起来，得出预测分数。此外，其他读者也可能对文章有不同的看法。因此，BBB可以通过改变评分的相对权重，使用相同的评分来预测不同读者的不同分数。

如图5所示，当预测与评分采用相同的尺度时，预测可以用矩阵填充法来建模，其中列是人，行是文章，而单元格则包含了人所发布的评分。矩阵中的许多单元格都是空的，因为读者尚未查看这些文章或选择不对其评分。在读者阅读相应文章之前，BBB就会预测缺失单元格的分数。

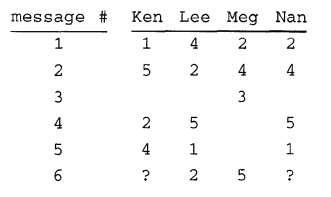
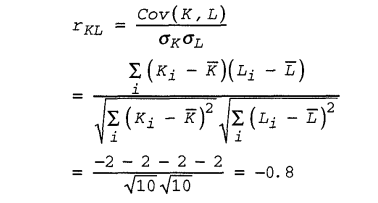


图5 一个评分矩阵示例

我们采用的所有评分方法都是基于一种启发式思维，即过去达成一致的人很可能会再次达成一致，至少是在同一新闻组的文章上。这种启发式有时会产生误导，但对大多数类型文章的偏好可能会随着时间的推移而相当稳定。

为了实现这一启发式方法，我们的BBBs首先会对以前文章的评分进行关联，以确定在对其中一篇文章进行预测时分配给其他每个人的权重。然后，他们使用这些权重来综合当前文章的评分。我们基于强化学习[12]、多元回归和最小化线性误差或平方误差的成对相关系数，研究了几种关联过去行为和使用由此产生的权重的技术。

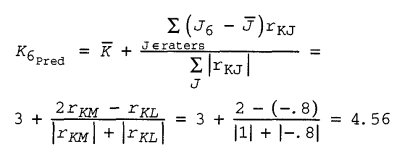
我们通过计算Ken在第6条（矩阵的最后一行）上的预测得分来说明其中一种相关性和预测技术。首先，我们计算相关系数[15]，权重介于-1和1之间，表示Ken与其他人在他们都评分的文章上的一致程度。例如，Ken与Lee的相关系数计算如下：



在上述公式中，所有总和和平均值都是根据Ken和Lee都评分的文章计算的。

同样，Ken与Meg的相关系数为+1，而与Nan的相关系数为0。也就是说，Ken倾向于不同意Lee的观点，而同意 Meg 的观点。他的评分与南的评分不相关。

要预测Ken对矩阵中最后一篇文章的评分，可根据以下公式对第6篇文章的所有评分进行加权平均：



这对Ken来说是一个合理的预测，因为这篇文章得到了过去同意他观点的人的高分和不同意他观点的人的低分。对Nan进行类似的计算，得出的预测值较低，为3.75，因为Nan过去与Lee有部分共识，Lee对文章的低评分部分抵消了Meg对文章的高评分。

评分预测系统对评分尺度的某些解释差异具有稳健性。如果两个用户完全相关，但一个用户只给3到5分，而另一个用户只给1到3分，那么第一个用户的5分将导致第二个用户的3分预测。如果两个用户完全相关，但第一个用户错误地认为1分是好分数，5分是坏分数，那么这两个用户将呈负相关，第一个用户的1分将导致第二个用户的5分预测。这就向用户清楚地解释了如何分配评分：分配你希望GroupLens为这篇文章预测的评分。

艾伦对五名受试者对新闻通讯文章的偏好进行了研究[1]，发现受试者之间的相关性非常小，因此对我们的基本假设提出了质疑，即过去同意的人很可能会再次同意。不过，如果受试者的样本更大，可能会发现一些受试者的评价有更大的重叠。更重要的是，有些人可能对某些主题有共同的兴趣，但对另一些主题却没有。两个人对技术文章的评价可能一致，但对笑话的评价可能不一致。我们的BBB为每个新闻组保留了单独的评级矩阵。

我们希望，随着BBB在计算相关性时使用更多过去的评级，预测的准确性会得到提高。明尼苏达大学有四人参加了早期版本的试点测试，使用的评分函数略有不同。虽然所有四位参与者都报告说，预测分数最终与他们的兴趣相当吻合，但他们确实观察到，在预测非常有用之前有一个启动间隔期。有必要进行进一步的实验和分析，以确定每个新用户的启动间隔可能有多长。

