



[12] 发明专利申请公开说明书

[21] 申请号 02156145.1

[43] 公开日 2003 年 7 月 9 日

[11] 公开号 CN 1428689A

[22] 申请日 2002.12.11 [21] 申请号 02156145.1

[30] 优先权

[32] 2001.12.11 [33] US [31] 10/020712

[71] 申请人 奥弗图尔服务公司

地址 美国加利福尼亚州

[72] 发明人 M·佩尼 W·戴维斯 D·格迪斯

J·杜克斯-施洛斯伯格

D·戴维斯

[74] 专利代理机构 中国专利代理(香港)有限公司

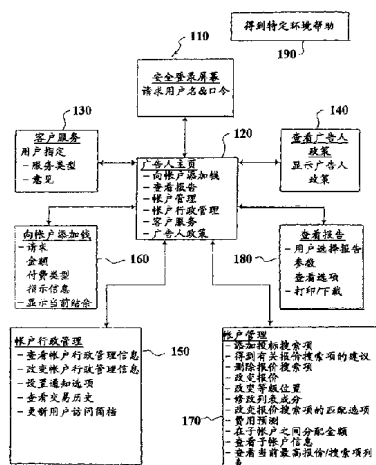
代理人 杨凯 罗朋

权利要求书 9 页 说明书 39 页 附图 20 页

[54] 发明名称 利用协作过滤和万维网蜘蛛式搜索推荐搜索项

[57] 摘要

在根据位置付费的搜索系统中，系统以两种方式之一或同时以两种方式向广告人提供管理其帐户的推荐。 第一种技术直接在广告人网站上寻找好的搜索项。 第二种技术涉及将广告人与其它相似的广告人进行比较并推荐其它广告人已选择的搜索项。 第一种技术被称为蜘蛛式搜索，而第二种技术被称为协作过滤。 在最佳实施例中，蜘蛛式搜索步骤的输出被用作到协作过滤步骤的输入。 然后以自然方式将两个步骤最后输出的搜索项进行交织。



1. 一种用于数据库搜索系统的方法，所述方法包括：
接收与所述数据库搜索系统上的广告人有关的搜索项列表；
5 根据所述数据库搜索系统上其它广告人的搜索项确定候选搜索项；以及
从所述候选搜索项中推荐更多的搜索项。
2. 如权利要求 1 所述的方法，其特征在于所述确定候选搜索项包括：
10 为搜索项分配等级；以及
利用广告人搜索项的所述分配等级，计算所述广告人和一个或多个所述其它广告人之间的相关性。
3. 如权利要求 2 所述的方法，其特征在于还包括：
预测候选搜索项将与所述广告人相关的可能性。
- 15 4. 如权利要求 3 所述的方法，其特征在于所述预测包括：
确定所述候选搜索项的质量量度，以及
根据所述质量量度预测候选搜索项的相关性。
5. 一种用于数据库搜索系统的方法，所述方法包括：
维护包括相关搜索项的搜索列表数据库；
20 接收与广告人有关的搜索项列表；
计算搜索项的等级；以及
根据所述算得的等级向所述广告人推荐更多搜索项。
6. 如权利要求 5 所述的方法，其特征在于所述计算等级包括：
为搜索项分配等级。
- 25 7. 如权利要求 5 所述的方法，其特征在于所述计算等级包括：
预测搜索项的等级。
8. 如权利要求 5 所述的方法，其特征在于所述接收搜索项列表包括：

从所述广告人接收初始搜索项列表。

9. 如权利要求 5 所述的方法，其特征在于所述接收搜索项列表包括：

5 识别所述数据库搜索系统上的现有广告人；以及
根据所述现有广告人的搜索项形成所述搜索项列表。

10. 如权利要求 5 所述的方法，其特征在于所述接收搜索项列表包括：

10 接收网站统一资源定位器（URL）；以及
对与所述 URL 有关的所述网站进行蜘蛛式搜索，以析取所述搜索项列表的搜索项。

11. 如权利要求 10 所述的方法，其特征在于所述对所述网站进行蜘蛛式搜索包括：

15 从所述网站的页面得到数据；
从所述数据记录候选搜索项；以及
确定各个候选搜索项的质量量度。

12. 如权利要求 11 所述的方法，其特征还在于还包括：
根据所述质量量度对所述候选搜索项进行排序；以及
仅推荐具有超出阈值的质量量度的候选搜索项。

20 13. 如权利要求 10 所述的方法，其特征还在于所述蜘蛛式搜索包括：

从所述网站的一个或多个页面得到数据；以及
检查所述一个或多个页面上的文本以得到候选搜索项。

25 14. 如权利要求 13 所述的方法，其特征还在于检查文本包括：
基本上检查所述一个或多个页面上的所有文本；以及
检查所述一个或多个页面上的元标记。

15. 如权利要求 13 所述的方法，其特征还在于接收网站 URL 包括：
接收所述广告人的 URL 作为所述网站 URL。

16. 如权利要求 13 所述的方法，其特征还在于接收网站 URL 包括：

从所述广告人接收所述网站 URL。

17. 如权利要求 5 所述的方法，其特征在于还包括：

从用户接收搜索查询；

在搜索列表数据库中搜索匹配搜索项；

5 通过根据与所述搜索列表有关的广告人报价金额对匹配搜索项进行格式编排来准备搜索结果；以及
将所述搜索结果传递给所述用户。

18. 一种数据库搜索系统，它包括：

搜索项数据库，每个搜索项与一个或多个广告人有关；以及

10 程序代码，它被配置为根据所述数据库中的搜索项为广告人推荐更多搜索项。

19. 如权利要求 18 所述的数据库搜索系统，其特征在于所述程序代码包括：

15 协作过滤代码，它被配置为根据与所述数据库搜索系统的其它广告人有关的搜索项推荐所述更多搜索项。

20. 如权利要求 19 所述的数据库搜索系统，其特征在于所述协作过滤代码包括皮尔逊相关码。

21. 如权利要求 19 所述的数据库搜索系统，其特征在于所述程序代码包括：

20 程序循环，用于迭代演算所述协作过滤代码。

22. 如权利要求 18 所述的数据库搜索系统，其特征在于还包括：

代码，用于在重复所述程序循环之前从所述广告人接收接受和拒绝表示。

25 23. 如权利要求 18 所述的数据库搜索系统，其特征在于所述程序代码包括：

蜘蛛代码，用于推荐所述更多搜索项。

24. 如权利要求 18 所述的数据库搜索系统，其特征在于所述程序代码包括：

蜘蛛代码, 在网站中寻找初始接受的搜索项; 以及
协作过滤代码, 提供所述推荐的更多搜索项。

25. 如权利要求 24 所述的数据库搜索系统, 其特征在於所述蜘蛛代码被配置为对所述广告人的网站进行蜘蛛式搜索。

5 26. 如权利要求 23 所述的数据库搜索系统, 其特征在於所述蜘蛛代码被配置为对所述广告人指定的网站进行蜘蛛式搜索。

27. 如权利要求 18 所述的数据库搜索系统, 其特征在於还包括:
过滤代码, 根据质量量度对候选搜索项进行过滤以生成所述推荐的更多搜索项。

10 28. 如权利要求 15 所述的数据库搜索系统, 其特征在於还包括:
搜索引擎程序代码, 它被配置为搜索所述数据库, 以响应于来自用户的搜索查询。

29. 一种用于储存广告人搜索列表的数据库搜索系统的数据库操作方法, 所述方法包括:

15 对指定网站进行蜘蛛式搜索, 以得到广告人的广告人搜索项初始列表;

利用其它广告人的搜索项过滤所述广告人搜索项初始列表; 以及

20 将所述广告人的搜索列表储存在搜索列表数据库中, 所述搜索列表由所述过滤了的搜索项形成。

30. 如权利要求 29 所述的数据库操作方法, 其特征在於所述指定网站包括广告人指定的网站。

31. 如权利要求 29 所述的数据库操作方法, 其特征在於所述指定网站包括所述数据库搜索系统指定的网站。

25 32. 如权利要求 29 所述的数据库操作方法, 其特征在於还包括:
为搜索项分配等级; 以及

利用广告人搜索项的所述分配等级, 计算所述广告人和一个或多个所述其它广告人之间的相关性。

33. 如权利要求 32 所述的数据库操作方法, 其特征在于还包括:
预测搜索项将与所述广告人相关的可能性。

34. 如权利要求 33 所述的数据库操作方法, 其特征在于所述预测
测包括:

5 为所述候选搜索项确定质量量度, 以及
根据所述质量量度预测候选搜索项的相关性。

35. 如权利要求 29 所述的数据库操作方法, 其特征在于蜘蛛式
搜索所述指定网站包括:

从所述指定网站的页面得到数据;

10 从所述数据记录候选搜索项; 以及
确定各个候选搜索项的质量量度。

36. 如权利要求 35 所述的数据库操作方法, 其特征在于还包括:
根据所述质量量度对所述候选搜索项进行排序; 以及
仅推荐具有超出阈值的质量量度的候选搜索项。

15 37. 如权利要求 29 所述的数据库操作方法, 其特征在于所述过
滤包括:

确定所述数据库搜索系统上所述广告人的网站和其它广告人的
网站之间的相关性;

20 利用所述相关性, 确定所述初始广告人搜索项列表中各个广告
人搜索项的等级; 以及

根据所述等级组织所述初始广告人搜索项列表的搜索项。

38. 如权利要求 29 所述的数据库操作方法, 其特征在于还包括:
向所述广告人显示所述组织的搜索项;
接收广告人对所述组织搜索项的接受表示;

25 根据所述接受表示调整所述广告人搜索项列表; 以及
过滤所述调整列表。

39. 如权利要求 38 所述的数据库操作方法, 其特征在于还包括:
迭代地显示所述组织的搜索项; 接收修订的广告人接受表示;

以及调整所述广告人搜索项列表。

40. 如权利要求 38 所述的数据库操作方法，其特征在于还包括：
从用户接收搜索查询；

在所述搜索列表数据库中搜索匹配搜索项；

5 通过根据与所述匹配搜索列表有关的广告人报价金额对匹配搜索项进行格式编排来准备搜索结果；以及
将所述搜索结果传递给所述用户。

41. 一种用于数据库搜索系统的方法，所述方法包括：

接收广告人的搜索项；

10 响应于所述收到的搜索项，生成更多相关搜索项的列表；以及
从所述更多相关搜索项的列表得到广告人所选的搜索项。

42. 如权利要求 41 所述的方法，其特征在于所述生成所述更多搜索项的列表包括：

为搜索项分配等级；以及

15 将来自所述收到的搜索项的一个或多个文本串与搜索项数据库相匹配。

43. 如权利要求 38 所述的方法，其特征在于所述生成所述更多搜索项的列表包括：

20 将来自所述收到的搜索项的一个或多个文本串与同义辞典相匹配。

44. 如权利要求 41 所述的方法，其特征在于还包括：

显示用于输入所选搜索项的一个或多个搜索列表的表格。

45. 如权利要求 42 所述的方法，其特征在于还包括：

25 输入所述所选搜索项作为各个所述一个或多个搜索列表中的缺省值。

46. 一种用于推荐搜索项的方法，所述方法包括：

接收描述广告人的信息；以及

根据所述广告人与其它广告人的相似性推荐搜索项。

47. 如权利要求 46 所述的方法，其特征在于还包括：
根据描述所述广告人的所述信息和描述所述其它广告人的信息
分配等级；以及
至少部分根据所述等级推荐搜索项。
- 5 48. 如权利要求 46 所述的方法，其特征在于还包括：
根据描述所述广告人的所述信息和描述所述其它广告人的信息
预测所述广告人和所述其它广告人的等级；以及
至少部分根据所述等级推荐搜索项。
- 10 49. 如权利要求 46 所述的方法，其特征在于还包括：
根据描述所述广告人的所述信息和描述所述其它广告人的信息
计算所述广告人和所述其它广告人的相关性；以及
至少部分根据所述相关性推荐搜索项。
- 15 50. 如权利要求 49 所述的方法，其特征在于所述计算相关性包
括：
计算皮尔逊相关性。
51. 一种用于推荐搜索项的方法，所述方法包括：
下载根源于特定统一资源定位器（URL）的网页；以及
推荐所述网页的正文中出现的搜索项。
- 20 52. 如权利要求 51 所述的方法，其特征在于还包括：
识别具有网站的广告人；
将所述广告人的网站指定为所述指定 URL。
53. 如权利要求 51 所述的方法，其特征在于还包括：
从广告人接收网站标识符作为所述指定 URL。
- 25 54. 一种用于推荐搜索项的方法，所述方法包括：
接收统一资源定位器（URL）；
下载根源于所述 URL 的网页；以及
推荐所述网页中出现的搜索项。
55. 如权利要求 54 所述的方法，其特征在于还包括：

识别具有网站的广告人;

接收所述广告人的网站作为所述 URL。

56. 一种用于推荐搜索项的方法, 所述方法包括:

5 计算候选搜索项的质量量度, 各个候选搜索项的所述质量量度是所述各个搜索项的万维网频率的函数和搜索项的搜索频率的函数; 以及

推荐那些所述算得的质量量度超过质量阈值的搜索项。

57. 如权利要求 56 所述的方法, 其特征在于还包括:

根据第二量度对所述候选搜索项排序。

10 58. 如权利要求 56 所述的方法, 其特征在于还包括:

自动校正所述质量阈值。

59. 一种用于数据库搜索系统的推荐搜索项的方法, 所述方法包括:

维护与广告人列表有关的搜索项数据库;

15 接收描述广告人的信息; 以及

向所述广告人推荐搜索项。

60. 如权利要求 59 所述的方法, 其特征在于还包括:

通过协作过滤确定候选搜索项; 以及

根据所述候选搜索项推荐搜索项。

20 61. 如权利要求 59 所述的方法, 其特征在于还包括:

下载根源于统一资源定位器的网页;

根据所述网页的正文中出现的项推荐所述搜索项。

62. 如权利要求 59 所述的方法, 其特征在于还包括:

接收统一资源定位器 (URL);

25 下载根源于所述 URL 的网页; 以及

根据所述网页中出现的项推荐搜索项。

63. 如权利要求 59 所述的方法, 其特征在于还包括:

计算候选搜索项的质量量度, 各个候选搜索项的所述质量量度

是所述各个搜索项的万维网出现频率的函数和搜索项的搜索频率的函数；以及

根据那些所述算得的质量量度超过质量阈值的搜索项推荐搜索项。

- 5 64. 如权利要求 59 所述的方法，其特征在于还包括：
从所述广告人接收对所述推荐搜索项的反馈；以及
根据所述反馈改变所述推荐搜索项。

利用协作过滤和万维网蜘蛛式搜索推荐搜索项

5 相关申请

本申请是 2001 年 7 月 24 日以 Davis 等人的名义提交的题为“用于影响计算机网络搜索引擎生成的搜索结果列表上的位置的系统和方法”序列号 09/911674 的申请的部分继续申请，该申请与本申请共同转让并通过引用将该申请的全部结合到本文中，并且该申请是 1999
10 年 5 月 28 日以 Darren J.Davis 等人的名义提交的申请序列号 09/322677 的继续申请，现在为美国专利号 6269361。

技术领域

本发明涉及数据库搜索系统，更具体地说，涉及一种利用协作
15 过滤和万维网蜘蛛式搜索推荐搜索项的数据库搜索系统和方法。

背景技术

美国专利 No.6269361 公开了具有广告人帐户的数据库。各帐户包括广告人的联系和计费信息。另外，各帐户包括至少一个具有至少三个部分的搜索列表：描述、包括一个或多个关键字的搜索项以及
20 报价。广告人可以在通过验证过程登录到他或她的帐户之后添加、删除或修改搜索列表。广告人通过首先选择要列出的与网站或其它信息源有关的搜索项来影响广告人帐户中搜索列表的位置。广告人将搜索项和描述输入到搜索列表中。广告人通过连续在线竞争投标过程影响搜索列表的位置。投标过程在广告人输入一个新报价时开始，
25 输入的报价最好是用于搜索列表的钱的金额。然后，所公开的系统将该报价与其它所有用于同一搜索项的报价进行比较，并为具有该搜索项的所有搜索列表产生等级值。投标处理产生的等级值确

定广告人的列表将出现在产生的搜索结果列表的什么位置，作为计算机网络上的搜索人或用户查询搜索项的响应。广告人更高的报价将导致更高的等级值和更有利的位置。这种系统被成为根据位置付费（pay-for-placement）的搜索引擎。

5 因此，在传统上，当用户在根据位置付费的搜索引擎上执行搜索时，根据各个广告人对用户搜索项的投标多少对结果进行排序。因为不同的用户将利用不同的词来寻找相同信息，因此重要的是广告人对大量不同的搜索项进行投标，以便使到其站点的流量最大。广告人的搜索项列表越好越多，广告人将看到越大的流量。

10 作为例示，海产食品卖主将希望不仅对词“seafood”投标，他还希望对诸如“fish（鱼）”、“tuna（金枪）”、“halibut（比目）”和“fresh fish（鲜鱼）”等项进行投标。慎重考虑后产生的列表将常常包括数百项。好的搜索项具有三个重要特性：它们适合于广告人的站点；它们足够流行，许多用户很可能用它们进行搜索，以及它们在广告人为在搜索结果中得到高等级而必须的投标量方面提供较好的值。愿意花时间考虑所有这些因素的广告人将得到好结果。

 很遗憾，极少的广告人理解如何建立好的搜索项列表，并且现在只有有限的工具能帮助它们。典型的现今技术是位于 <http://inventory.overture.com> 的由 Overture Services, Inc.提供的搜索项建议工具（STST），STST 根据字符串匹配提供建议。给定一个词，STST 返回包括该词的所有搜索项的排序列表。该列表按照用户在上个月搜索这些项的次数来排序。在海鲜的例示中，如果广告人输入词“fish”，他的结果将包括如“fresh fish”、“fish market（鱼市场）”、“tropical fish（热带鱼）”以及“fish bait（鱼饵）”的项，但没有如“tuna”或“halibut”的词，因为它们不包括字符串“fish”。为了建立他的初始搜索项列表，新的广告人将经常输入若干词到 STST 中，然后对它返回的所有项投标。

 这种方法存在三个问题。首先，尽管 STST 找到很多好的项，比

如“fresh fish”和“fish market”，但它也找到许多不好的项，如“fishing（钓鱼）”、“tropical fish”和“fish bait”，这些项和广告人的站点没有联系。这些项为搜索引擎提供商带来了额外的工作，因为它的编辑人员必须过滤掉广告人提供的不合适的项。其次，STST 漏掉了许多好的项，如“tuna”和“halibut”。这导致了广告人的流量损失以及提供商的收入损失，因为每个投标应有助于提高搜索项的价格并增加提供商的收入。第三，广告人容易简单地忽略掉他应该输入到 STST 的某个词，由此失去适合其站点的搜索项的整个空间。这些遗漏项也将导致广告人的流量损失以及提供商的收入损失。

STST 的改进型式是可以在 <http://users.idealab.com/~charlie/advertisers/start.html> 中找到的 GoTo Super Term Finder(到超级项目查找器，STF)。此工具跟踪两个列表：用于广告人站点的好词的接受列表，以及不好的词或与广告人站点或其内容无关的词词的拒绝列表。STF 显示包括在第一列表中而不在第二列表中的词的所有搜索项的排序列表。如同 STST 一样，结果列表按用户在上个月搜索这些项的次数来排序。在海鲜的例示中，如果接受列表包括词“fish”而拒绝列表包括词“bait”，则输出将显示如“fresh fish”和“tropical fish”的项，但不会出现“fish bait”。广告人可在迭代过程中利用这个输出精炼其接受和拒绝列表。

尽管 STF 较 STST 进行了改进，但它仍存在类似问题。在海鲜例示中，许多搜索项包括词“fish”，但与海鲜站点无关。广告人仍必须亲自对这些项进行识别并一个一个否定。除非否定项共享公共的词，否则广告人在 STF 中必须要做的工作量和在 STST 中一样多。两个工具都存在的弱点在于不能够识别如“tuna”或“halibut”这样的好搜索项。可以有許多这样的语义相关项；它们可能甚至同时出现在广告人的网站上。但想起每一个仍是广告人的责任。STST 和 STF 的问题在于它们两个都是根据语法属性查找搜索项，并且它们迫使广告人自己考虑根词。显然需要更好的方法，一种考虑单词的意义

并且能够通过查看广告人的网站来自动识别这些词的方法。

发现语义相关项的系统是 Wordtracker(词跟踪系统), 该系统可以在 <http://www.wordtracker.com> 中找到。给出某个搜索项, Wordtracker 以两种方法推荐新的项。首先, Wordtracker 通过在同义辞典中查找词来推荐一些词。其次, Wordtracker 通过利用称为“横向搜索”的算法搜寻词来推荐一些词。横向搜索通过两个流行万维网搜索引擎运行原始搜索项。然后, 它下载前 200 个网页的结果, 从页面的元标记“关键词”和“描述”中析取所有项并返回按各个项出现在这些标记中的频繁程度排序的列表。

Wordtracker 较之 STST 和 STF 仅有少量改进。在海鲜例示中, 如果广告人搜索“fish”, 则他很可能看到包括“tuna”和“halibut”的结果, 但他仍将看到不好的项, 如与其站点无关的“tropical fish”和“fish bait”。更具体的搜索“seafood(海鲜)”将去掉这些不好的项中的一些, 但引进了其它来自海鲜餐厅的如“restaurant(餐厅)”和“steak(鱼排)”的项。与 STF 不同的是, Wordtracker 没有办法滤去这种不好的项并精炼搜索。也没法提供好项的广泛列表, 这是因为万维网搜索引擎在搜索项多于一个时效果不好。这两个局限是严重的, 因为广告人能够识别确切描述其站点并且其他人喜欢的单一搜索项的情况非常罕见。Wordtracker 还存在这样的问题: 元关键词并非总是表示网站。因为没有编辑检查, 因此网站的设计者常常包括谬误的关键词, 企图使他们的网页在搜索引擎上更突出。搜索引擎本身也是有局限性的, 并且可以在它们 200 个列表中返回许多与广告人站点无关的网页。最后, 如 STST 和 STF 一样, Wordtracker 仍需要广告人考虑他自己的搜索项作为开始。

已知这些缺陷, 显然需要更好的工具, 一种可能找到用于广告人站点的所有好搜索项而免除不好的项的方法。

发明内容

5 仅以说明为目的，本发明的实施例以两种方式中的一种或两种提供搜索项推荐。第一种技术涉及直接在广告人网站上查找好的搜索项。第二种技术涉及将广告人和其它相似的广告人进行比较并推荐其它广告人已经选择了的搜索项。第一种技术被称为“蜘蛛式搜索”而第二种技术被称为“协作过滤”。在最佳实施例中，蜘蛛式查找步骤的输出被用作协作过滤步骤的输入。然后以自然方式对两个步骤最后输出的搜索项进行交织。

上述最佳实施例作为说明提供。此部分中不应有内容被看作是对权利要求的限制，由权利要求书定义本发明的范围。

10

附图说明

图 1 是说明大网络和根据本发明用于产生根据结果付费（pay-for-performance）的搜索结果的系统和方法的实施例之间的关系的方框图；

15

图 2 是用于本发明的一个实施例的菜单、显示屏幕和输入屏幕图；

图 3 是说明在本发明的一个实施例中执行的广告人用户登录过程的流程图；

20

图 4 是说明在本发明的一个实施例中执行的管理用户登录过程的流程图；

图 5 是在本发明的一个实施例中使用的帐户记录数据的图；

图 6 是向本发明的一个实施例中使用的帐户记录中添加钱的方法的流程图；

图 7 是由本发明的一个实施例产生的搜索结果列表的例示；

25

图 8 是说明本发明的一个实施例中使用的改变报价过程的流程图；

图 9 是图 8 的改变报价过程中所有的显示屏幕的例示；

图 10 是说明用于在根据位置付费的搜索引擎上向广告人推荐搜

索项的方法的流程图;

图 11 是说明通过对网站进行蜘蛛搜索来对搜索项评级的方法的流程图;

图 12-14 是说明通过协作过滤来对搜索项评级的方法的流程图;

5 图 15-17 是说明计算两个广告人之间的皮尔逊相关性的流程图;
以及

图 18-20 是说明对来自蜘蛛搜索和协作过滤的预测进行组合的流程图。

10 具体实施方式

公开了在基于客户机/服务器的计算机网络系统上由诸如广告人的站点发起人确定用于产生根据结果付费的搜索结果的方法和系统。提供以下描述使本领域的所有技术人员能够实现和使用本发明。出于解释的目的, 陈述了具体术语以提供对本发明的透彻理解。具体应用的描述只作为例示而提供。对最佳实施例的各种修改对本领域的技术人员将是显而易见的, 并且本文所定义的一般原理可适用于其它实施例和应用, 并不背离本发明的精神和范围。因此, 本发明无意限于所示实施例, 而应依照与本文所公开的原理和特点一致的最广范围。

20 现在参考附图, 图 1 是用于本发明的最佳实施例的配置为客户机/服务器体系结构的分布系统 10 的例示。“客户机”是一类或一组中使用与其无关的另一类或一组的服务的一个成员。在诸如因特网的计算机网络环境中, 客户机是请求由称为服务器程序的另一进程提供的服务的进程(一般说来是程序或任务)。客户进程利用请求的
25 服务, 并不需要知道有关另一服务器程序或服务器本身的任何工作细节。在连网系统中, 客户进程一般在访问由运行相应服务器进程的另一计算机提供的共享网络资源的计算机上运行。但是, 还应该注意, 客户机程序和服务器程序在同一计算机上运行也是可能的。

“服务器”一般是可通过诸如因特网的通信媒体访问的远程计算机系统。客户进程可以在第二个计算机系统中有效，并通过允许多台客户机利用服务器的信息收集能力的通信媒体与服务器进程通信。因此，服务器基本上作为计算机网络的信息提供者。

5 因此，图 1 的方框图显示了一种分布系统 10，它包括多个客户计算机 12、多个广告人万维网服务器 14、帐户管理服务器 22 以及搜索引擎万维网服务器 24，所有这些都连接到网络 20。网络 20 在下文中将一般地称作因特网。尽管本发明的系统和方法对因特网特别有用，但应该理解，客户计算机 12、广告人万维网服务器 14、帐户管理服务器 22 以及搜索引擎万维网服务器 24 可以通过多个不同类型的网络之一连接在一起。这样的网络可包括局域网（LAN）、其它广域网（WAN）以及通过电话线访问的区域网络，比如商业信息服务。客户和服务进程甚至可包括同时在一个计算机上执行的不同程序。

15 客户计算机 12 可以是常规的个人计算机（PC）、工作站或任何其它规格的计算机系统。客户计算机 12 一般包括一个或多个处理器、存储器、输入/输出装置以及诸如常规的调制解调器的网络接口。广告人万维网服务器 14、帐户管理服务器 22 和搜索引擎万维网服务器 24 可以相似地配置。但是，广告人万维网服务器 14、帐户管理服务器 22 和搜索引擎万维网服务器 24 可分别包括通过单独的专用网络连接到的多个计算机。事实上，网络 20 可包括成百上千的计算机单独网络。

25 客户计算机 12 可执行诸如 NAVIGATOR、EXPLORER 或 MOSAIC 浏览器程序的万维网浏览器程序 16，以定位储存在广告人服务器 14 上的网页或记录 30。浏览器程序 16 允许用于输入要检索的特定网页 30 的地址。这些地址被称为统一资源定位器，或 URL。另外，一旦检索到网页，浏览器程序 16 可以在用户“点击”到其它网页的超级链接时提供到其它页面或记录的访问。这类超级链接位

于网页 30 中并为用户提供自动方式来输入另一页的 URL 并检索该页。页面可以是包括内容纯文字信息或者诸如软件程序、图形、音频信号、视频等更复杂的数字编码多媒体内容的数据库记录。

在本发明的最佳实施例中，如图 1 所示，客户计算机 12 利用超文本传输协议 (HTTP) 提供的功能通过网络 20 与各种网络信息提供者进行通信，其中包括帐户管理服务器 22、搜索引擎服务器 24 和广告人服务器 14，但是也可以使用诸如 FTP、SNMP、TELNET 和若干本领域已知的其它协议的其它通信协议。优选的是将搜索引擎服务器 24、帐户管理服务器 22 和广告人服务器 14 定位在万维网上。

如上所述，本发明的最佳实施例中考虑了至少两种类型的服务器。考虑的第一种服务器是包括计算机存储媒体 32 和处理系统 34 的帐户管理服务器 22。数据库 38 存储在帐户管理服务器 22 的存储媒体 32 上。数据库 38 包括广告人帐户信息。从以下描述将了解，本发明的系统和方法可以用软件实现，其中软件作为可执行指令存储在帐户管理服务器 22 的诸如存储器或大容量存储装置的计算机存储媒体上。在客户计算机 12 上运行的常规浏览器程序 16 可被用于访问储存在帐户管理服务器 22 上的广告人帐户信息。到帐户管理服务器 22 的访问最好通过未示出的防火墙完成，所述防火墙保护帐户管理和搜索结果位置程序以及帐户信息不被外部篡改。可以通过对标准通信协议的增强提供更大的安全性，比如安全 HTTP 或安全套接层。

考虑的第二服务器类型是搜索引擎服务器 24。搜索引擎程序允许网络用户在导航到能够通过其浏览器程序 16 向搜索引擎服务器 24 提交查询的其它万维网服务器上的搜索引擎万维网服务器 URL 或站点时，键入关键字查询，从而在万维网上可得到的上百万网页中识别感兴趣的网页。在本发明的最佳实施例中，搜索引擎服务器 24 生成包括、至少部分包括从帐户管理服务器 22 进行的投标过程的结果中得到格式化的相关条目的搜索结果列表。搜索引擎服务器 24 生成

到包括与用户在客户计算机 12 上输入的搜索项有关的信息的文件的
超级文本链接的列表。搜索引擎万维网服务器将此列表以网页的形式
发送给网络用户，并显示于在客户计算机 12 上运行的浏览器 16
之上。目前，搜索引擎万维网服务器的最佳实施例可以通过导航到
5 URL <http://wwwlgoto.com/>上的网页而找到。另外，搜索结果列表网
页，其例示显示在图 7 中，将在下面进行更详细地讨论。

搜索引擎万维网服务器 24 连接到因特网 20。在本发明的最佳实
施例中，搜索引擎万维网服务器 24 包括搜索数据库 40，该数据库中
包括用于响应于用户查询而生成搜索结果的搜索列表记录。另外，
10 搜索引擎万维网服务器 24 还可连接到帐户管理服务器 22。帐户管理
服务器 22 也可连接到因特网。本发明的搜索引擎万维网服务器 24
和帐户管理服务器 22 处理位于客户计算机 12 上用户的不同信息需
要。

例如，位于客户计算机 12 的一类用户可以是网络信息提供商，
15 比如广告网站发起人或拥有位于广告人万维网服务器 14 的广告人网
页 30 的所有者。这些广告网站发起人或广告人可能希望访问驻留在
帐户管理服务器 22 上的存储器 32 中的帐户信息。广告网站发起人
可通过驻留在帐户管理服务器 22 上的帐户参与与其它广告人的竞争
投标过程。广告人可对与广告人的网站内容相关的任何数量的搜索
20 项投标。在本发明的一个实施例中，在将包括搜索项和广告人网站
URL 的搜索列表插入到数据库 40 之前，通过人工编辑过程确定投标
的到广告人站点的搜索项的相关性。在本发明的另一实施例中，利
用在帐户管理服务器 22 的处理器 34 上执行的计算机程序估计搜索
列表中到相应网站的投标搜索项的相关性，其中计算机程序将根据
25 一组预定编辑规则估计搜索项和相应网站。

在利用广告人投标的搜索项的搜索被执行时，更高的报价得到搜
索引擎 24 生成的搜索结果列表页上更有利的位置。在本发明的最佳
实施例中，广告人的报价金额包括每次通过搜索结果列表页上的超

级链接广告人位置被访问时从广告人帐户上扣除的金额。搜索者用计算机输入装置“点击”超级链接，从而开始检索与广告人的超级链接有关的信息的检索请求。可取的是，每次访问或“点击”在搜索结果列表超级链接上将重定向到搜索引擎万维网服务器 24，从而将“点击”与广告人的帐户标识符相关联。这个对搜索者不明显的重定向动作将在利用搜索者点击的搜索结果列表超级链接访问广告人的 URL 之前，访问编码在搜索结果页中的帐户标识信息。这个帐户标识信息以及来自检索请求的信息作为检索请求事件被记录在广告人的帐户中。由于通过这种机制得到的信息最终将帐户标识符与某个 URL 匹配，且其匹配方式是利用本领域已知的常规服务器系统记录所不可能的，因此将维持准确的帐户借记记录。更可取的是，搜索结果列表页上广告人网站描述和超级链接是通过广告人的列表是已付费列表的指示完成的。更可取的是，各个已付费列表显示“广告人费用”，这是对应于广告人为每次通过搜索结果列表转介到该广告人站点所支付的“每点击价格”的金额。

客户计算机 12 上的第二类用户可包括在万维网上寻找特定信息的搜索者。搜索者可通过其浏览器 16 访问驻留在万维网服务器 24 上的搜索引擎网页 36。搜索引擎网页 36 包括一个查询框，搜索者可在其中键入包括一个或多个关键字的搜索项。或者，搜索者可通过超级链接到搜索引擎万维网服务器 24 并定位在储存于远程万维网服务器的网页的查询框来查询搜索引擎万维网服务器 24。当搜索者结束了搜索项的输入时，搜索者可通过点击提供的超级链接将查询传递给搜索引擎万维网服务器 24。然后，搜索引擎万维网服务器 24 将生成搜索结果列表页并将该页传递给在客户计算机 12 的搜索者。

搜索者可点击与搜索结果页上各列表相关的超级文本连接，以访问相应的网页。超级文本链接可访问因特网上任何地方的网页，并包括到位于广告人万维网服务器 14 上的广告人网页 18 的已付费列表。在本发明的最佳实施例中，搜索结果列表还包括没有作为广告

人投标结果放置而是由诸如 INKTOMI、LYCOS 或 YAHOO!搜索引擎等常规万维网搜索引擎生成的未付费列表。未付费超级文本连接还可包括由编辑组人工编入数据库 40 的连接。最可取的是，未付费列表跟在搜索结果页上付费广告人列表后面。

5 图 2 是表示提供给通过常规浏览器查询 16 访问帐户管理服务器 22 的广告人的菜单、显示屏幕和输入屏幕的简图。广告人在将帐户管理服务器 22 的 URL 输入到图 1 的浏览器查询 16 中后，其调用在服务器 22 的处理系统 34 上运行的登录应用，如以下在图 2 的屏幕 110 中所讨论的。一旦广告人登录完毕，处理系统 34 提供一个菜单 120，
10 该菜单具有若干选项和为广告人提供的其它服务。这些将在下面进行更详细讨论的项目使例程被调用，从而或者执行广告人的请求，或者在执行广告人的请求之前请求其它信息。在本发明的一个实施例中，广告人可通过菜单 120 访问多个选项，其中包括请求客户服务 130、查看广告人政策 140、执行帐户管理任务 150、添加钱到广告人帐户 160、管理出现在搜索引擎上的帐户广告 170 以及查看活动报告 180。特定环境帮助 190 也可以在菜单 120 以及所有上述选项中
15 得到。

 在图 3 和 4 中显示了用于两种用户的本发明的最佳实施例的登录程序。图 3 显示了广告人的登录程序 270。图 4 显示了管理和维护本
20 发明的系统和方法的管理人的登录程序 290。如上所述，客户计算机 12 上的广告人或管理人必须首先在步骤 271 或 291 使用浏览器程序来访问帐户管理服务器。在广告人导航到登录页的 URL 从而在步骤 272 或 292 开始登录过程后，帐户管理服务器 22 的处理系统 34 在步骤 274 或 294 调用登录应用。根据该应用，处理器提供请求广告人或管理人的用户名称和口令的输入屏幕 110（图 2）。在步骤 276 或
25 296 将这些信息项提供给本领域已知的安全应用，根据储存在帐户管理服务器 22 的存储器 32 中的数据库中储存的帐户信息进行验证。

 根据图 3，在将用户确认为广告人后，为广告人提供图 2 和菜单

5 屏幕和仅仅对相应广告人帐户的有限读/写访问权限，如步骤 278 所示。还可以在步骤 280 将广告人登录事件 278 记录在审计追踪数据结构中，作为记录在数据库中的广告人帐户记录的一部分。审计追踪最好被实现为数据库 38 中的一系列条目，其中每个条目对应于访问了广告人的帐户记录的事件。非常可取的是，帐户记录的审计追踪信息可以由帐户所有人和其它合适的管理者浏览。

10 但是，如果在图 4 的步骤 295 中将用户验证为管理者，则为管理者提供对所有广告人帐户的特定管理访问权限，如步骤 296 所示。在步骤 297 中将管理者登录事件 296 记录在管理者帐户记录的审计追踪数据结构部分。该审计追踪最好被实现为数据库 38 中的一系列条目，其中每个条目对应于访问了管理者的帐户记录的事件。更可取的是，管理者的审计追踪信息可以由帐户所有人和其它合适的管理者浏览。

15 此外，不像在步骤 282 中向验证过的广告人用户提供一般的广告人主菜单，在步骤 298 中为已验证过的管理者提供访问广告人帐户的数据库 38 的权利。可取的是，向管理者提供能够使管理者选择广告人帐户进行监视的数据库搜索界面。例如，该界面可包括查询框，管理者可在其中键入对应于管理者希望访问的帐户的账号或用户名或联系人名称。当管理者在步骤 299 中选择广告人帐户进行监视时，20 管理者被引导到图 2 的广告人主页面 120，该页也可以由广告人看到。

对位于帐户管理服务器 22 的帐户信息 32 的访问被限于系统上存在帐户记录的用户，因为只为那些用户提供了有效的登录名称和口令。在帐户管理服务器 22 的数据库 38 中存储了口令和登录名称信息以及用户的其它帐户信息，如图 1 所示。包括登录用户名和口令的帐户信息通过单独的在线登记过程输入到图 1 的数据库 38 中，25 该过程在本发明的范围之外。

图 5 是显示数据库中包括在各个广告人帐户记录 300 中的信息类型图。首先，广告人帐户记录 300 包括用户名 302 和口令 304，用于

如上所述的在线验证。帐户记录还包括联系信息 310（如联系人名称、公司名称、街道地址、电话、电子邮件地址）。

5 联系信息 310 最好在广告人在通知选项中请求通知关键广告人事件时被用于指向到广告人的通信，这一点将在下面进行讨论。帐户记录 300 还包括计费信息 320（如当前结余、信用卡信息）。计费信息 320 包括在广告人选择向广告人帐户加钱的选项时被访问的数据。另外，诸如当前结余的某些计费信息可以触发在通知选项中请求通知的事件。帐户记录 300 的审计追踪选择 325 包括被访问的帐户记录 300 的所有事件的列表。管理者或广告人每次访问或修改帐户记录 300 时，描述帐户访问和/或修改事件的短条目将被附在启动该事件的管理者或广告人帐户的审计追踪选择 330 上。然后，审计追踪信息可用于帮助建立账号所有人在账号上的交易历史。