似乎可以开发出更好的评分机制。除了更好的矩阵填充技术外，使用他人的评分和文章内容进行预测可能也会有所帮助。此外，考虑到人们在评分前阅读文章所花费的时间也会有所帮助，我们的BBB收集了这些信息，但并未使用。

幸运的是，GroupLens的架构是开放的：任何人都可以实现另一种BBB，只要它以上述格式发布评级文章，并以与我们的BBB相同的方式与客户交流即可。我们希望，替代性BBB的开发将成为未来研究的一个活跃领域。如下所述，我们的下一个试点测试将产生评级集，我们将把这些评级集提供给其他希望评估替代评分算法的人。

**使用评分**

如何最好地使用BBB生成的评分取决于新闻客户端。有些新闻客户端可能会过滤掉分数低于阈值的文章。有的会根据分数对文章进行排序。还有些客户可能只是用数字或图形显示分数。为了与易用性设计目标保持一致，开发人员应根据客户端的整体设计修改每个新闻客户端。

新闻客户端的一个趋势是显示新闻组中未读文章的摘要。摘要的每一行都包含一篇文章的信息，通常是作者、主题行和长度。用户浏览摘要，并要求显示那些看起来有趣的文章全文。我们修改过的三个新闻客户端都使用了这种显示技术。

我们实施的三个修改版客户端在摘要显示中对分数的使用略有不同。修改后的NN客户端显示文章的顺序与普通NN客户端相同，即文章到达新闻服务器的顺序。它只是增加了一列包含预测分数的内容。在该客户端的第一个版本中，分数是以数字显示的。

修改后的Gnus客户端使用预测分数来改变摘要中文章的显示顺序。Gnus按线程对文章进行聚类。修改后的Gnus客户端根据线程中文章的最大预测得分对线程进行排序。不过，在每个主题中，文章仍按时间顺序显示，以保持讨论的流畅性。与修改后的NN一样，分数显示在摘要的附加栏中。

明尼苏达试点测试包括Gnus和NN客户端的用户。不出所料，参与者倾向于认为他们自己的新闻阅读器的排序和显示机制是最好的，但所有人都很高兴看到分数预测被纳入到标准格式中。

不过，有几位用户注意到，要直观地扫描预测结果以找出偏高的预测结果有些困难。改进版的NN客户端（图6）将预测值四舍五入到最接近的整数，并以字母等级（A-E）的形式报告，这是美国大学学生所熟悉的等级。

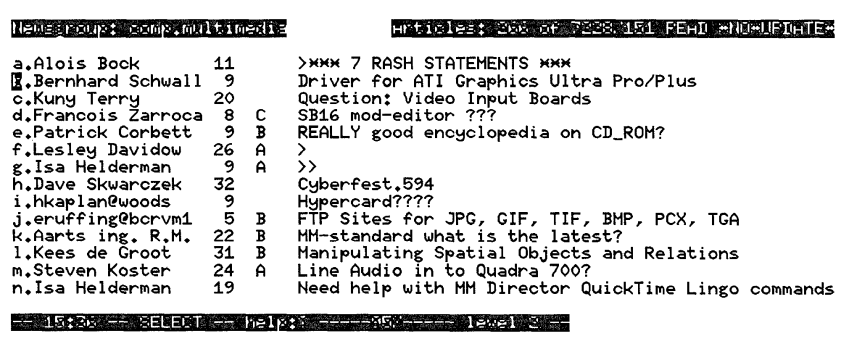


图6 改良后的 NN 客户端。第三列显示文章的行数。第四列将评分预测以字母等级显示，由 Better Bit Bureaus 给出的数字预测转换而来（5 对应 A，4 对应 B，依此类推）。若无人对文章进行评估，则不做预测。

修改后的NewsWatcher客户端以柱状图而非数字的形式显示预测分数（图7），这样更便于直观地扫描高分文章（柱状图较长）。除此之外，它还沿用了原始NewsWatcher客户端的惯例。文章被分成多个线程，摘要显示最初只显示每个线程中第一篇文章的标题行。用户可以向下旋转与线程相关的三角形，查看其余文章的标题行。