15 广告信息选择 330 包括进行本发明的在线投标过程所需的信息，其中在搜索引擎生成的搜索结果列表中为网站描述和超级链接确定位置。每个用户帐户 300 的广告数据 330 可以被组织为零或多个子帐户 340。各个子帐户 340 包括至少一个搜索列表 344。各个搜索列表对应于对搜索项的投标。广告人可利用子帐户组织对多个搜索项的多次投标，或者组织为多个网站的投标。子帐户在广告人设法跟踪目标市场部分的性能方面也特别有用。引进子帐户的超结构，以方便广告人设法组织其广告成果，并且不影响本发明的工作方法。或者，广告信息部分不必包括子帐户的添加组织层，但可以仅包括一个或多个搜索列表。

25 搜索列表 344 对应于搜索项/投标配对，并包括进行在线竞争投标过程的关键信息。可取的是，各个搜索列表包括以下信息：搜索项 352、网站描述 354、URL 356、报价 358 和标题 360。搜索项 352 包括可具有公共词的一个或多个英文关键词（或任何其它语言）。各关键词又包括字符串。搜索项是竞争在线投标过程的对象。广告人选择与广告人网站的内容有关的搜索项进行投标。理想的是，广

告人可选择寻找有关广告人网站的信息的搜索者很可能输入的项作为搜索项目标，尽管也可以选择较少的共同搜索项来保证用于投标的相关搜索项的综合性覆盖。

5 网站描述 354 是广告人网站的内容的简短文字描述(最好少于 190 个字符) 并可能作为广告人输入的一部分显示在搜索结果列表中。搜索列表 344 还可以包括网站标题 360，该标题可作为到广告人条目的超级链接标题显示在搜索结果列表中。URL 356 包括广告人网站的统一资源定位器地址。当用户点击广告人的搜索结果列表条目中提供的超级连接时，URL 被提供给浏览器程序。浏览器程序又通过
10 上述的重定向机制访问广告人站点。URL 还可以作为广告人条目的一部分显示在搜索结果列表中。

报价 358 最好是广告人对某个列表的出价金额。可取的是，在每次用户执行对相应搜索项执行搜索并利用搜索结果列表超级链接将搜索者引导到广告人网站时，从广告人的预付帐户上扣除该金额或者将该金额记录在用于收费的广告人帐户中。最后，等级值是最好
15 由图 1 所示的帐户管理服务器 22 的处理系统 34 在每次广告人进行一次报价或搜索输入搜索查询时动态生成的值。广告人的搜索列表的等级值确定在对相应搜索项执行搜索时搜索结果列表中生成的广告人条目的放置位置。可取的是，等级值是由与报价 358 的直接联系确定的顺序值；报价愈高，等级值愈高，以及搜索结果列表上更有利的放置位置。最可取的是，将等级值 1 分配给最高报价，并且
20 接连的更高顺序值(如 2, 3, 4, ...)与接连的更低等级相关并被分配给接连的更低报价。

一旦登录完毕，广告人可执行在图 2 的菜单 120 中设置的若干直接任务，包括浏览对广告人的规则和政策列表，以及请求客户服务帮助。这些项目将调用例程来执行请求。例如，当选择了“客户服务”时，将显示输入屏幕 130 以便广告人选择请求的客户服务类型。另外，可以在屏幕 130 上提供表格，以便广告人可将客户意见输入
25

到基于万维网的输入表格中。

当选择了“浏览广告人政策”时，将通过图 1 的帐户管理服务器 22 的处理系统 34 调用例程。如图 2 所示，例程将显示信息化网页 140。网页 140 陈述当前有效的广告人政策（如“所有搜索列表描述必须明确地与搜索项有关”）。

图 2 的菜单 120 还包括“帐户管理”选择，该选择允许广告人浏览和改变广告人的联系信息和计费信息或在适合的情况下更新广告人的访问简档（如果有的话）。提供本领域众所周知并类似于上述那些的基于万维网的表格来更新帐户信息。

“帐户管理”菜单还包括使广告人能够浏览广告人帐户的交易历史的一个选择。在“浏览帐户历史”选择中，广告人可调用浏览过去帐户交易的列表的例程（如向帐户加钱、添加或删除投标搜索项或改变报价）。可以实现其它例程以允许广告人显示特定类型或发生在特定时间的交易历史。可以从上述的图 5 的审计追踪列表 325 得到交易信息。可以用软件实现可点击按钮，可以提供本领域已知的基于万维网的表格和/或菜单以允许广告人指定这些限制。

另外，图 2 的“帐户管理”菜单 150 包括允许广告人设置通知选项的选择。在这种选择下，广告人可选择将使系统在某些关键事件发生时通知广告人的选项。例如，广告人可选择设置一种选项，该选项使系统在广告人帐户结余低于特定水平时给广告人发送常规的电子邮件消息。以这种方式，广告人可接收“警告”，从而在帐户被暂停之前（意味着广告人的列表将不再出现在搜索结果列表中）补充帐户。另一个广告人可能希望通知的关键事件是广告人的列表在为特定搜索项所生成的搜索结果列表中的位置的变化。例如，广告人可能希望在如果另一广告人对特定搜索项的出价比他高（意味着广告人的列表将出现在比从前低得多的搜索结果列表页的位置上），系统就给它发送常规的电子邮件消息。当系统指定关键事件之一发生时，由各个受影响的搜索列表触发数据库搜索。然后，系

统将根据广告人的帐户中指定的通知选项执行适当的通知例程。

回来参考图 2，菜单 120 中还出现一个允许广告人将钱添加到广告人帐户的选择，以便广告人用帐户中的资金来支付通过搜索结果页转介到广告人站点的费用。可取的是，只有在其帐户中具有资金的广告人可将其已付费列表包括在生成的任何搜索结果列表中。最可取的是，满足所选商业标准的广告人可选择不考虑帐户结余而定时地支付反映搜索引擎生成的到广告人站点的实际转介引起的费用的帐单金额，而不是维持正的帐户结余。图 6 中更详细地显示了在调用“向帐户加钱”时执行的进程，该进程从步骤 602 开始。当在步骤 604 中点击“向帐户加钱”选择时，调用一个函数，该函数接收表示广告人的数据并从数据库检索广告人帐户。然后，执行进程将广告人的缺省计费信息储存并在步骤 606 中为广告人显示该缺省计费信息。显示的计费信息包括要添加的缺省金额、缺省付款方式以及缺省票据信息。

在本发明的最佳实施例中，广告人可通过利用信用卡在线并基本上实时地添加资金，尽管利用其它付款方式也肯定在本发明的范围内。例如，在本发明的备选实施例中，广告人可通过本领域已知的的电子资金验证机制从诸如付款卡的广告人银行帐户将所需金额转帐来向其帐户添加资金，其方式类似于授予 Gifford 的美国专利 No. 5724424 中陈述的方式。在本发明的另一备选实施例中，广告人可利用常规的基于纸的支票来向其帐户添加资金。在那种情况下，添加的资金可以通过人工输入在帐户记录数据库中更新。票据信息包括有关付款方式的更多细节。例如，对于信用卡，票据信息可包括有关信用卡名称的数据（如 MasterCard、Visa 或 American express）、信用卡号、信用卡的到期日以及信用卡的计费信息（如计费名称和地址）。在本发明的最佳实施例中，出于安全性目的，只向广告人显示部分信用卡号。

从例如储存在帐户数据库中的不变状态得到显示给广告人的缺省

值。在本发明的一个实施例中，储存的计费信息值可包括广告人在上次（最近）调用并完成对广告人帐户添加钱的进程时设置的值。在基于万维网的表中向广告人显示缺省计费信息。广告人可点击基于万维网的表上适当的文本输入框并改变缺省计费信息。在广告人完成改变后，广告人可点击表上提供的超级链接“提交”按钮从而在步骤 608 中请求系统更新计费信息和当前结余。一旦广告人请求了更新，系统就调用一个确认广告人提供的计费信息的函数，并重新显示给广告人进行确认，如步骤 610 所示。确认计费信息以只读形式显示并且广告人可能不能改变该信息。

10 确认步骤如下所述地工作。如果付款是从广告人的外部帐户上扣除的，则可以利用授予 Gifford 的美国专利 No. 5724424 中陈述的系统对付款进行验证、授权和完成。但是，如果付款方式是用信用卡，系统则调用确认算法，该算法利用授予 Stein 等人的美国专利 No. 5836241 中陈述的方法确认信用卡号。确认算法还通过与当前系统日期和时间的直接比较确认到期日。另外，该函数在广告人确认之前将新的值存储在临时实例中。

一旦广告人确定显示数据是正确的，广告人可点击页面上提供的“确认”按钮来表示帐户应该在步骤 612 中更新。在步骤 612 中，系统调用一个函数，该函数将钱添加到适当的帐户结余中，更新广告人的计费信息并将计费信息附在广告人的付款历史上。将广告人的更新计费信息从临时实例存储为不变状态（如帐户记录数据库）。

在步骤 612 调用函数中，可以由系统在步骤 614 调用信用卡付款函数。在本发明的备选实施例中，可以根据付款方式的更新值通过定义多种付款方式来调用诸如付款卡付款的其它付款函数。

25 如果付款方式是信用卡，则在步骤 616 立即贷记用户帐户，用户的信用卡已经在步骤 610 进行了确认。显示给出添加钱交易状态的屏幕，显示交易号和新的当前结余，以反映刚刚完成的信用卡交易所添加的金额。

在本发明的备选实施例中，在已经将钱添加到帐户之后，可以在步骤 616 结束添加钱的进程时在子帐户之间分配添加到帐户的金额。如果广告人没有子帐户，则帐户中所有的钱是总拨款。但是，如果广告人具有超过一个的子帐户，则系统将显示确认和缺省消息，提示广告人“在子帐户之间分配钱”。

在图 6 的步骤 616 之后，钱添加到广告人帐户时可调用菜单选择“在子帐户之间分配钱”，或者可以在图 2 所示的“帐户管理”菜单 170 中调用该菜单选择。可从广告人主页 120 访问“帐户管理”菜单 170，如图 2 所示。该“在子帐户之间分配钱”菜单选择允许广告人在广告人的子帐户中分配广告人帐户的当前和任何未决结余。然后，系统更新帐户结余。当前结余分配将实时进行，但将以持久状态储存未决的结余分配。在处理未决结余的付款时将调用例程来更新子帐户结余，以反映未决结余分配。如果请求，这时将向广告人发送自动通知。这种直观在线帐户管理和分配允许广告人快速而有效地管理它们的在线广告预算。广告人可用资金补充其帐户并分配它们的预算，所有这些都一个简单的基于万维网的对话中。基于计算机的实现避免了耗时、高成本的手工输入广告人的帐户交易。

“在子帐户之间分配钱”例程从广告人在上述执行点通过调用适当的菜单选择表示意欲分配钱时开始。当广告人表示意欲分配时，由系统调用一个函数以确定当前结余中是否有资金未决（即未激活的赊帐）还没有分配到广告人的子帐户中，并显示结余选择选项。在本发明的最佳实施例中，建立帐户实例并从持久状态设置未决当前结余帐户字段。

如果没有未分配的未决资金，系统可显示该帐户总的以及各个子帐户的当前可用结余。然后，广告人在子帐户之间分配当前可用结余并提交更新结余的请求。调用计算并显示子帐户结余的当前累计的函数。当前累计被存储在临时变量中，该变量被设置为特定广告人的所有子帐户的当前结余总和。该函数还确认新的可用子帐户结

余，以保证总数不超过授权金额。如果新的广告人设置可用于子帐户结余不超过授权金额，则调用将更新持久状态的所有子帐户结余的函数并以只读格式显示该更新。

5 如果当前帐户结余中存在未决资金，未决资金必须从当前结余单独分配。然后，未决资金将在收到资金时添加到可用当前结余中。因此，函数必须提示广告人在分配未决资金或分配可用资金之间进行选择。分配未决资金选择的工作方式与上述分配可用资金选择的方式非常相似。在广告人选择分配未决资金后，调用例程以显示帐户和子帐户的当前未决结余。广告人在活动之间分配未决子帐户结余并提交更新结余的请求。调用计算并显示未决子帐户结余的当前
10 累计的函数。该函数还确认新的未决子帐户分配，以保证分配不超过任何授权金额。未决分配的当前累计被设置为广告人的所有子帐户的当前未决结余之和。如果新的用户设置未决子帐户结余或这些结余的总数不超过任何授权金额，则该函数将更新不变状态中所有
15 未决子帐户分配，如数据库中的广告人帐户，并以只读格式显示该更新。

如上所述并如图 2 所示，可以从广告人主菜单 120 调用显示帐户管理菜单 170 的例程。除了上述“在子帐户之间分配钱”选择以外，剩下的选择全部在某种程度上使用数据库上广告人帐户中出现的搜索列表，并还可能影响搜索结果列表中广告人的条目。因此，在这一点上需要对搜索引擎生成的搜索结果列表的进一步描述。
20

当远程搜索者访问搜索引擎万维网服务器 24 上的搜索查询页并根据上述程序执行搜索请求时，搜索引擎万维网服务器 24 最好生成并显示搜索结果列表，其中搜索结果列表中各搜索列表的搜索项字段中的“标准化”条目正好与远程搜索者输入的标准化搜索项查询
25 匹配。查询和搜索列表中使用的搜索项的标准去掉了搜索和网站发起人输入的搜索项的公共不规则性，比如大写字母、复数，以便生成相关结果。但是，确定搜索列表的搜索项字段和远程搜索者输入

5 的搜索项查询之间匹配的备选方案都肯定在本发明的范围内。例如，可以采用本领域已知的字符串匹配算法来生成匹配，其中搜索列表搜索项和搜索项查询的关键字具有相同的字根但不完全相同（例如 computing（计算）对 computer（计算机））。或者，同义词的同义词典数据库可以存储在搜索引擎万维网服务器 24，从而可以为具有同义词的搜索项生成匹配。还可以采用局部方法学来精炼某些搜索。例如，对“面包房”或“食品杂货店”的搜索可以限制在所选城市、邮政编码或电话区号内的那些广告人。可以通过帐户管理服务器 22 上存储器 32 中存储的广告人帐户数据库的交叉引用得到这种信息。10 最好，可以采用国际化方法学来精炼美国以外用户的搜索。例如，可以通过例如广告人帐户数据库的交叉引用来生成特定国家或语言的搜索结果。

图 7 显示了本发明的一个实施例中所用的搜索结果列表显示的例示，该例示是对项“zip drives（zip 驱动器）”的搜索而得到的前几个条目的显示。如图 7 所示，搜索结果列表中诸如条目 710a 的单条目由网站的描述 720 和超级链接 730 组成，其中所述描述最好包括标题和简短文字描述，而超级链接将在搜索者点击时将搜索者的浏览器引导到所描述的位置所在的 URL。URL 740 也可以显示在搜索结果列表条目 710a 中，如图 7 所示。当浏览图 7 的搜索结果项显示 710 的远程搜索者选择或“点击”搜索结果项显示 710 的超级链接 730 时，发生搜索结果项的“点击直通”。为了完成“点击直通”，搜索者的点击应该被记录在帐户管理服务器中并经上述重定向机制重定向到广告人的 URL。

25 搜索结果列表条目 710a-710h 还可以显示广告人的搜索列表的等级值。等级值是顺序值，最好是由图 1 的处理系统 34 生成并分配给搜索列表的数字。可取的是，通过用软件实现的进程分配等级值，该等级值建立报价、等级和搜索列表的搜索项之间的关系。该进程收集所有与特定搜索项匹配的搜索列表，对搜索列表案最高到最低

的报价顺序排序，并将等级值分配给各个按顺序的搜索列表。最高的报价得到最高的等级值，次最高的报价得到次最高的等级值，一直进行到最低报价得到最低等级值，最可取的是，最高等级值为 1，然后按顺序地将递增的顺序值（如 2, 3, 4, ...）分配给的递减等级。

5 图 7 中说明了等级值和报价之间的校正，其中已付费搜索列表条目 710a 到 710f 显示该条目的广告人报价 750a 到 750f。可取的是，如果具有同一搜索项的两个搜索列表页具有同一报价，则为较早收到的投标分配更高等级值。未付费列表 710g 和 710h 不显示报价并显示在最低等级已付费列表的后面。可取的是，如果搜索结果页中列表数量不足以填满 40 个位置，则显示未付费列表。通过利用本领域已知的对象分布数据库和文本搜索算法的搜索引擎生成未付费列表。这种搜索引擎的例示可以由 Inktomi 公司操作。远程搜索者输入的原始搜索查询被用于通过常规搜索引擎生成未付费列表。

15 如图 2 的活动管理菜单 170 所示，为广告人提供若干选择来管理搜索列表。首先，在“改变报价”选择中，广告人可改变当前帐户中搜索列表的投标。图 8 中显示了系统调用的改变投标函数的进程。在广告人通过选择“改变投标”菜单选项来表示意欲改变投标时，系统搜索数据库中用户的帐户并显示整个帐户或广告人帐户中的缺省子帐户的搜索列表，如步骤 810 所示。搜索列表可以被组合到广告人定义的子帐户中并可包括一个或多个搜索列表。一次可能只显示一个子帐户。显示还最好应该允许广告人改变所选子帐户，如步骤 815 所示。然后，屏幕显示将显示所选子帐户的搜索列表，如步骤 820 所示。

25 图 9 显示了步骤 810 中显示给广告人的屏幕显示的例示并将在下面进行讨论。为了改变报价，广告人用户可通过为搜索项将新的报价输入到新投标输入字段中来为广告人已具有投标搜索项指定新投标。如上所述，在图 8 的步骤 820 中向广告人显示其输入的投标改变。为更新显示页的投标，广告人在图 8 的步骤 830 请求更新改变

结果。广告人可将这种请求通过各种方式发送给帐户管理服务器，其中包括点击按钮图形。

如图 8 的步骤 840 所示，在接收到更新广告人投标的请求后，系统为每个显示的搜索列表、等级值和成为匹配搜索项字段的最高等级搜索列表所需的报价计算新的报价。可取的是，系统接着在步骤 850 显示改变显示屏。在用户确认改变之后，系统通过将变化值写入数据库中的帐户来更新持久状态。

以表格形式显示搜索列表数据，其中一个搜索列表对应于表 900 的一行。搜索项 902 显示在最左列，然后是当前的报价 904，以及搜索列表的当前等级 906。当前等级后面是题为“成为 #1 的报价”列 907，该列被定义为成为显示的搜索项的最高等级搜索列表所需的报价。每行最右边的列包括最初设置为当前报价的新投标输入字段 908。

如图 9 所示，搜索列表可以显示为“子帐户”。个子帐户包括一个搜索列表组，其中多个子帐户存在于一个广告人帐户。各个子帐户可以显示在具有单独页的单独显示页上。广告人最好应该能够通过操作图 9 所示显示屏上的下拉菜单 910 来改变显示的子帐户。另外，无法在一页中完整显示的搜索列表组可以分成多页，这些页可以通过操作下拉菜单 920 单独浏览。同样，广告人最好应能够通过直接点击位于图 9 的显示页面上的下拉菜单 920 来改变显示的页。广告人可通过将搜索列表的新报价输入到新投标输入字段 908 来指定所显示搜索列表的新投标。为了更新广告人输入的改变的结果，广告人点击按钮图形 912，从而将更新请求传送到帐户管理服务器，该服务器如上所述更新投标。

图 2 的“帐户管理”菜单中列出的许多其它选择作为上述“改变投标”功能的变型。例如，如果广告人选择“改变等级位置”选项，可向广告人显示类似于“改变投标”功能中所用的图 9 的显示的显示。但是，在“改变等级位置”选项中，将用“新等级”字段代替

“新投标”字段，广告人在其中输入搜索项新的所需等级位置。在广告人请求更新等级后，系统接着用本领域技术人员方便得到的各种算法中的任何算法计算新的报价。例如，系统可调用定位搜索数据库中具有所需等级/搜索项组合的搜索列表的例程，检索所述组合的相关报价，然后计算高出 N 分的报价；其中例如 N=1。在系统计算新的报价并向广告人显示了只读的确认显示后，系统在收到广告人的认可后更新报价和等级值。

图 2 的帐户管理菜单上的“修改列表成分”选择还可以生成类似于图 9 格式的显示。当广告人选择“修改列表成分”选项时，广告人可通过为各个搜索列表建立的基于万维网的表输入对搜索列表的 URL、标题或描述的改变。类似于上述过程，用于 URL、标题和描述字段的表开始可包括作为缺省值的老 URL、标题和描述。在广告人输入了所需改变后，广告人可向系统发送请求来更新改变。然后，系统显示只读的确认屏幕，并接着在广告人认可改变之后将这些改变写入持久状态（如用户帐户数据库）。

可以实现与以上描述类似的过程，以改变与搜索列表有关的任何其它外围选项；例如，改变与投标搜索项有关的匹配选项。改变所需的任何对投标或等级的重新计算还可以用与上述过程相似的方式确定。

在“删除投标搜索项”选项中，系统检索广告人帐户中的所有搜索列表并将搜索列表显示于类似图 9 所示的结构和格式中。各个搜索列表条目可包括广告人点击的复选框而不是新的投标字段。然后，广告人通过点击来将复选 (X) 标记放置在要删除的搜索项后面，尽管也可以使用本领域已知的任何其它方法来从网页上的列表中选择一项或多项。在广告人选择了所有要删除的搜索列表并请求系统更新改变之后，系统最好显示所请求改变的只读确认，并仅在广告人认可了这些改变之后更新广告人的帐户。从搜索数据库 36 中去掉了“删除”的搜索列表并且这些搜索列表将不会出现在后续搜索中。

但是，搜索列表将保留作为广告人帐户记录的部分，用于计费 and 帐户活动监视目的。

在“添加投标搜索项”选项中，系统为广告人提供具有若干与搜索列表中的元素对应的输入字段。然后，广告人输入对应于各个搜索列表元素的各个字段信息，包括搜索项、网站 URL、网站标题、
5 网站描述和报价以及任何其它相关信息。在广告人完成了数据输入并向系统表示了之后，系统向广告人返回只读确认屏幕。然后，系统建立新的搜索列表实例并在收到广告人的认可之后将其写入帐户数据库和搜索数据库。

10 可取的是，图 2 的“帐户管理”菜单 170 为广告人提供“得到有关投标搜索项的建立”的选择。在这种情况下，广告人将投标搜索项输入到显示给广告人的表格驱动查询框。系统读取广告人输入的搜索项并生成另外的相关搜索项的列表以帮助广告人定位与广告人网站的内容有关的搜索项。可取的是，利用诸如应用于以软件实现的
15 投标搜索项数据库和/或同义词典数据库的字符串匹配算法的方法生成另外的搜索项。广告人可从系统生成的列表中选择要投标的搜索项。在这种情况下，系统向广告人显示上述用于“添加投标搜索项”的输入字段，其中有一个用于输入各个所选搜索项的搜索列表的表格。可取的是，将所选搜索项作为缺省值插入到各个搜索列表
20 的表格中。如果需要，也可以将其它搜索列表成分的缺省值插入到表格中。因此，在一个实施例中，公开的系统接收与数据库搜索系统上广告人有关的搜索项列表，根据数据库搜索系统上其它广告人的搜索项确定候选搜索项，并从候选搜索项中推荐另外的搜索项。在另一实施例中，公开的系统提供得到广告人的搜索项，响应于收
25 到的搜索项，生成另外的相关搜索项列表，并从另外的相关搜索项列表中得到广告人的所选搜索项。

图 2 的“帐户管理”菜单 170 最好还为广告人提供“费用预测”选择。在此选择中，广告人指定他希望预测“日运转费用”和“满

期剩余天数”的搜索列表或子帐户。系统根据成本预测算法计算预测，并在只读屏幕上向广告人显示预测。预测可以利用本领域已知的若干不同算法进行计算。但是，由于通过用报价乘以特定时段内在该报价下搜索列表收到的点击次数来计算搜索列表的成本，因此

5 每种成本预测算法一般必须确定搜索列表每月估计的点击次数。对搜索列表的点击可以通过本领域众所周知的软件实现的计数机制来跟踪。可以跟踪一段时间上对所有搜索列表的点击。这种数据可用于生成估计的每个月的总点击数以及用于各个搜索项。对于特定搜索项，确定每天的估计搜索次数，并将之乘以点击成本。然后，将

10 这个乘积乘以所述搜索列表的等级下平均点击次数与平均出现次数之比得到“日运转费用”。可以用当前结余除以日运转费用得到帐户资金的耗尽或“满期”预测天数。

本发明的一个实施例使成本预测算法基于简单的预测器模型，该模型假定每个搜索项以相似的方式工作。该模型假定广告人的搜索

15 列表的等级在整个月将保持恒定没有波动。该算法具有以下优点：实现简单并且计算速度快。预测器模型基于以下事实：特定搜索者列表的例如总点击次数的点击直通率，或者转介，被认为是搜索列表等级的函数。所以，该模型假设各个搜索项的使用曲线，也就是所绘制的搜索列表的点击次数随搜索列表等级变化的曲线与所有搜索

20 项的使用曲线相似。因此，所有搜索项的所有点击总数在一段时间上外插的已知值、给定等级下所有搜索项的所有点击总次数，以及所选搜索项的所有点击总次数可以采用简单的比例来确定给定等级下所选搜索项的所有点击总次数。然后，用为所选等级下所选搜索项估计的每日所有点击的总次数乘以广告人对该等级下搜索项的

25 当前报价，以确定日平均预测。另外，如果已知特定搜索项或搜索项排序与一般模式明显不同，则针对搜索项、广告人或其它参数的校正值得被引进到精细调节预测成本估计中。

最后，图2的“帐户管理”菜单170提供查看与广告人活动有关

的信息的若干选择。“查看子帐户信息”选择显示与所选子帐户有关的只读信息。“查看搜索项列表”选择显示广告人所选搜索项和相应的 URL、出价和等级的列表，其中搜索项最好按子帐户分组。广告人还可以查看从来自系统在收到来自广告人的请求搜索项后生成的只读显示的搜索项列表中选择的一组搜索项的当前最高投标。

对于要求更全面搜索列表活动报告的广告人，可以从图 2 的广告人主页 120 中选择“查看报告”选项。在本发明的一个实施例中，“查看报告”选项生成在当前日期前最多一年的全面报告。例如，可得到即时之前 7 天中的每天的日报告、之前四周的周报告、之前 12 个月的月报告以及之前四个季度的季报告。还可以根据广告人的兴趣得到另外的报告。其它预定报告类型可包括在以下时段的活动跟踪：从帐户初始起，本年度截至现在为止、年度、本季度截至现在为止、本月截至现在为止以及本周截至现在为止。报告类别可包括：详细报告，可根据广告人帐户、根据搜索列表和根据 URL 查看；以及摘要报告，可根据广告人帐户和根据子帐户查看。报告可包括标识数据，比如广告人帐户和子帐户名称、报告覆盖的日趋和报告类型。另外，包括可包括关键搜索列表帐户数据，比如当前结余，未决当前结余、平均日帐户借记以及运行费用。此外，包括还可包括关键数据，比如搜索项、URL、投标、当前等级和点击次数、对搜索项进行的搜索次数、出现次数（搜索列表出现在搜索结果列表中的次数）、以及点击直通率（定义为点击次数/出现次数）。可取的是，报告至少可以用 HTML 查看选项得到，用于通过浏览器程序、打印或下载进行查看。但是要注意，可以得到其它查看选项，比如 Adobe Acrobat、PostScript、ASCII 文本、，电子数据表交换格式（如 CSV，制表符定界的）和其它众所周知的格式。

当广告人选择了“查看报告”选项时，系统调用显示可得到的报告类型、数据、排序和浏览选项的函数。系统最好建立具有以下字段的报告实例：报告类型、报告日期、报告排序和查看选项，其中

所有字段初始设置为零。一旦广告人定义了上述参数，系统根据广告人设置参数调用生成请求报告，并根据查看选项参数显示该报告的函数。

5 最后，本发明的最佳实施例实现环境特定帮助选项，广告人可能在登录后的任何时候请求这种帮助。帮助选项可以实现为位于系统生成显示页上的小图标或按钮。广告人可以点击显示页上的这个图标或按钮图形来请求帮助，其中系统在这个图标或按钮图形上生成并显示了固定在用户正在查看的特定显示的函数的帮助页。帮助可以实现为单独的显示页、可搜索索引、对话框或本领域中众所周知的任何其它方法。

10 图 10-20 说明在诸如以上参考图 1-9 所述的根据位置付费的市场体系中为位置发起人或广告人提供搜索项推荐的方法和装置的特定实施例。所公开的实施例提供了用于数据库搜索系统的方法。方法包括维护包括相关搜索项的搜索列表数据库，接收与广告人有关的搜索项列表，向广告人推荐另外的搜索项。其它公开的实施例提供用于数据库搜索系统的数据库操作方法，所述数据库搜索系统储存包括广告人所选的搜索项的广告人搜索列表。该方法包括蜘蛛式搜索特定网站以得到广告人的初始广告人搜索项列表。该方法还包括利用其它广告人的搜索项过滤初始广告人搜索项列表，并储存在广告人的搜索列表数据库搜索列表中，用过滤了的搜索项形成搜索列表。

25 公开的实施例还包括数据库搜索系统，该系统包括搜索项数据库，该数据库中的每个搜索项与一个或多个广告人相关。程序代码被配置为根据数据库中的搜索项为广告人推荐另外的搜索项。再者，本文所公开的实施例提供一种用于数据库搜索系统的方法，其中包括得到广告人的搜索项，并且作为响应，生成另外的相关搜索项列表。该方法还包括从另外的相关搜索项列表得到广告人所选搜索项。

在这里给出的实施例中，利用蜘蛛式搜索和协作过滤来识别可能

推荐给广告人的搜索项。以下介绍首先分别描述蜘蛛式搜索和协作过滤技术，然后给出可以如何将这两种技术进行组合。

5 蜘蛛式搜索是用于下载生根于统一资源定位器（URL）的网站的简单技术。程序下载 URL 给出的主页，然后扫描其上到其它页的超级链接并下载这些页。蜘蛛式搜索过程继续到程序达到预定链接深度、下载预定页数或达到某种其它停止标准为止。下载页面的顺序可以是广度优先或深度优先。在广度优先蜘蛛式搜索中，程序将新的 URL 添加到要下载的页列表的最后；在深度优先蜘蛛式搜索中，程序将新的 URL 添加到要下载的页列表的前面。这些算法对本领域的技术工程师是直观且众所周知的。有关这些技术的更多信息可通过查询（ResearchIndex, <http://citeseer.nj.nec.com>, Cho, Molina, and pages）“有序 URL 中的有效爬行”和“人工智能原理”（Nilsson, ISBN 0934613109）来找到。

10

15 本文所述的一些实施例利用蜘蛛式搜索来寻找直接出现在广告人网站上的搜索项。从广告人找到的根开始，根据本发明的实施例的方法和系统首先横向下载页面并在其中扫描搜索项。它记录它找到的每一个提供商数据库表示在上个月已搜索的项。作为例示，如果页面上的文字包括短语“tropical fish store（热带鱼商店）”，则程序将找到六个项：“tropical, ”、“fish, ”、“store, ”、“tropical fish”、“fis store”和“tropical fish store”。程序利用质量量度给这些项打分，将高于特定阈值的项添加到其推荐列表中。在最佳实施例中，质量量度考虑两个因素：搜索项在万维网上的常见度，以及用户用它进行搜索的频率。当程序累积了足够的推荐时，它按它们的质量或它们出现在下载页中的次数对它们进行排序并返回列表。

20

25 当前实施例的蜘蛛式搜索组件在三个方面不同于以前的工具。首先，它直接查看广告人网站中的页面，而不是下载广告人网站中没有的可能完全无关的其它页。其次，它查看网页上的所有文字，而不是仅查看 DESCRIPTION(描述)和 KEYWORD(关键字)标记中的

词。第三，它利用其质量量度去掉不好的搜索项，甚至不把它们提供给广告人。

5 协作过滤是根据用户相似性进行推荐的技术。作为例示，如 Amazon.com 的公司利用协作过滤技术推荐书，一旦客户通过 www.amazon.com 上的在线服务购买了几本书，Amazon.com 就将该客户与它数据库中的其它客户进行比较来推荐新书。当它发现另一进行了许多相同购买的客户时，它向该客户推荐各个选择。当前实施例将这种思路延伸到为根据位置付费搜索引擎上广告人推荐搜索项。

10 例如，假定典型提供商具有 50000 个广告人的数据库。该数据库的一部分可能看起来是这样的：

	鱼	金枪鱼	比目鱼	饵	虫	汽车
Joe 的鱼	X	X	X	-	-	-
Rick 的汽车店	-	-	-	-	-	X
Bill 的滑车	X	-	-	X	X	-

15 表中的 X 表示广告人对某一项进行了投保。在海鲜的例示中，开始对“fish”感兴趣的广告人与 Joe 和 Bill 相似，并且程序将推荐“tuna, ”、“halibut, ” “bait, ” 和 “worms” 。如果广告人精炼其搜索项以包括“tuna” 但去掉“bait” ，则他就不再与 Bill 相似，并且程序将停止推荐“worms” 。与 STF 一样，当前的发明允许广告人迭代地接受和拒绝一些词直到他对推荐列表感到满意。

20 定量分析上，协作过滤计算新广告人和所有现有广告人之间的皮尔逊相关性。为了计算这种相关性，为广告人/项表格中的各个条目分配数字等级。在一个可能的实施例中，最高等级是 5，表示某项是广告人站点的最佳描述，而最低等级为 0，表示某项无关。在最佳实施例中，广告人为他投标了的每一项取得等级 5，而为其它每个项得到 UNKNOWN（未知）等级。新广告人为广告人已接受的项取得

等级 5，为他拒绝的项得到 1，而为每个其它项得到 2。因此，新广告人和所有现有广告人之间的皮尔逊相关为：

$$\rho_a = \frac{\sum_t (r_{n,t} - \bar{r}_n)(r_{a,t} - \bar{r}_a)}{\sigma_n \sigma_a}$$

在此公式中， n 是新广告人， ρ_a 是他与广告人 a 的相关性， $r_{n,t}$ 是他分配给项 t 的等级，而 \bar{r}_n 和 σ_n 是其等级的平均和标准偏差。有下标 a 的项具有对现有广告人的相应意义。对所有搜索项求和。用广告人等级的平均代替 UNKNOWN 等级，从而从等式中消掉所有具有 UNKNOWN 的项。相关范围在 -1 和 1 之间，其中 0 表示不相关，正相关表示两个广告人具有相似的等级。该公式是统计学所公知的并且是本领域的技术人员所熟悉的。更多的细节可通过查询“工程师和科学家的统计方法手册”（Wadsworth[ed], ISBN 007067678X）得到。

一旦协作过滤计算出新广告人和所有现有广告人之间的相关，它预测各项对新广告人是好搜索项的可能性。它通过计算各项的平均等级进行预测，其中广告人对平均的贡献由他对新广告人的相关而确定。具有高相关性的广告人得到全加权；具有低相关性的广告人得到很低的加权；而具有零相关性的广告人没有加权。用于这种预测的一个公式是：

$$e_t = \bar{r}_n + \frac{\sum_a (r_{a,t} - \bar{r}_a) \rho_a}{\sum_a \rho_a}$$

在这个公式中， n 是新广告人而 e_t 是对项 t 的估计等级。剩下的各项与上一公式具有相同的意义。对所有现有广告人求和。同样用广告人的已知等级的平均代替 UNKNOWN 等级，从而从等式中将其消掉。公式是加权和，他估计与原始等级一样在相同的 0 到 5 范围内的等级。如果所有高度相关的广告人对某项给出高等级，则该项

得到高的估计。协作过滤的输出是按其估计等级排序的搜索项列表。

5 这些公式提供了根据相似性计算等级的直接技术。存在许多相似的公式和变型。例如，在进行预测时，更好的作法常常是仅仅对 10 - 20 个最密切相关广告人求加权平均，而不是对所有广告人进行加权平均。还有用于提高计算效率的技术，或者不用相关量度或距离量度进行协作过滤的技术。这些变型很容易在有关协作过滤的文献中找到，同时本发明的实施例不限于任何这些变型。有关不同协作过滤算法的优点和缺点的更多细节可以在 GroupLens 的网站 <http://www.cs.umn.edu/Research/GroupLens> 中找到。

10 已知蜘蛛技术和协作过滤的核心构建模块，根据本发明的一个实施例的完整系统和方法如下：从初始的接受和拒绝搜索项列表开始，运行协作过滤算法，允许广告人接受和拒绝新的项，然后重新运行协作过滤。当广告人对其接受项列表感到满意时此过程结束。该技术用以下三种方式得到其初始的接受项列表：或者直接来自于广告人，或者来自现有广告人的投保列表，或者来自通过对新广告人的网站允许万维网蜘蛛所返回的推荐列表。这最后一种方法是最佳实施例。在利用万维网蜘蛛时，它推荐的搜索项得到在 4.9 到 2.1 的线性范围内变化的初始等级。无论什么时候本发明向广告人显示推荐，它将原始蜘蛛推荐和协作过滤的输出进行交织，因为来自这两种技术
15 的推荐常常是互补的。交织公式对万维网蜘蛛的推荐的加权随着广告人接收和拒绝更多的项而愈来愈少。

20 在典型应用中，新广告人将从其网站的 URL 着手并经过 3 - 5 次接受和拒绝项的迭代。只要其网站类似于现有广告人的那些网站，系统将快速地识别它们并给出高质量的推荐。推荐即使在没有单个广告人是最佳匹配的情况下也将是好的，因为加权和允许系统组合来自许多广告人的推荐。并且在没有广告人与新广告人相似的情况下，万维网蜘蛛仍通过直接在广告人网站上寻找搜索项来给出好的
25 推荐。与现有技术相比，本发明的实施例提供了对好搜索项的极好

覆盖并消除了不好的搜索项。

现在参考附图，图 10 是说明用于向根据位置付费的搜索引擎上的广告人推荐搜索项的方法的流程图。该方法可以在服务器或其它与根据位置付费的搜索引擎有关的数据处理装置上实现。该方法可以
5 可以被实现为能够在数据处理装置上和数据库的储存数据或其它存储部分一起操作的软件代码。广告人利用通过因特网链接到服务器上的诸如远程个人计算机的任何适当装置访问服务器来运行程序。以上连同图 1 给出了一个适当系统的例示实施例。方法从框 1000 开始。

在框 1002 中，系统提示广告人选择建立初始的接受搜索项列表的
10 输入方法。该列表可直接来自广告人的输入、来自广告人指定的统一资源定位器（URL）、或者来自广告人指定的先有广告人。在提示广告人他希望使用的方法后，程序进入图 10 所示的三条路径之一。