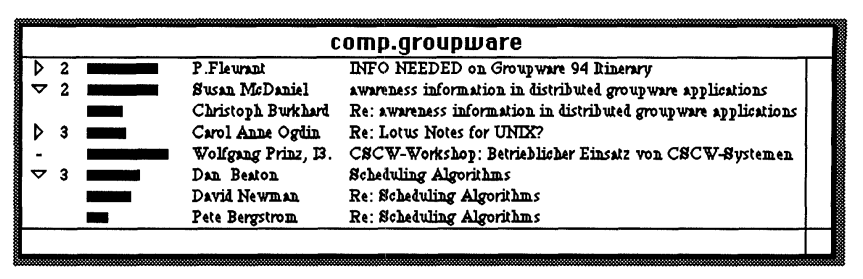


图7 改良后的NewsWatcher客户端以条形图形式显示预测分数。免责声明：这些分数是为演示目的随机生成的。在实际情况中，我们预计 Pete Bergstrom（本文作者之一）的文章预测分数会高得多。

**规模问题**

规模问题还需要进一步研究，以了解随着规模的扩大，性能会发生怎样的变化。就GroupLens而言，有几个相关的性能指标：预测质量、用户时间、BBB计算时间和磁盘存储以及网络流量。

第一个衡量标准是分数预测的质量。我们预计预测质量会随着用户数量的增加而提高，因为预测算法将获得更多数据。

另一个衡量标准是用户发布评分和接收预测所需的等待时间。在GroupLens的早期版本中，BBB的功能被整合到了新闻客户端中。独立BBB的一个主要优势是它可以预先获取评分并预先计算预测，而不是在用户启动新闻客户端时进行计算。因此，随着GroupLens的发展，即使计算评分需要花费更多的CPU时间，用户所花费的时间也应基本保持不变。

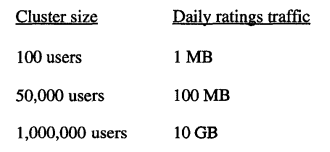
对于许多可能的预测公式，CPU时间的增长速度甚至会超过用户数量的线性增长速度。为了减少CPU时间，BBB可以只使用评级矩阵的一部分，在计算时间与预测质量之间进行权衡。

尽管每个评分都很短，但每篇新闻都可能被许多评分者阅读和评分，因此评分的总量可能超过新闻的总量。为了最大限度地减少存储需求，BBB可以采用在评级到达时使用和丢弃评级的算法，而不是存储评级。

有三种基本技术可以减少网络流量：减少评级的大小、减少评级的数量、减少每个评级的发送地点。我们的BBB在一篇文章中批量发送多个评级，这是减少每个评级存储量的第一步，但还可以进一步压缩。可以通过限制每篇文章的评分总数或具有相似资料的用户的评分数量来减少评分数量。

在GroupLens架构中，BBB与新闻客户端的分离减少了每个评分的目的地数量：每个新闻客户端只接收分数预测，而不是所有有助于这些预测的单个评分。

通过向某些BBB而不是其他BBB发送评级，可以进一步减少每个评级的目的地数量。例如，可以根据地理位置或兴趣对BBB进行分组，只在分组内交换评级。每个群组的规模必须足够小，以限制所分发的评级信息量，但又必须足够大，以提供有效的同行群组。下表估算了不同集群规模的每日网络流量，假设每个用户每天对100篇文章进行评级，每个评级大约需要100个字节。为了便于比较，目前的网络新闻流量约为每天100MB。



**GroupLens架构总结**

GroupLens的核心是用于发布评级的开放式架构。该架构规定了BBB批量生产的评级格式、Usenet传播评级的方式，以及在新闻客户端和BBB之间提供预测和评级的接口。除此之外，该架构是完全开放的。BBB和新闻客户端可以自由替换，从而为预测评级以及收集评级和展示预测的用户界面提供了实验环境。

**持续实验**

斯伦贝谢公司和明尼苏达大学之前的两次试点测试都只涉及评级的本地共享。这些测试改进了新闻客户端的整体架构和用户界面，这在前面已经讨论过。下一步，我们计划在今年夏天进行更大规模的分布式测试。我们在麻省理工学院和明尼苏达州的新闻服务器上建立了一个新闻组，并通过该新闻组建立了两个（略有不同的）更好的位局，它们通过该新闻组交流收视率。

这项测试的目的不是为了证明人们更愿意阅读使用了我们的协同过滤器的网络新闻，而不是不使用它们。我们认为，这样的评估应该至少再等一个迭代设计周期。相反，我们的目标是找出可能出现的任何意想不到的扩展问题，并收集有助于评估其他分数预测算法的数据集。

衡量任何算法有效性的主要基准是其预测从评分矩阵中删除的值的能力。乍看之下，任何大量的评分集都可能有助于创建这样一个基准。但仔细观察，完整的评分矩阵比稀疏的评分矩阵更有价值。例如，假设用户只阅读了少量文章，并根据从BBBX获得的分数预测对这些文章进行了评分。如果用户阅读了不同的文章，就会产生一个稀疏的评分矩阵。现在，假设我们希望将X与另一种可预测用户不同评分的Y进行比较。我们可以在用户阅读过的文章中比较Y和X的预测，但样本是有偏差的。也许在Y得分的情况下，用户会阅读其他文章并喜欢它们。

为了进行无偏见的比较，我们要求下一次试点测试的每位参与者都阅读训练集中的所有文章并打分。训练集将包含测试中每个新闻组的一些文章。由于用户将使用化名进行评分，因此我们可以与其他研究人员共享训练集中的评分。此外，我们还将保留训练集中文章的全文。这样，我们就可以对执行内容过滤或内容过滤与协同过滤相结合的BBB以及仅使用其他用户评分的BBB进行评估。

**社会影响**

协同过滤可能会给已经快速发展的网络新闻社区带来许多社会变化。例如，有节制的新闻组的效用可能会下降。新的社会模式必须发展起来，以鼓励对社会有益的行为，例如审查那些已经获得了一些低评分的文章。最后，如果GroupLens能够有效地创建具有共同兴趣的同伴群体，那么这些同伴群体是否具有渗透性，或者地球村是否会分裂成不同的部落？

**对网络新闻行为的改变**

GroupLens有可能改变我们现在所知的网络新闻。首先，个人用户选择阅读的文章质量应该会提高。更重要的是，随着越来越多的用户依赖GroupLens，Usenet上低质量文章的总数可能会大幅减少。因为很少有人会阅读这些文章，所以发布这些文章的动机也会降低。GroupLens还可以取代或补充其他既定的网络新闻行为。

经过审查的新闻组

GroupLens可能会减少对经过审查的新闻组的需求。GroupLens的优势在于“审查员”可以是群体，而不仅仅是个人，并且每个用户可以依赖不同的审查员，而不是为整个组设置一个单一的审查员。

某些新闻组可能会选择同时使用审查员和GroupLens。新闻组的审查员将对文章提交进行初步筛选。同行评分将允许进一步的过滤。

新闻组拆分

目前，新闻组从广泛的主题开始，随着流量的增加而拆分为更窄的主题。例如，新闻组rec.sport.football最终拆分为子组：澳大利亚、加拿大、橄榄球、职业、大学、幻想、雾，并且为NFL每个球队设立了一个组。这些拆分是一种内容过滤形式，由用户发起和管理。

GroupLens用户可能会发现许多这样的拆分并不重要，甚至在某些情况下是不希望的。随着时间的推移，用户将发现自己只阅读他们最感兴趣的新闻组的子集，因为他们与具有类似兴趣的同行组相关联。通过GroupLens而不是通过显式的内容过滤发生的兴趣拆分将自然出现，无需额外的用户或管理努力。允许通过GroupLens而不是通过显式内容过滤发生拆分，允许更多的跨领域的兴趣文章的交流。例如，由Bills球迷发布的关于即将与Cowboys进行的足球比赛的有趣文章也会到达Cowboys球迷，前提是文章发布在更专业的新闻组rec.sport.football.bills中。

杀文件

杀文件是某些新闻客户端实现的内容过滤机制。然而，许多强烈不喜欢特定主题或特定作者的用户并不使用杀文件，因为他们发现该机制复杂且繁琐。GroupLens可能成为实现相同目的的更简单的方法。用户的同行组将对这些文章给予低评分，因此只有少数用户需要阅读它们。

**激励机制**

个人通过GroupLens提供评分需要付出额外的努力，尽管这只是一个适度的投入。这些评分使其他用户能够选择有趣的文章。这是一条双向路：每个人都可以成为评分的生产者和消费者。

当某人阅读并评分一篇文章时，有提供诚实评分的动机，因为不诚实的评分将导致BBB为该用户做出糟糕的未来预测。然而，没有评分文章的动机。相反，有等待他人评分而不是自己阅读和评分文章的动机。一定的利他主义或内疚感可能导致大多数人“尽自己的一份力”，但产生的评分数量可能会少于社会最优水平。