如果广告人选择直接指定初始的搜索项列表，在框 1004 从广告
15 人接收一些项，在一个例示实施例中，程序显示一个文字框，其中广告人可以输入用逗号隔开的初始项列表。如果广告人选择指定 URL 作为初始的搜索项列表的来源，则提示广告人输入网站 URL。系统在框 1008 运行蜘蛛算法来从该站点提取搜索项。这种蜘蛛算法的例示实施例将在下面结合图 11 - 13 进行描述。如果广告人选择指定先
20 有广告人作为初始的搜索项列表的来源，则在框 1010 从广告人接收现有广告人的标识信息。新广告人选择一个现有广告人，然后在框 1012 程序将广告人已投标的项列表设置为初始接受项列表。

该方法现在进入其主循环，其中包括框 1014、1016、1018、1020。
在每次迭代过程中，它运行协作过滤算法，在框 1016 显示推荐搜索
25 项的排序列表，并在框 1018 允许广告人接收和拒绝一些项。在例示实施例中，包括推荐搜索项的网页被发送给广告人，为广告人提供用户界面与系统进行交互。广告人通过点击各项后面适当的复选框来接受和拒绝各项。当他完成改变时，他点击一个按钮来将数据页

发送到服务器并重新运行协作过滤算法。广告人可以按他所希望的次数继续进行迭代，在框 1014 重复循环，直到他对他接受的项感到满意。然后，他在框 1020 点击最后按钮退出循环，并储存或打印其所选搜索项。可取的是，与广告人的通信是利用诸如 TCP/IP 的合适的数据传输协议通过因特网进行的。可以用其它数据通信信道取代。该方法在框 1022 结束。

图 11 是显示应用于执行蜘蛛算法的方法的流程图。该算法可以被在例如图 10 的框 1008 被调用。该方法从框 1100 开始。调用该过程来传递作为广告人网站的根的 URL。从这个 URL 开始，过程进入包括框 1102、1104、1106、1108 的循环。该过程利用广度优先蜘蛛算法下载页面。对于它下载的各页面，它在框 1104 扫描页面上的文字以找到在上个月已被用作搜索项的每个短语。在最佳实施例中，这种扫描通过建立识别正则表达式 $s_1|s_2|\dots|s_n$ 的有限态机来进行，其中各个 s_i 是有效搜索项。程序利用该状态机一次扫描页面上的一个字符，并在它找到各个搜索项时就发送它。因为状态机仅取决于当前的有效搜索项组，因此最佳实施例只在用户已搜索项的项的数据库改变时定期构建它。用于构建这种有限态机的算法很容易在文献中得到并在诸如 grep（查找正则表达式）的公共搜索应用程序中出现，如 Aho 和 Hopcraft 的“计算机算法设计”（ISBN 0201000296）中所述。它们对计算机系统领域普通技术专业人员是众所周知的。

每次蜘蛛在某页上找到新的项，它在框 1106 将该项添加到它在网站上已找到的项的列表中。它跟踪它在阵列 COUNT(T)中看到各项的次数。循环在框 1108 重复。当装置找到 1000 项时，如框 1102 的循环控制所示，下载和扫描过程结束。可以使用其它阈值或循环控制技术。图 11 的循环操作仅是例示性的。

下一步骤是滤除不好的项。这是在包括框 1110、1112、1114 的循环中进行的。不好是主观量度，并可存在许多实现可使用的可能量度。在最佳实施例中，质量量度取决于两个量：某项出现在万维

网上的文件中的频率，以及用户用它进行搜索的频率。在框 1112 评估质量量度。该方法通过查询返回包含该项的文件的数目的搜索引擎来找到万维网上某项的频率。它通过检查提供商数据库中的信息来找到用户用它进行搜索的频率。上述实施例中采用的质量量度是这两个数字之比的 \log ，如图 11 的框 1112 所示。

为实现高质量评定，项必须是在搜索人中流行的项，但不是在万维网文件中太常见，否则是无用的搜索项。因为质量量度改变迟缓，最好实施例仅定期对它们计算计算并对结果进行高速缓存。可以用其它质量量度取代。

一旦该方法算得它找到的 1000 项的质量，循环退回到框 1114，并且该方法在框 1116 删除或放弃所有在预定质量阈值以下的项。该阈值可以是变量，随时间而变，因为它取决于万维网上索引了多少页以及多少用户利用提供商的搜索引擎进行搜索。在最佳实施例中，程序通过检查好搜索项的临界上已知项的质量来自动校正阈值。它为这些项的平均质量设定阈值。准确的项列表取决于搜索引擎提供商并不限于特定实施例。

蜘蛛算法中的最后步骤是在框 1116 按照高于质量阈值的项在蜘蛛下载和扫描了的页中出现的频率对它们进行排序。这些计数被存储在 COUNT[1] 阵列中，排序列表是蜘蛛算法的输出。在一个典型实施例中，质量过滤删除了大约 80% 的项，同时算法大约返回 200 项。蜘蛛方法在框 1118 结束。

图 12 是给出一种用于执行协作过滤算法的方法的流程图。该方法从框 1200 开始。在框 1202 和框 1204，对新广告人和现有广告人的等级值初始化。用于执行这些操作的实施例在以下结合图 13 和 14 进行描述。在框 1206 中，控制进入包括框 1206、1208、1210 的循环。在此循环中，该方法通过图 11 的协作过滤算法处理所选搜索项并在框 1208 计算各项的新广告人的估计等级。这种等级预测方法的一个实施例在以下结合图 18-20 进行描述。在处理了所有搜索项后，

循环在框 1210 退出。在算法结束时，在框 1212 按各项的预测等级对它们进行排序。该方法返回最后列表，作为其推荐的分级列表，然后在框 1214 结束。

5 在此算法和以下算法中，存在许多实现可包括的效率优化。例如，它可以仅返回前 100 个搜索项而不是整个列表，或者它可以高速缓存计算结果以避免重复工作。所有这些优化对计算系统设计领域的普通技术专业人员都是显而易见的，并且这里所示的实施例不依赖于实现使用的特定优化。

10 图 13 是说明用于初始化现有广告人等级值的最佳算法的流程图。该算法是对每个广告人/搜索项对的循环。对于每一对，如果广告人以对某项投标，程序设置等级为 5，否则为 UNKNOWN。等级被存储在 $V[A][T]$ 阵列中，以便使程序的其它部分能够对它们进行访问。

15 该方法从框 1300 开始。利用广告人变量 A 在框 1302 进入广告人处理循环。利用项变量 T 在框 1304 进入项处理循环。在框 1306，该方法确定与广告人变量 A 有关的广告人是否已对与变量 T 有关的项投标。如果没有，在框 1308，等级 $V[A][T]$ 被设置为等级值阵列中的 UNKNOWN 值。如果广告人已对该项投标，则在框 1310 将阵列条目 $V[A][T]$ 设置为 5，5 是任意挑选的值。

20 在框 1312，项变量加 1 或者改变为选择下一项。控制仍在包括框 1304、1306、1308、1310、1312 的循环中，直到处理了与变量 A 有关的变量的所有搜索项。接着，在框 1314，广告人变量 A 加 1 或者改变到新选广告人，然后循环为新选广告人继续处理搜索项。在所有广告人被处理了所有搜索项之后，方法在框 1316 结束。

25 图 14 是给出初始化新广告人的等级值的最佳算法的流程图。该算法是对每个搜索项的循环。对于每一项，如果新广告人已接受某项，则程序设置等级为 5，而如果他拒绝某些则设置为 1。如果他两样都没做，而蜘蛛推荐了该项，则程序将等级设置为蜘蛛的估计等

级。如果这三种情况都没有发生，则程序将等级设置为 2。

该方法从框 1400 开始，在框 1402，利用项变量 T 作为循环变量进入循环。在框 1404 确定广告人是否已接受与广告人的搜索项的变量 T 有关的项。如果是，在框 1406，广告人和项的等级 V[A][T] 被设置为等级阵列中的值 5。控制进行到框 1418 以选择循环变量 T 的下一项。如果广告人没有接受当前搜索项 T，在框 1408 确定广告人是否拒绝了它。如果是，在框 1410，将广告人和项的等级 V[A][T] 设置为值 1 同时控制进行到框 1418，以对循环变量加 1。如果广告人没有拒绝项 T，在框 1412 确定蜘蛛算法是否推荐了与变量 T 有关的项。如果是，在框 1414 将广告人和项的等级 V[A][T] 的值设置为等于蜘蛛算法建立的等级。否则，广告人和项的等级 V[A][T] 被设置为值 2。然后，控制进行到框 1418，以对循环变量加 1。在所有项都处理完毕后，该方法在框 1420 结束。

图 15 是说明用于计算两个广告人之间的皮尔逊相关的算法的流程图。该算法是在整个搜索项上的循环。对于各项，程序累计各个值，以允许它计算皮尔逊相关公式：

$$\rho_a = \frac{\sum_t (r_{n,t} - \bar{r}_n)(r_{a,t} - \bar{r}_a)}{\sigma_n \sigma_a}$$

X 变量累积分子的值，同时 Y 变量累积分母的值。在程序在所有搜索项循环之后，它利用流程图中的最终表达式计算相关性。

该方法从框 1500 开始。在框 1502 对变量 X、Y1 和 Y2 进行初始化。在框 1504 进入循环，以处理搜索项列表中的各个搜索项。在框 1506 中，利用等级算法计算变量 X1 和 X2。等级算法计算广告人分配给搜索项的的等级。适当等级算法的一个实施例在以下结合图 16 进行描述。在框 1508 中将 X1 和 X2 的值与 X 的上一值进行组合，以产生当前 X 值。在框 1510，利用算得的 X1 和 X2 更新 Y1 和 Y2 的值。在框 1512，控制循环回到框 1504，直到处理完所有搜索项。

然后，如框 1514 所示计算皮尔逊相关。方法在框 1516 结束并返回皮尔逊相关的值。

图 16 是显示用于计算广告人分配给某项的等级的算法的流程图。如果记录在 $V[A][T]$ 阵列中的等级不是 UNKNOWN，则算法简单地将该等级返回，否则，它返回广告人的平均等级。

该方法从框 1600 开始。传递两个变量：广告人变量和项变量。在框 1602 确定与广告人和项有关的等级是否为未知的。如果不是，则在框 1604 将等级设置为等于等级阵列中的等级值。若干变量是未知的，则在框 1606 将等级设置为等于广告人的平均等级。一种用于计算广告人的平均等级的方法在下面结合图 17 进行描述。返回等级同时该方法在框 1608 结束。

图 17 是显示用于计算广告人的平均等级的算法的一个实施例的流程图。该算法是在每个搜索项上的循环。对于具有已知等级的各个搜索项，程序将该等级添加到和 S 中，同时计数器 N 加 1。在循环结束时，平均等级是简单的比率 S/N 。

该方法从框 1700 开始。在框 1702 对和变量 S 和计数变量 N 进行初始化。在框 1704 进入循环，根据循环变量选择广告人的列表的搜索项。在框 1706，确定储存在等级阵列中的搜索项等级是否具有 UNKNOWN 值。如果不是，则在框 1708 将等级 $V[A][T]$ 的值加到和变量 S 中同时计数变量 N 加 1。控制进行到框 1710，在那里重复该循环直到处理了搜索项的广告人的列表中所有搜索项。在框 1712 计算平均等级作为 S 与 N 之比。在框 1714，该方法结束同时返回平均等级。

图 18 是显示万维网蜘蛛和协作过滤的组合推荐算法的一个实施例的流程图。项的组合等级是蜘蛛等级和协作过滤等级的加权和。首先，当广告人尚未接受或拒绝任何项时，算法对协作过滤的等级加权两倍于它对蜘蛛推荐的加权。随着接受和拒绝项的增加，蜘蛛等级的加权按比例下降。存在许多可能的生成各个等级的组合等级

的公式，而本发明不限于这些公式中的任何一个。

在图 18 的实施例中，方法从框 1800 开始。在框 1802 中，变量 N 被设置为等于广告人接受的推荐搜索项的数量，同时变量 M 被设置为等于广告人拒绝的推荐搜索项的数量。在框 1804，调用两个例程来从蜘蛛和来自协作过滤的预测等级计算预测等级。这些例程的例示实施例在以下分别结合图 19 和 20 进行描述。在框 1806，将预测组合并且该方法在框 1808 结束同时返回结果。

图 19 是用于计算蜘蛛等级的算法的一个实施例的流程图。如果蜘蛛没有找到某个项，或如果该项没有通过其质量过滤，则搜索将为其分配等级 2。其余项得到 4.9 到 2.1 线性刻度范围内的等级。蜘蛛推荐的第一项得到等级 4.9，而它推荐的最后一项得到等级 2.1。有许多其它可能的公式用于根据蜘蛛的分级推荐生成等级，并且本发明不限于这些公式中的任何一个。

方法从框 1900 开始。在框 1902 确定蜘蛛找到的项是否传递到了项变量 T 中的方法。如果是，则在框 1904 将变量 N 设置为等于蜘蛛找到的项数，而变量 M 被设置为等于项 T 在蜘蛛返回的推荐排序列表中的位置。

在框 1906，根据所述公式计算来自蜘蛛的预测等级。在框 1908，如果蜘蛛没有找到项 T，则来自蜘蛛的预测等级被设置为等于 2。该方法在框 1908 结束同时返回来自蜘蛛的预测等级。

图 20 是给出用于计算协作过滤的等级的算法的一个实施例的流程图。该算法是在每个广告人上的循环。对于每个广告人，程序累计各个值，以允许它根据以下公式计算等级：

$$e_t = \bar{r}_n + \frac{\sum_a (r_{a,t} - \bar{r}_a) \rho_a}{\sum_a \rho_a}$$

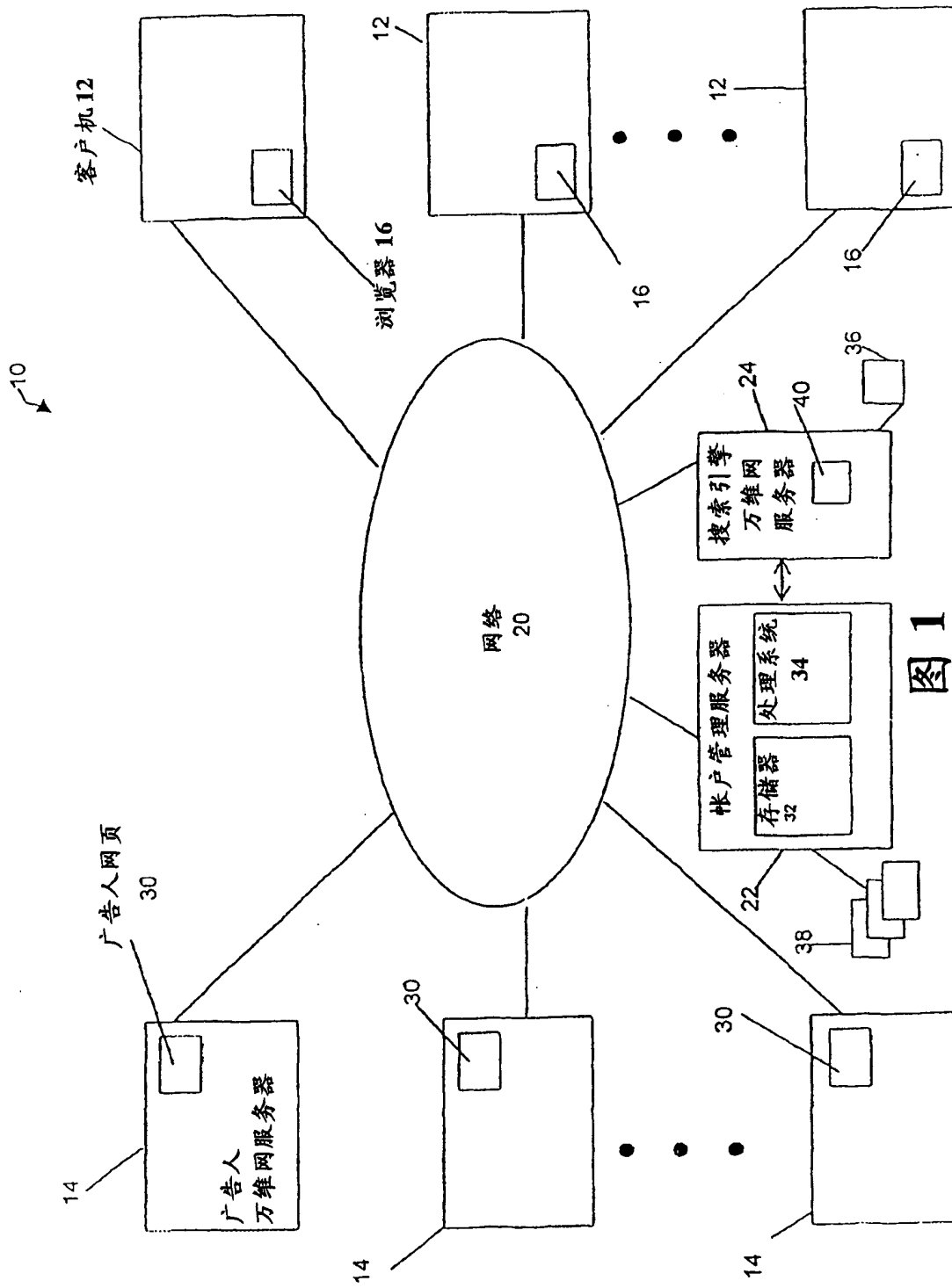
变量 X 累积分子的值，同时变量 Y 累积分母的值。在最后步骤中，该算法利用流程图中给出的表达式计算最后等级。这个最后等

级可以在等级 0 到 5 以外，但它可能在这个标度上正确地解释。

该方法从框 2000 开始。在框 2002 对变量 X 和 Y 进行初始化。在框 2004 进入循环，经由该循环处理一个广告人的一次迭代。在框 2006 中，利用 W 和 XA 的值更新 X 和 Y 的值。在框 2010，控制回到在框 2004 的循环开始，以处理下一广告人。在处理了所有广告人后，利用框 2012 中的公式和以上结合图 17 的平均等级算法计算来自协作过滤的预测。方法在框 2014 结束并返回来自协作过滤的预测。

根据以上所述，可以看出，本发明的实施例提供了一种方法和设备，用于向根据位置付费的搜索系统上的广告人推荐搜索项。该方法和装置根据广告人位置的内容并通过将广告人与其它相似的广告人进行比较并推荐他们已经选择的搜索项来提供搜索项推荐。以这种方式，系统推荐好的搜索项，或者具有与广告人位置或位置内容有关的项，同时避免没有这种关系的不好的项。该系统与广告人进行交互，允许他决定在何时搜索项的集合满足其要求。但是，对搜索项进行识别和分级的过程是自动的，并且根据广告人位置的实际页面并基于与其它广告人的比较。