明尼苏达四人试点测试包括一个高流量新闻组，rec.arts.movies。文章的数量如此之多，以至于每个参与者都不愿阅读每日总量的四分之一。该新闻组很快从测试中被移除。可能较大的用户群体将为高流量列表（如rec.arts.movies）生成评分，但与较大群体合作时，依赖“尽自己的一份力”的心态更加困难。

还有其他更微妙的激励问题可能出现。例如，在正面和负面评分的效果之间存在不对称性。如果前几个读者对一篇文章的评分过高，其他人将阅读该文章并给予较低的评分。另一方面，如果一篇文章的前几个评分是负面的，其他人可能会因为它最初的负面评分而不去阅读它，即使他们原本会给予它较高的评分。

为了避免这种情况，可能需要为一些人提供外部激励，以阅读和评分那些最初评分较低的文章。外部激励可以是金钱、声誉或仅仅是访问他人的评分：那些没有贡献自己应有份额的评分的人可能被拒绝访问BBB的预测。

**地球村**

现在的新闻组，就像以前的报纸和地方电视节目一样，为其读者社区提供了共同的历史。有了 GroupLens，用户可以选择只阅读与他们有许多共同兴趣的小群体的文章。随着时间的推移，这可能会导致地球村分裂成许多小部落，每个部落都形成一个虚拟社区，但又彼此孤立。

某种形式的分裂是不可避免的，甚至是可取的，因为没有一个用户能够跟上每天产生的大量新闻。问题在于，子群是封闭的还是可渗透的。预测渗透性的一个理由是，许多群组会在短时间内形成，然后解散[3]。另一种观点认为，许多用户将参与多个分组，从而为最佳创意跨越兴趣小组的界限提供了一种机制。

**结论**

共享评价在各种活动中都很有用。我们向朋友、同事和专业评论员询问他们对书籍、电影、期刊文章、汽车、学校和社区的意见。显然，某种形式的共享评价也应该有助于筛选电子信息流，如网络新闻。目前尚不清楚这些评价应该采取何种形式，如何收集和传播，以及如何用于选择文章进行阅读。

GroupLens是一种有前景的方法。一个单一的数字给出了对特定读者相关的文章所有维度的综合评分。我们修改了三个新闻阅读客户端，以方便输入这样的数字评分。我们还修改了客户端显示主题行的方式，以包含基于他人评分的预测分数。

当然，由于兴趣或质量评估的不同，读者对特定文章的看法会有差异。为了适应意见分歧，不是所有读者都会对特定评价者给予同等信任。我们实现的算法会根据读者和评价者过去观点的相关性自动确定对每个评价的重视程度。这有两个好处：读者不必事先知道应该信任哪些评价者，即使评价者选择保持匿名，他们的评价也可以变得值得信赖。

GroupLens架构允许新用户连接和新的评分服务器上线，无需全球协调。新用户只需使用修改后的新闻客户端，并连接到评分服务器。用户无需说服她的网络新闻服务器管理员修改新闻服务器、运行任何额外软件，甚至无需要求承载任何额外的新闻组。新的评分服务器只需访问承载评分新闻组的新闻服务器。

此外，架构是开放的。任何希望修改新闻客户端以允许输入评价或使用预测分数的人都可以这样做，只要客户端遵循我们为与评分服务器通信而建立的协议即可。任何希望改进我们的评分服务器评分预测的人都可以这样做。可能有比我们更好的方法来关联过去的评价。也可能有将评价与内容过滤结合使用的方法。例如，在关联过去的评价时，评分算法可能只考虑与当前文章在某种程度上相似的过去文章的评价。我们的下一次试点测试应生成可用于评估替代预测方法的数据集。

只有进一步的测试才能揭示GroupLens是否收集了合适的评价，并以人们喜欢的方式使用它们。如果简单的数字评价被证明是足够的，那么架构将扩展到大量的评分服务器和用户。如果不是，那么我们的测试数据将有助于开发和评估其他共享和使用评价的机制。

目前，人们阅读新闻文章并对其做出反应，但这些反应被浪费了。GroupLens是挖掘这一隐藏资源的第一步。

**致谢：**Shumpei Kumon在CSCW 92上的主旨演讲激发了我们对声誉在社交过滤中的实际应用的研究。感谢Lorin Hitt和Carl Feynman就如何基于过去的相关性进行评分预测进行了有益的讨论。Peter Foltz和Sue Dumais慷慨地提供了他们关于内容过滤实验生成的一个测试评分集。感谢Chris Avery、Joe Adler、Yannis Bakos、Erik Brynjolfsson、David Goldberg、Bill MacGregor、Tom Malone、David Maltz、Vahid Mashayekhi、Lisa Spears、Doug Terry、Mark Uhrmacher和Zbigniew Wieckowski。

# **参考文献**

1. Allen, R.B. User Models: Theory, Method, and Practice. International Journal of Man-Machine Studies, 32, (1990), pp. 511-543.
2. Belkin, N.J. and Croft, B.W. Information Filtering and Information Retrieval: Two Sides of the Same Coin? CACM, 35, 12 (1992), pp. 29-38.
3. Brothers, L., Hollan, J., Nielsen, J., Stornetta, S., Abney, S., Furnas, G. and Littman, M. Supporting Informal Communication via Ephemeral Interest Groups. In Proceedings of CSCW 92 (1992, New York: ACM), pp. 84-90.
4. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K. and Harshman, R. Indexing by Latent Semantic Analysis. Journal of the American Society for Information Science, 41, 6 (1990), pp. 391-407.
5. Foltz, P.W. and Dumais, S.T. Personalized Information Delivery: An Analysis of Information Filtering Methods. Communications of the ACM, 35, 12 (1992), pp. 51-60.
6. Goldberg, D., Nichols, D., Oki, B.M. and Terry, D. Using Collaborative Filtering to Weave an Information Tapestry. Communications of the ACM, 35, 12 (1992), pp. 61-70.
7. Hill, W.C., Hollan, J.D., Wroblewski, D. and McCandless, T. Edit Wear and Read Wear. In Proceedings of CHI 92 Conference on Human Factors in Computing Systems (1992, New York: ACM), pp. 3-9.
8. Kahn, R.E. and Cerf, V.G. The Digital Library Project, Volume 1: The Wold of Knowbots. An Open Architecture for a Digital Library System and a Plan for Its Development. CNRI, 1985 Preston White Drive, Suite 100, Reston, VA 22091 Tech Report (March, 1988).
9. Karlqren, J. Newsgroup Clustering Based on User Behavior-- A Recommendation Algebra. Swedish Institute of Computer Science #SICS-T--94/04-SE (March, 1994).
10. Kawell, L.J., Beckhardt, S., Halvorsen, T. and Ozzie, R. Replicated Document Management in a Group Communication System. In Proceedings of CSCW 88 (1988, New York: ACM).
11. Kumon, S. From Wealth to Wisdom: A Change in the Social Paradigm. In Proceedings of CSCW 92 (1992, New York: ACM), pp. 3.
12. Maes, P. and Kozierok, R. Learning Interface Agents. In Proceedings of AAAI 93 (1993, San Mateo, CA: American Association for Artificial Intelligence).
13. Malone, T.W., Grant, K.R., Turbak, F.A., Brobst, S.A. and Cohen, M.D. Intelligent Information Sharing Systems. Communications of the ACM, 30, 5 (1987), pp. 390-402.
14. Maltz, D.A. Distributing Information for Collaborative Filtering on Usenet Net News. MIT Department of EECS MS Thesis (May, 1994).
15. Pindyck, R.S. and Rubinfeld, D.L. Econometric Models and Economic Forecasts. Macgraw-Hill, New York, 1991.
16. Salton, G. and Buckley, C. Term-Weighting Approaches in Automatic Text Retrieval. Information Processing and Management, 24, 5 (1988), pp. 513-523.
17. Salton, G. and Buckley, C. Improving Retrieval Performance by Relevance Feedback. Journal of the American Society for Information Science, 41, 4 (1990), pp. 288-297.
18. Sheth, B. A Learning Approach to Personalized Information Filtering. MIT Department of EECS MS Thesis (February, 1994).
19. Stodolsky, D.S. Invitational Journals Based Upon Peer Consensus. Roskilde University Centre, Institute of Geography, Socioeconomic Analysis, and Computer Science. ISSN 0109-9779-29 #No. 29/ 1990 (, 1990).
20. Suchak, M.A. GoodNews: A Collaborative Filter for Network News. MIT Department of EECS MS Thesis (February, 1994).
21. Wiederhold, G. Mediators in the Architecture of Future Information Systems. IEEE Computer, March, (1992), pp. 38-49.