尽管给出并描述了本发明的特定实施例，但可以进行修改。因此，所附权利要求书旨在涵盖这类改变和修改，权利要求书是本发明的真正精神和范围。



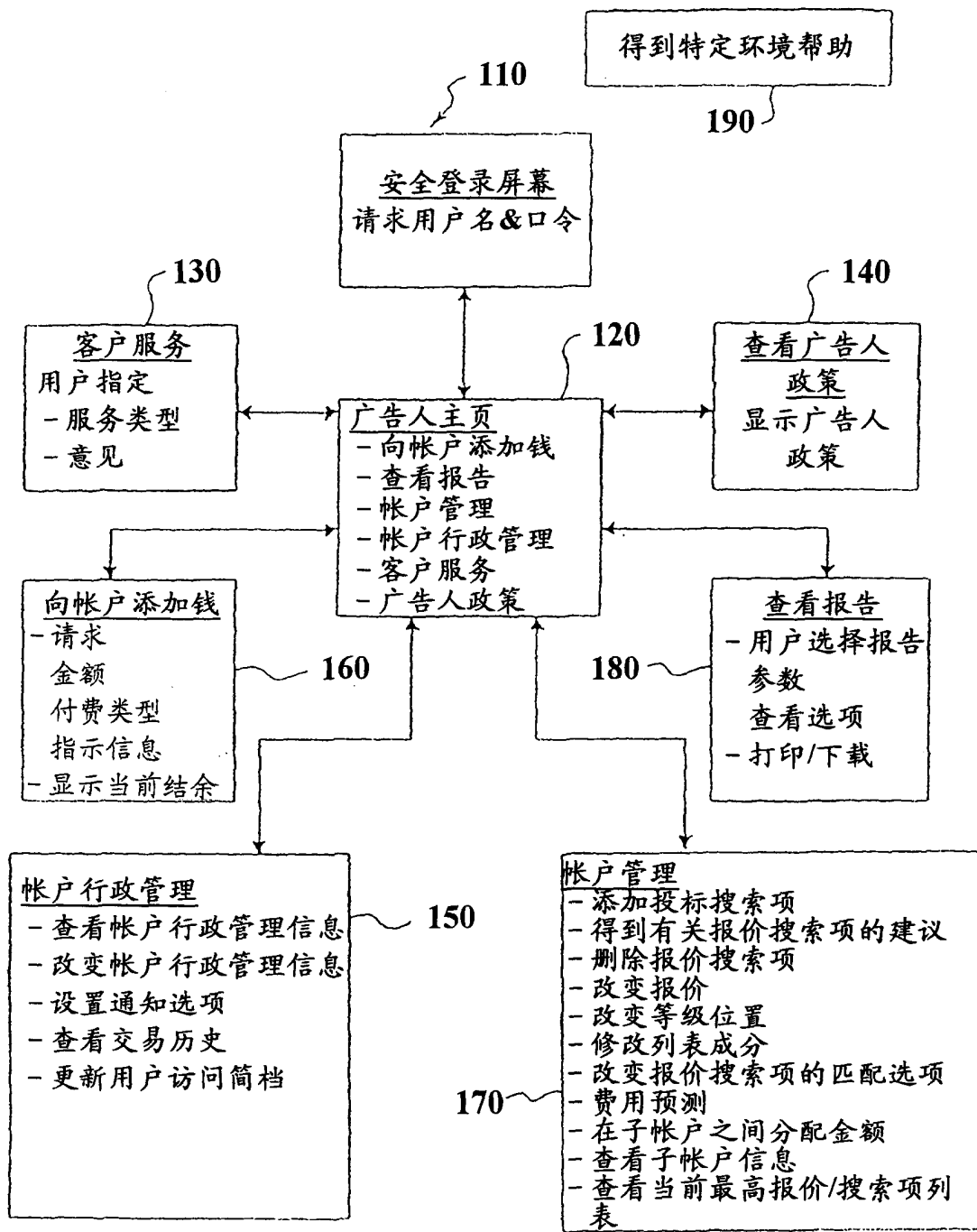


图 2

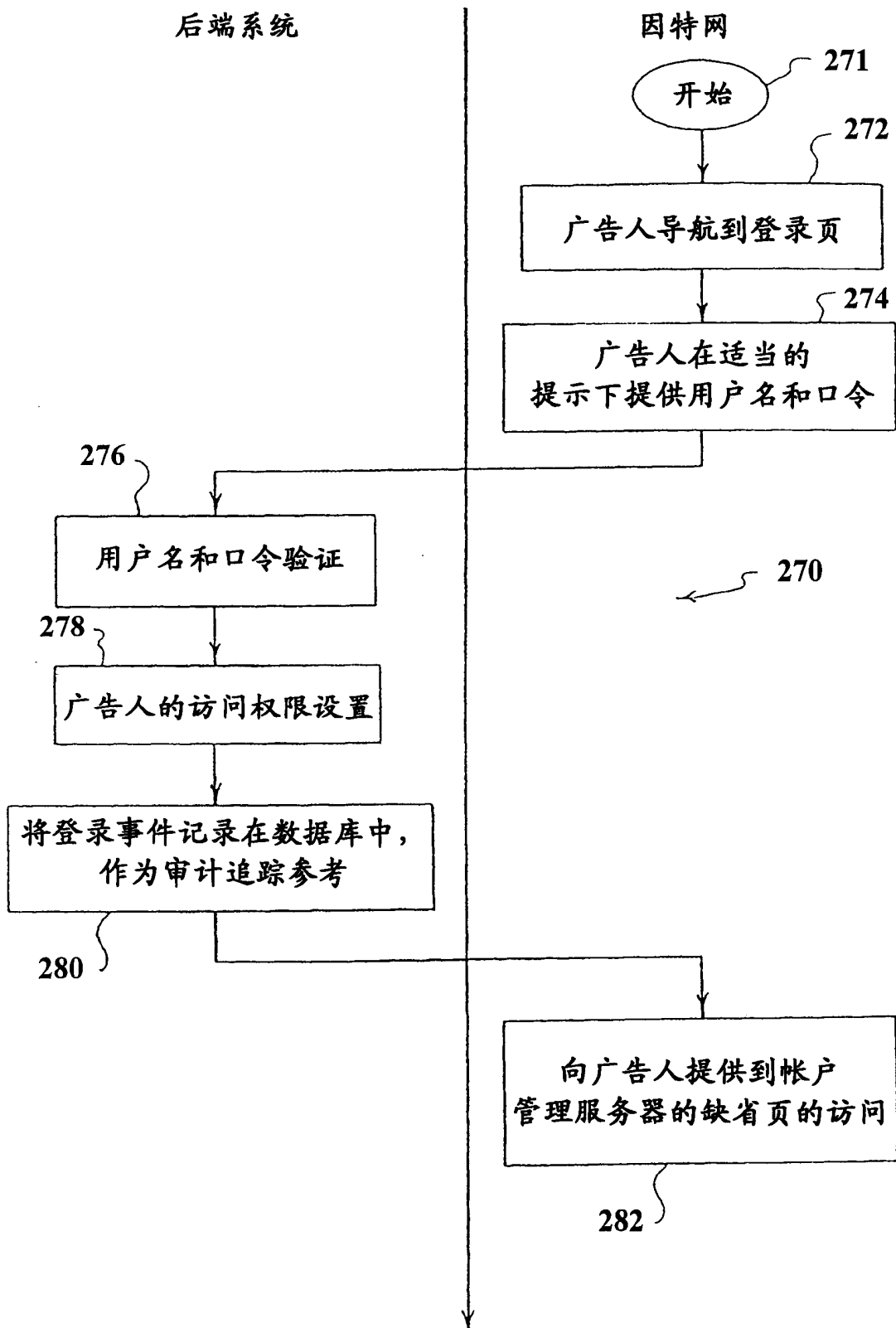


图 3

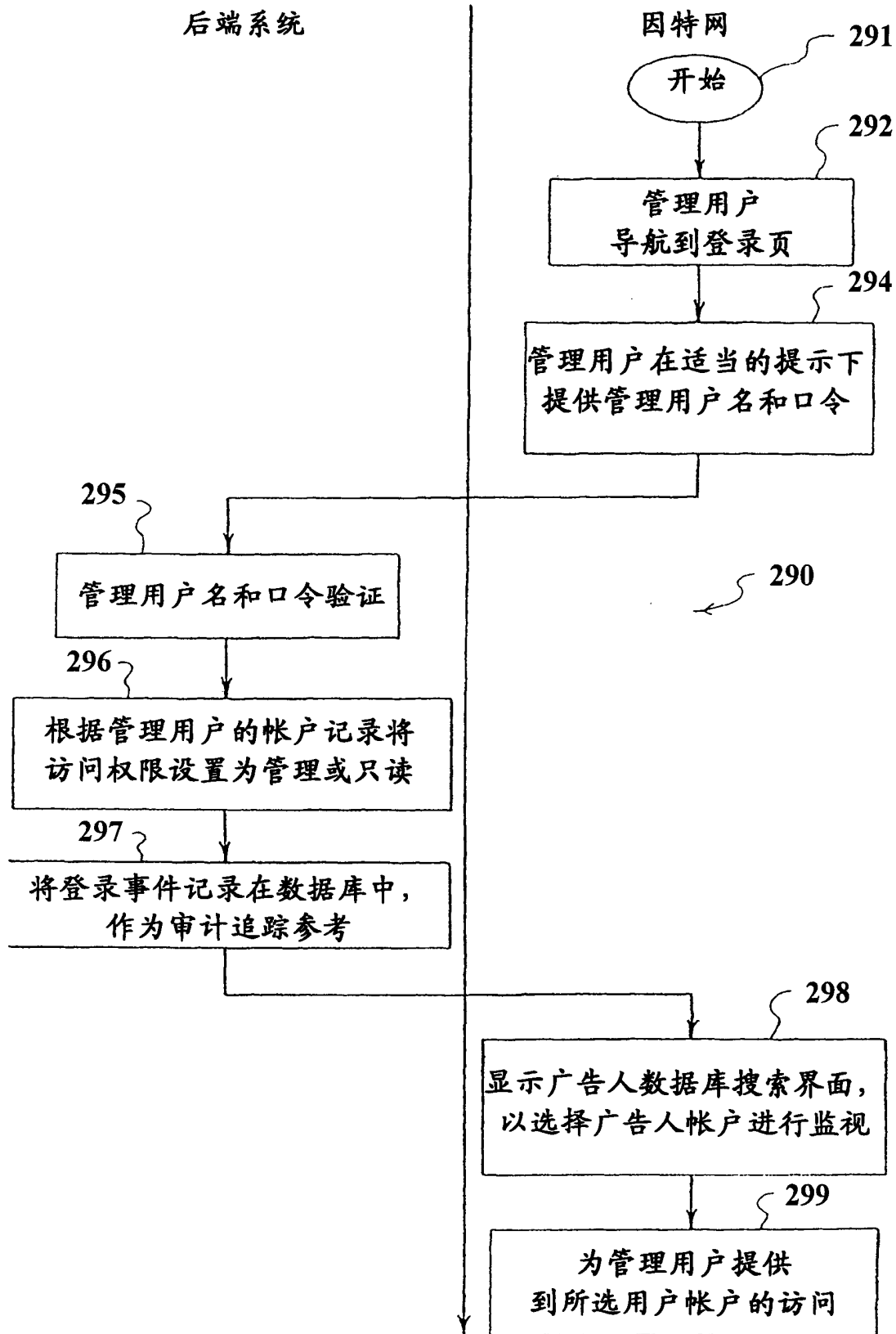


图 4

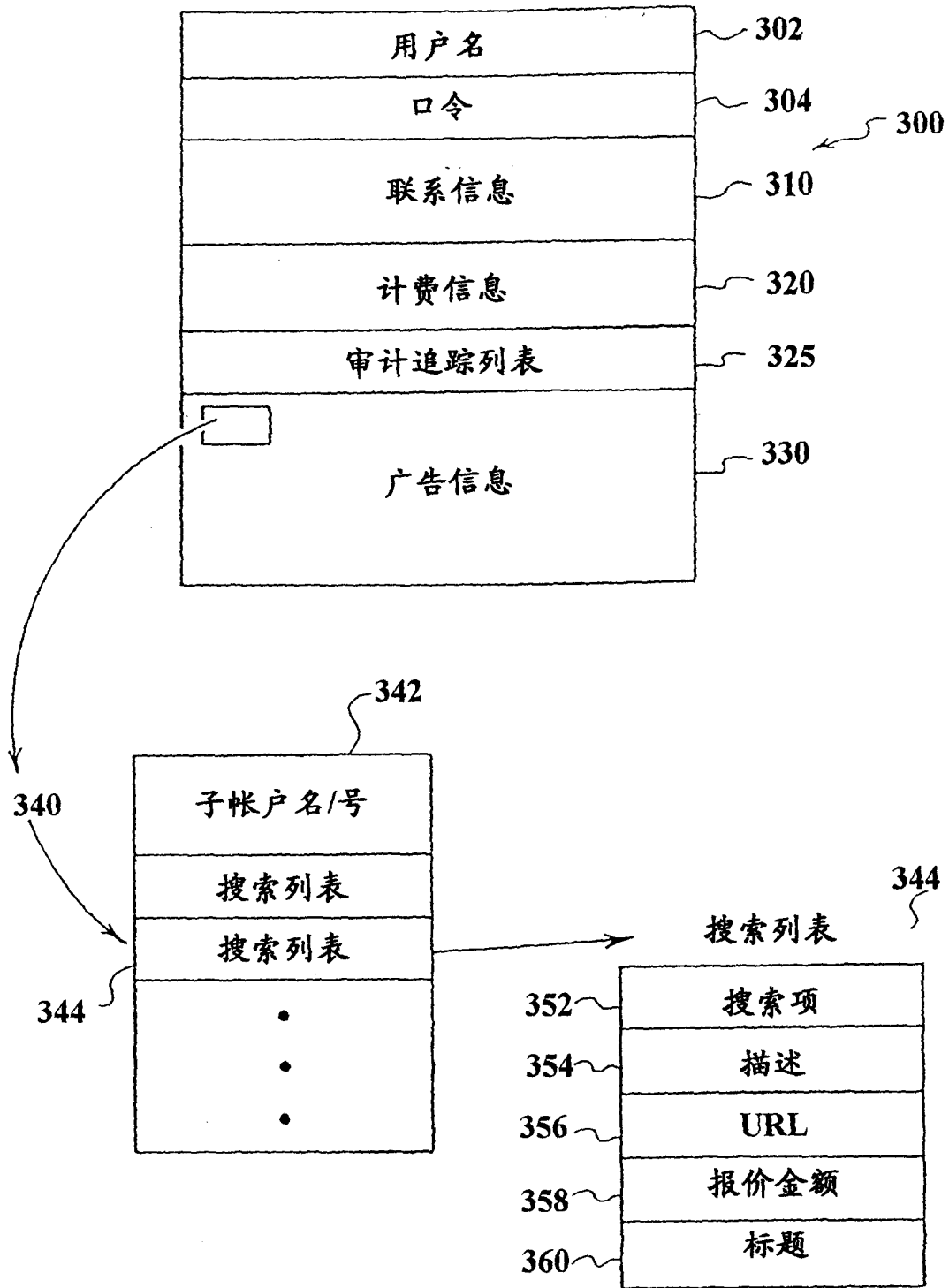


图 5

图 6

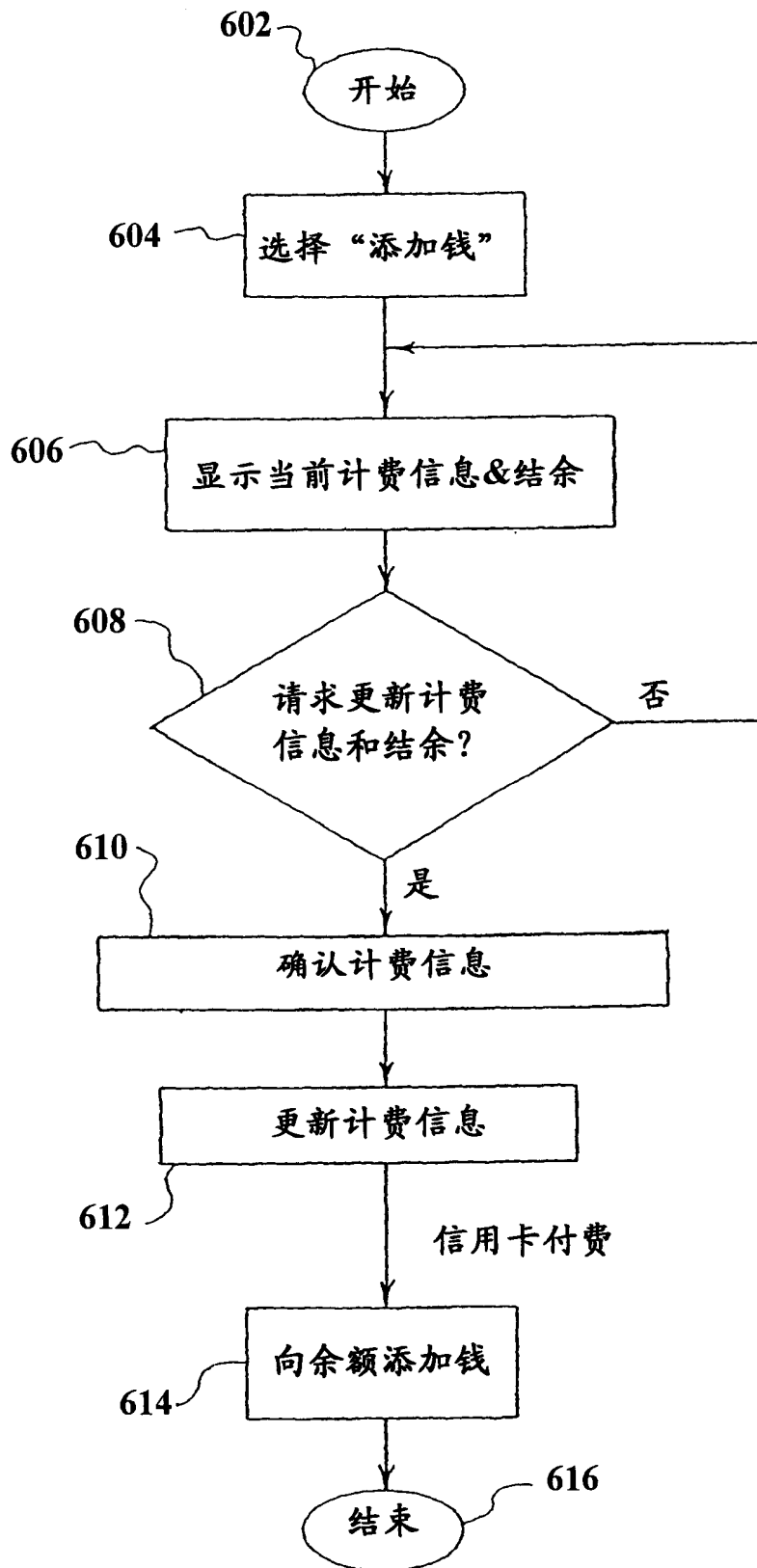


图 7

GO TO.com ZIP DRIVES Find It!

Type what you're looking for and click Find It!

Try Go To Remote! **730** Add Go To to Favorites

- 720** 1. We Buy And Sell Zip Drives **720**

Buy Sell or Trade used computers, memory, peripherals, hardware and software. We offer the best brand names, competitive prices, large inventory, and reliability. We can ship anywhere.
<http://www.micro-tradingpost.com/> (Cost to advertiser: \$0.08) **710a**
- 740** 2. MacnFind.com - Click Here! **750a**

Online guide to apple computer retailers. Find IMacs, power macintosh G3s, powerbooks, mac software and much more!
<http://www.mactofind.com/> (Cost to advertiser: \$0.07) **750b** **760b** **710b**
- 760c** 3. Custom Computer Systems, Inc (CCS) **750c**

CCS, a wholesale distributor of computer systems, parts & accessories. Online shopping and same day shipping! 800.379.1227
<http://www.wccsi.com/> (Cost to advertiser: \$0.06) **710c**
- 760d** 4. Best Price on Zip Drives **750d**

ATMAN - USA.COM Computer on-line shopping. Huge selection, competitive price, reliable service. Major credit cards are welcome.
<http://www.atman-usa.com/> (Cost to advertiser: \$0.02) **710d**
- 760e** 5. Computer Beat...One - Step Web Directory **750e**

A One-Stop Web Directory featuring top computer, PC, Mac web sites on the Internet. Spotlights Computer News, Help Desks, HTML, Internet and Computer Shopping!
<http://www.search-beat.com/> (Cost to advertiser:\$0.01) **710e**
- 760f** 6. Macnet Online Computer Store **750f**

Computer retailer of hardware software peripherals consumable and accessories for PC/Macintosh users. See Powermac, Imac, DVD, digital camera, monitors, printers, scanners, networking, PDA, notebook, and more.
<http://www.applemacnet.com/> (Cost to advertiser: \$0.01) **710f**
- 760g** 7. hard drives for sale, western digital, maxtor, seagate, samsung, ibm, lomega zip drives. Ide, scsi, PSA Inc reseller **710g**

Offering computer systems, memory, cpu sales, motherboards, computer upgrades at volume and dealer pricing. A distributor, reseller, broker for all pc needs from corporations to individual personal computer users. PSA Inc. - computer sales and upgrades.
<http://www.psa.com/>
- 760f** 8. Tape Drives, Zip Drives, Floppy Drives **710h**

Wholesale distributor in Cleveland Ohio of computer hardware, computer systems, computer components, peripherals, cases, keyboards, mice, tape backup and floppy drives. Free technical support.
<http://www.bitsandbyteinc.com/>

[http://www.../;\\$scsionk\\$MEPODQAA40YLQFIEE1APUQ?type=home&Keywords=zip+drive](http://www.../;$scsionk$MEPODQAA40YLQFIEE1APUQ?type=home&Keywords=zip+drive)

图 8

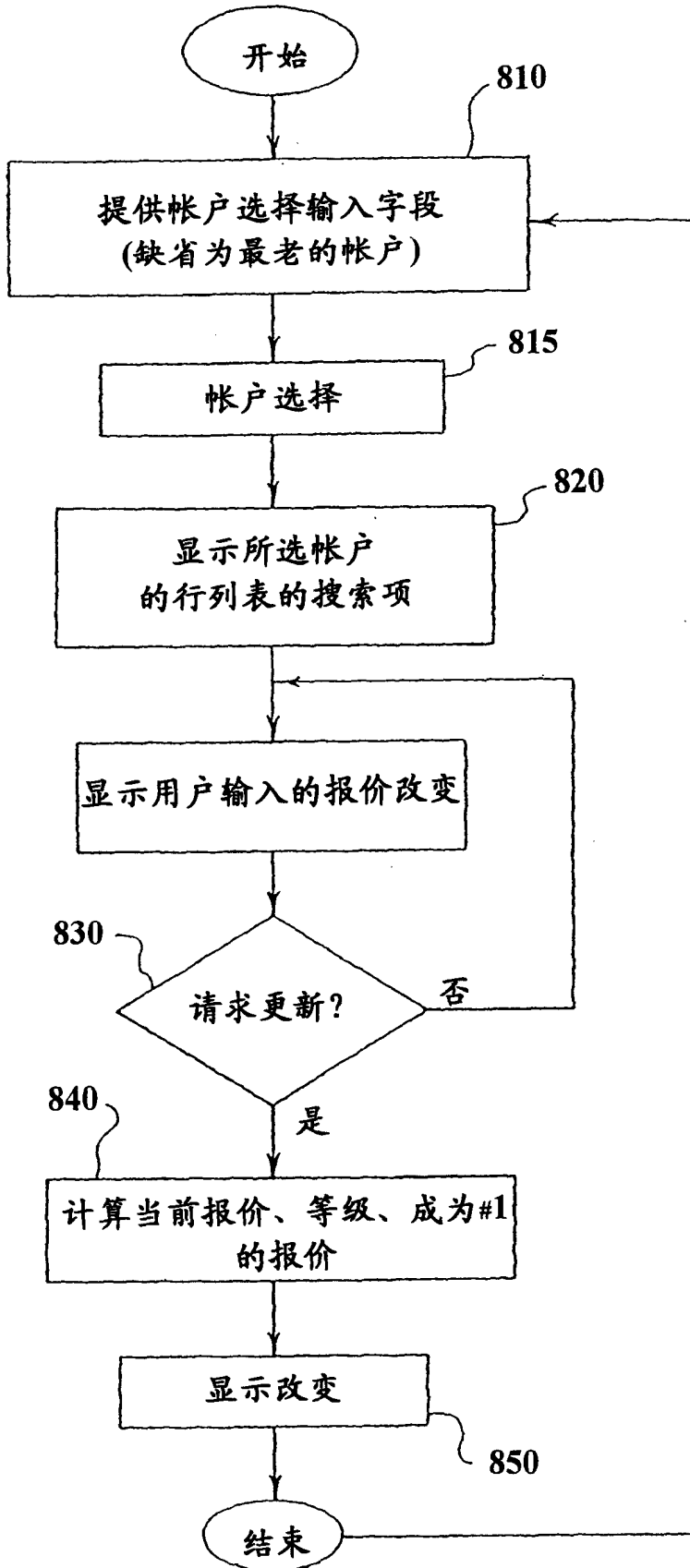


图 9

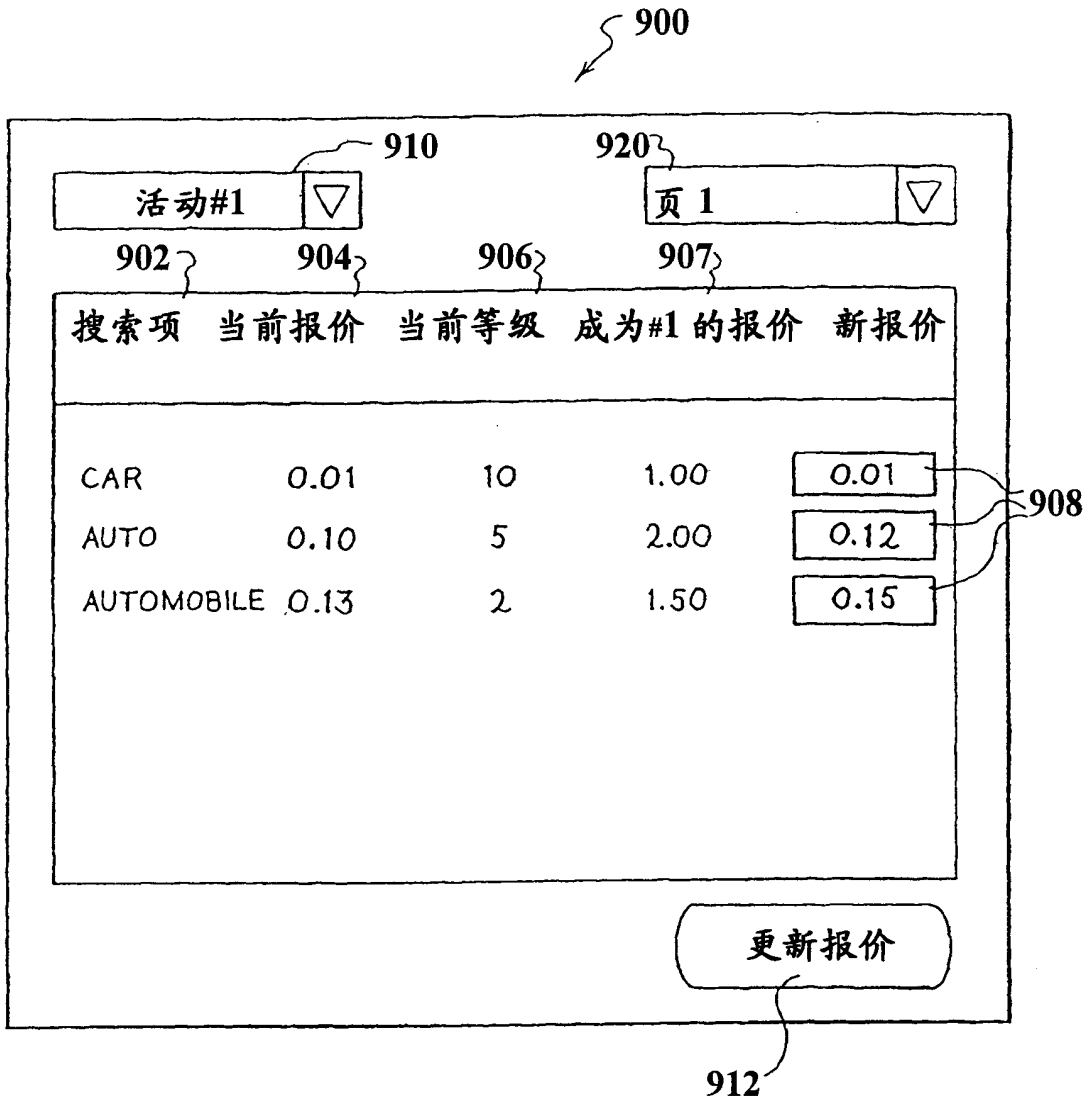


图 10

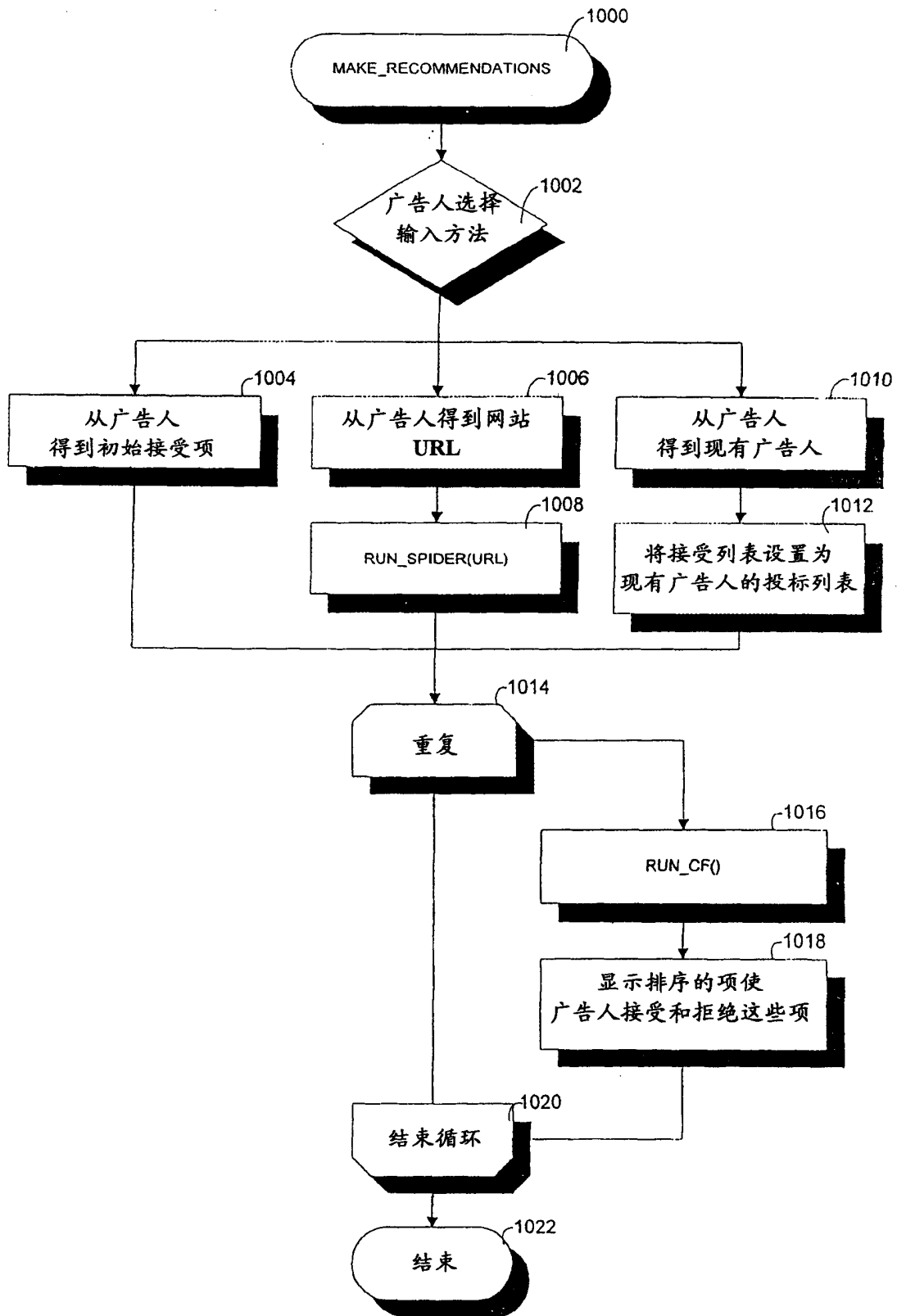


图 11

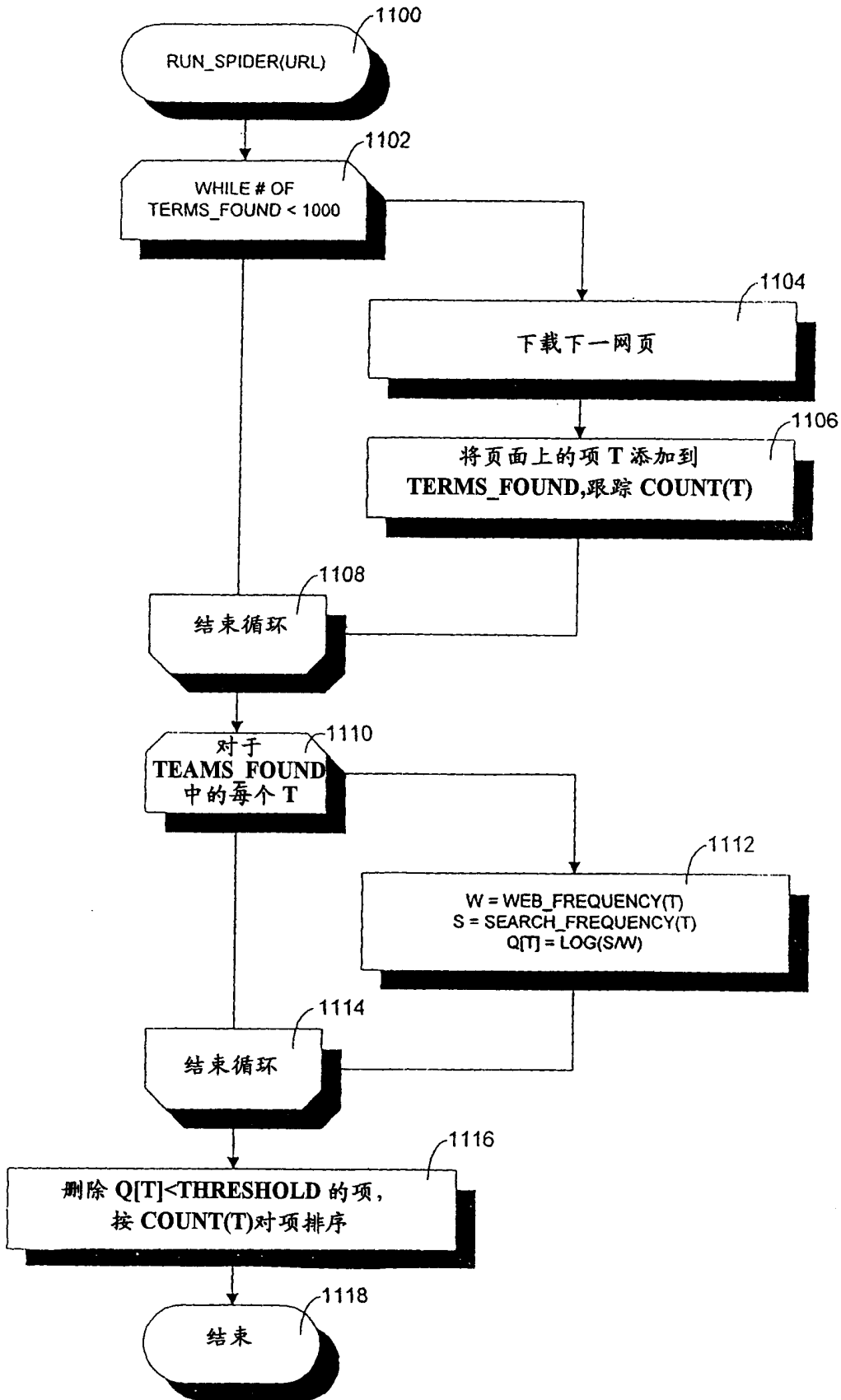


图 12

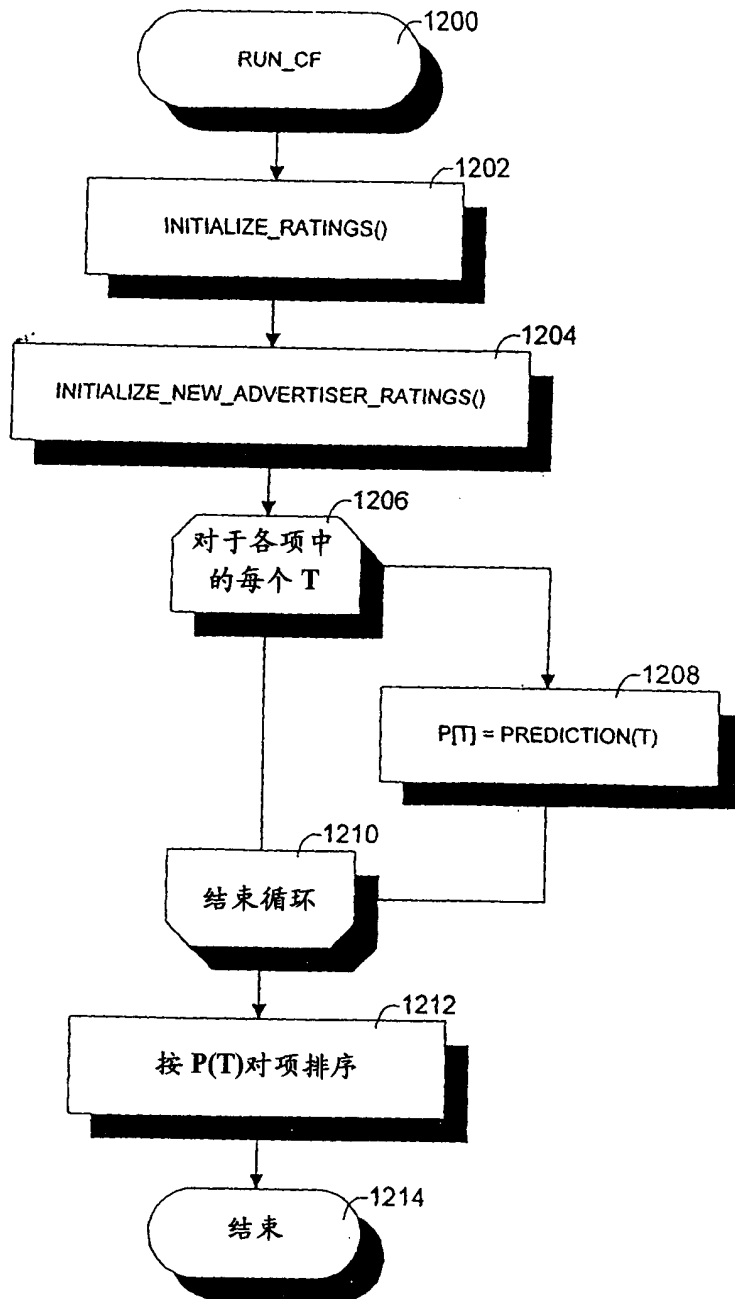


图 13

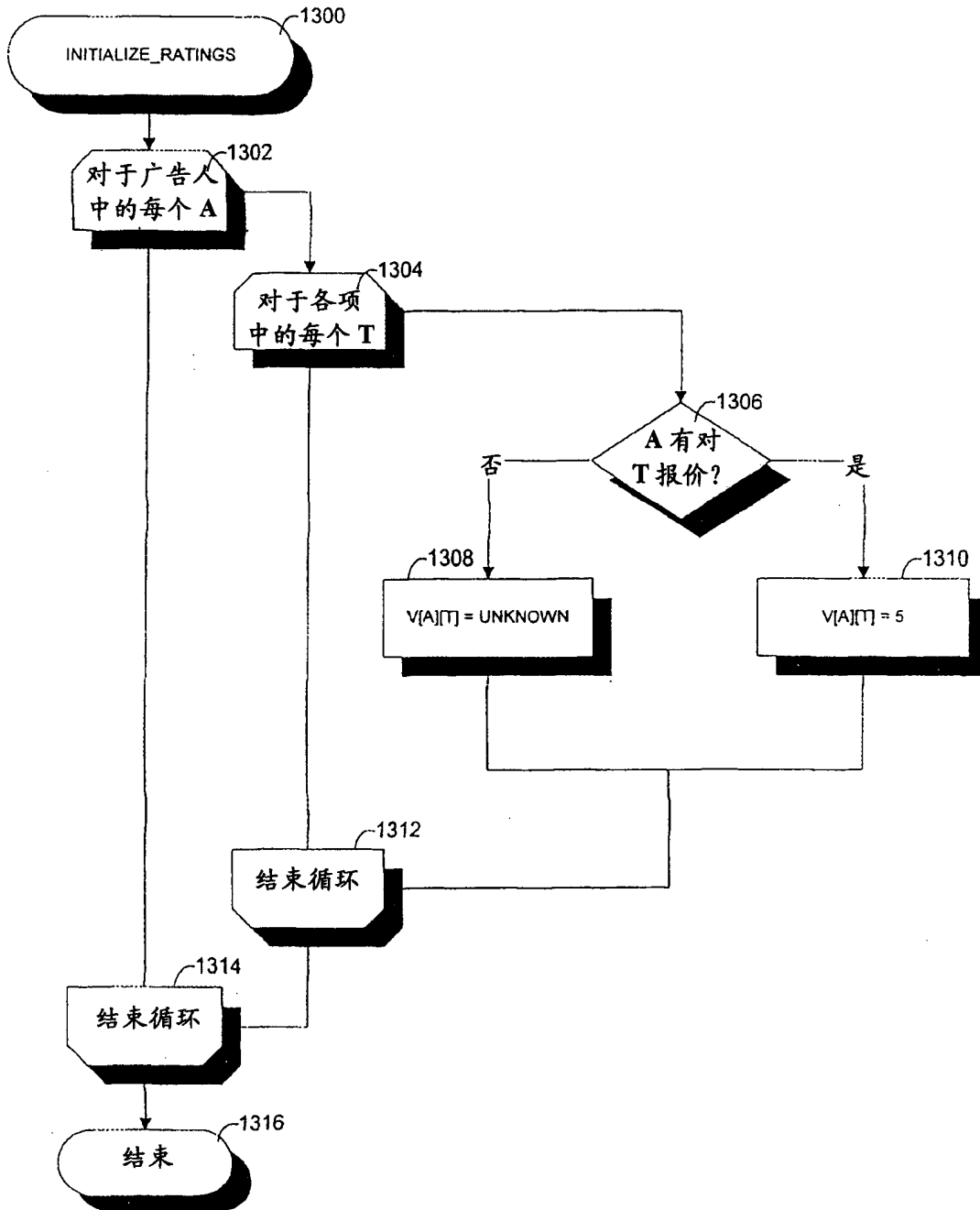


图 14

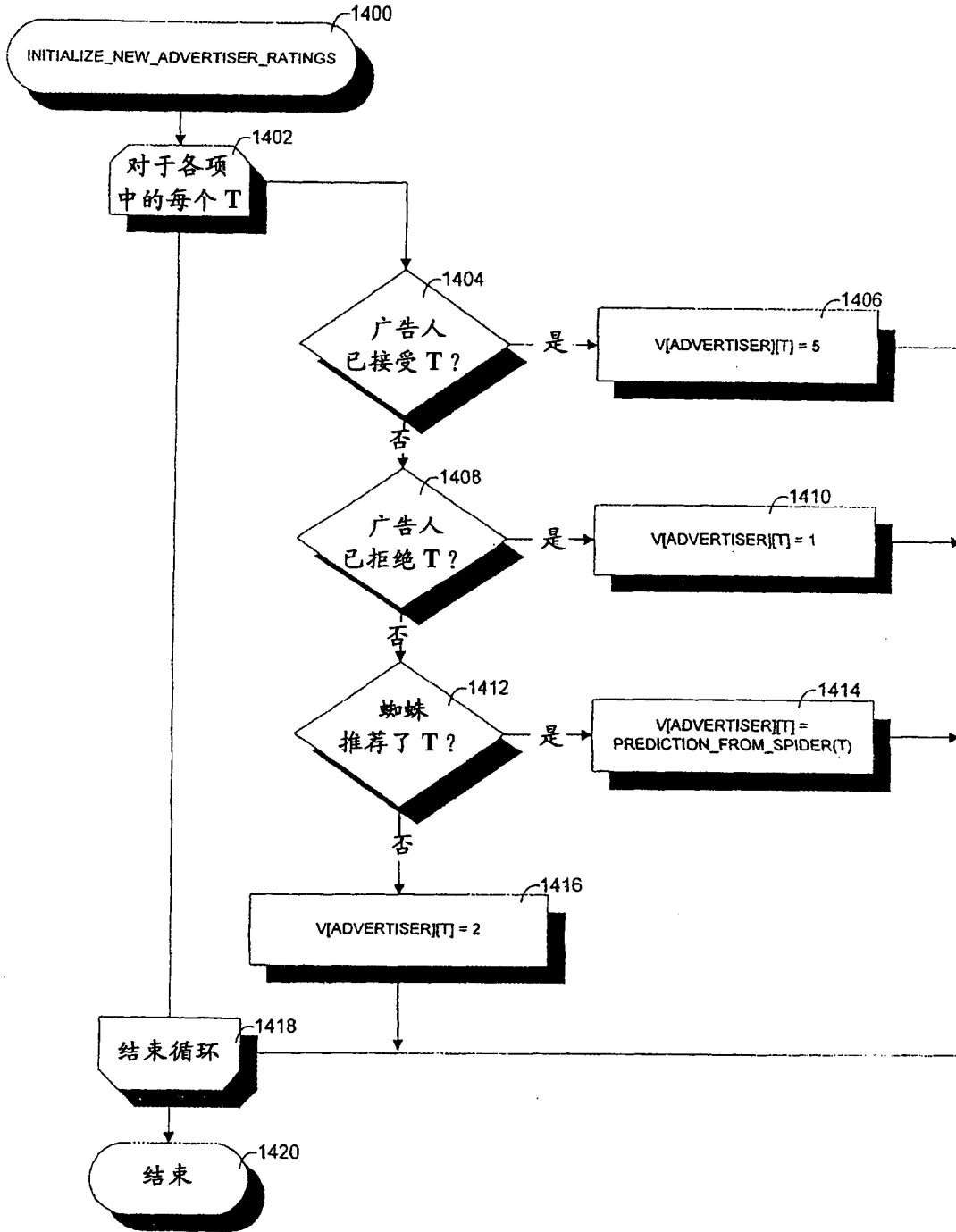


图 15

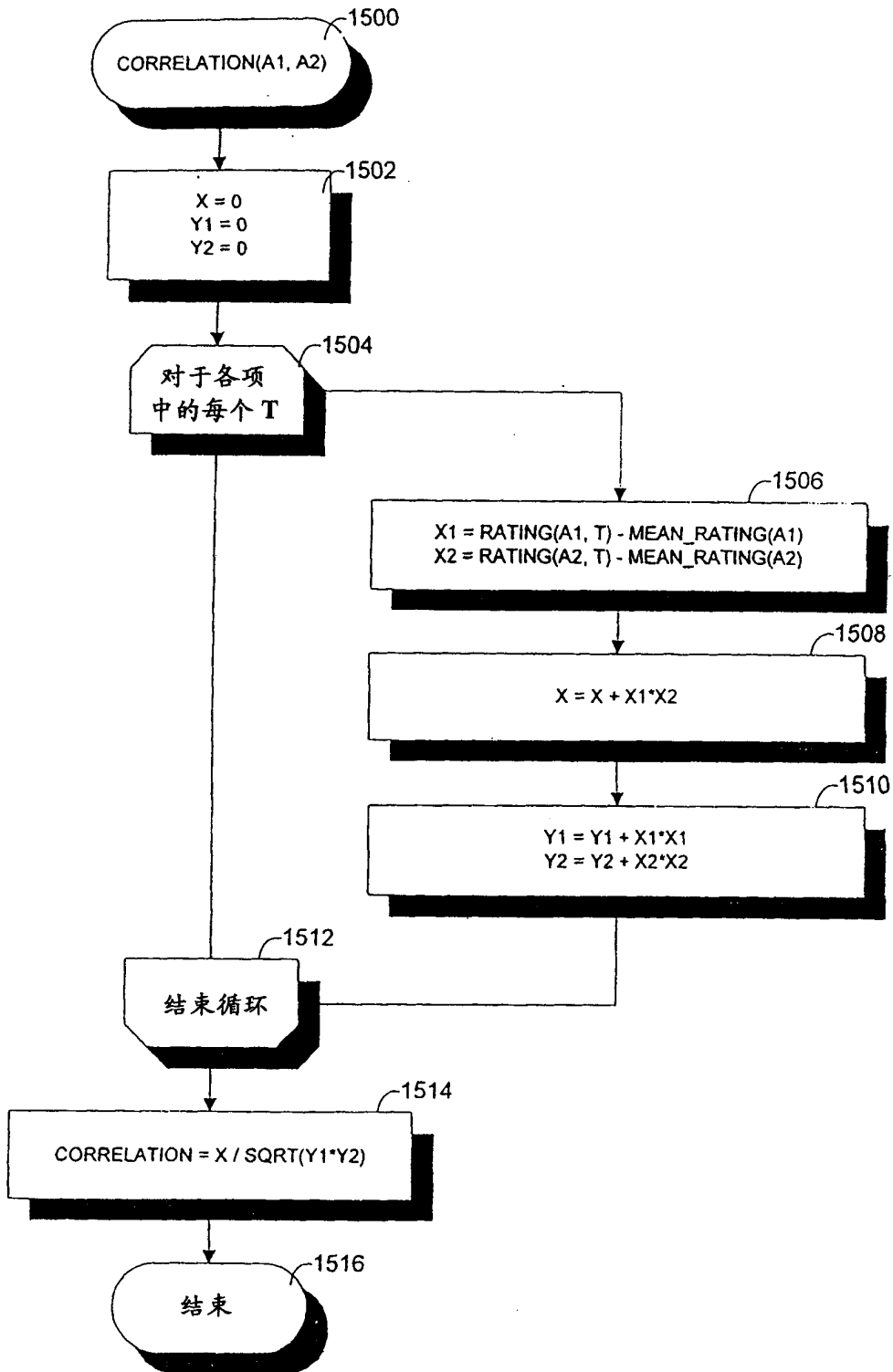


图 16

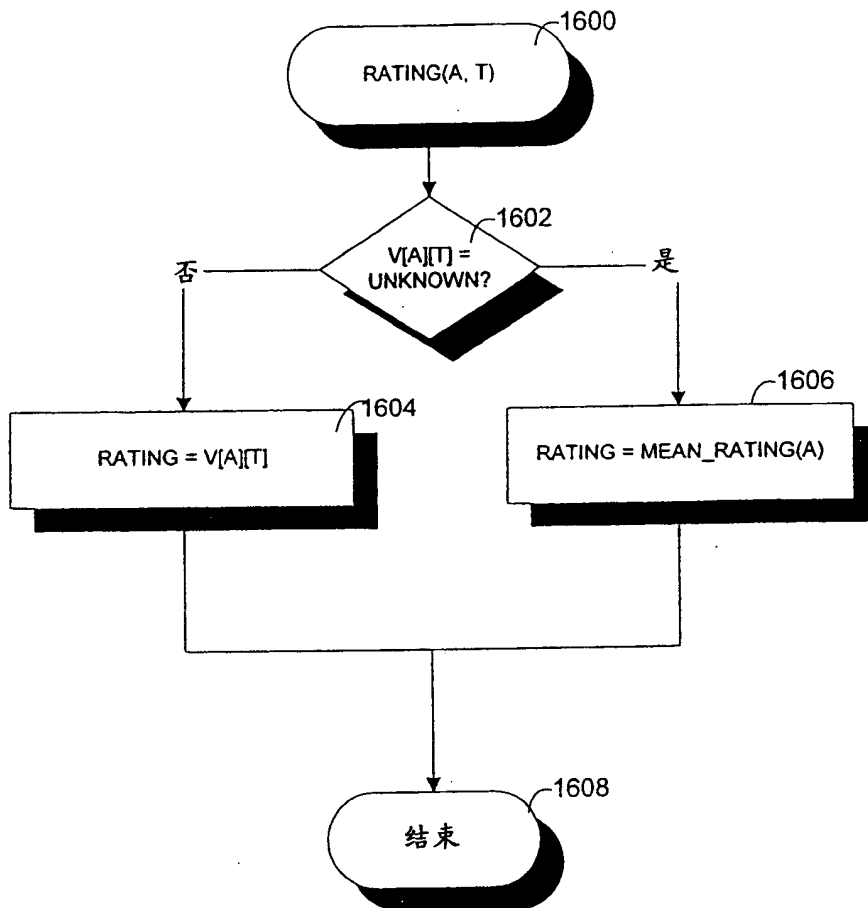


图 17

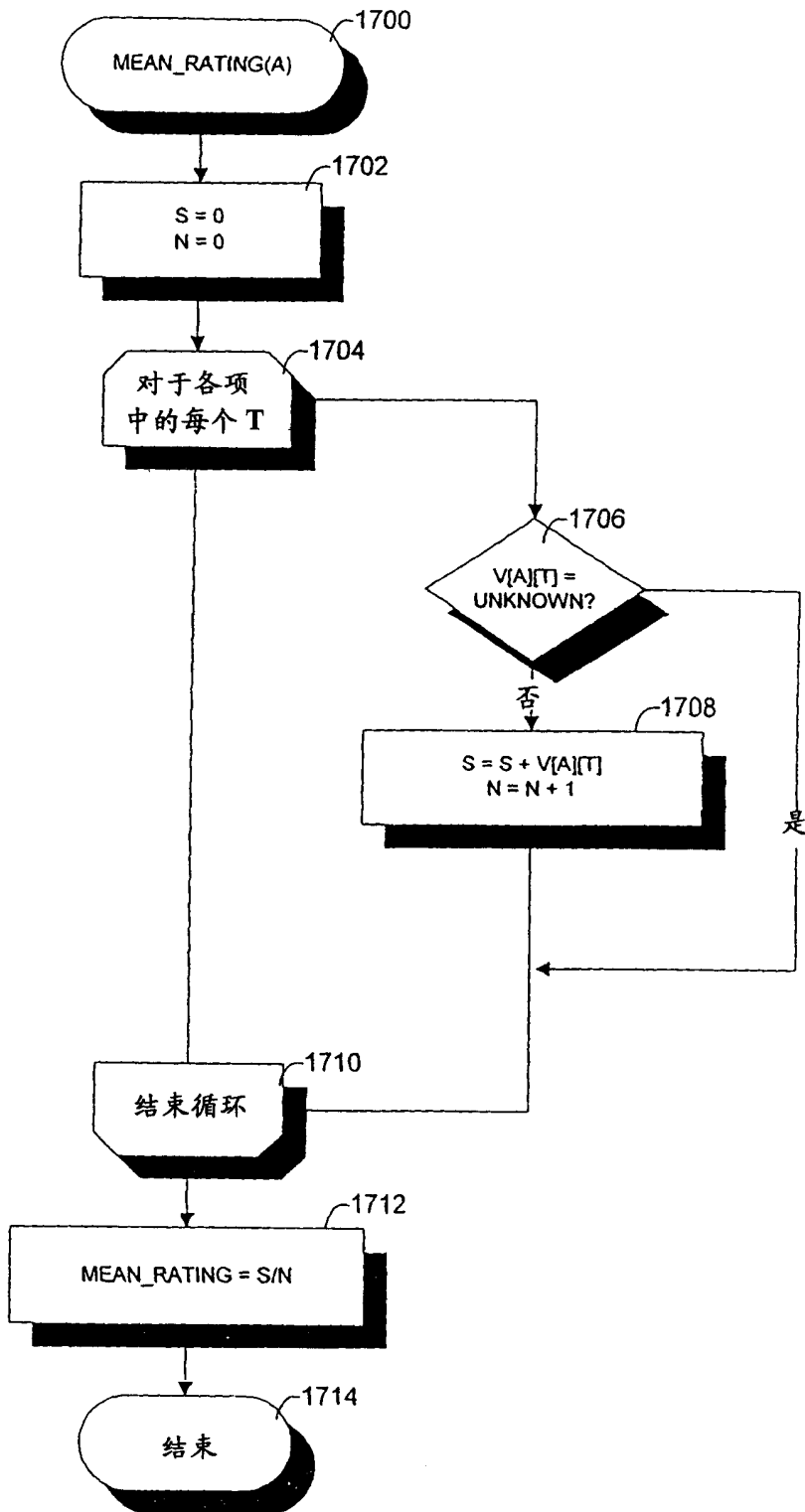


图 18

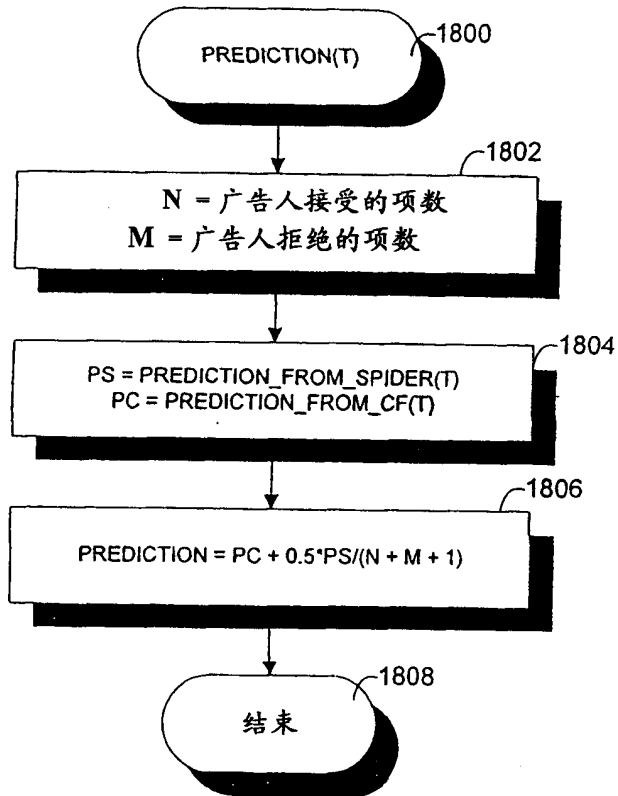


图 19

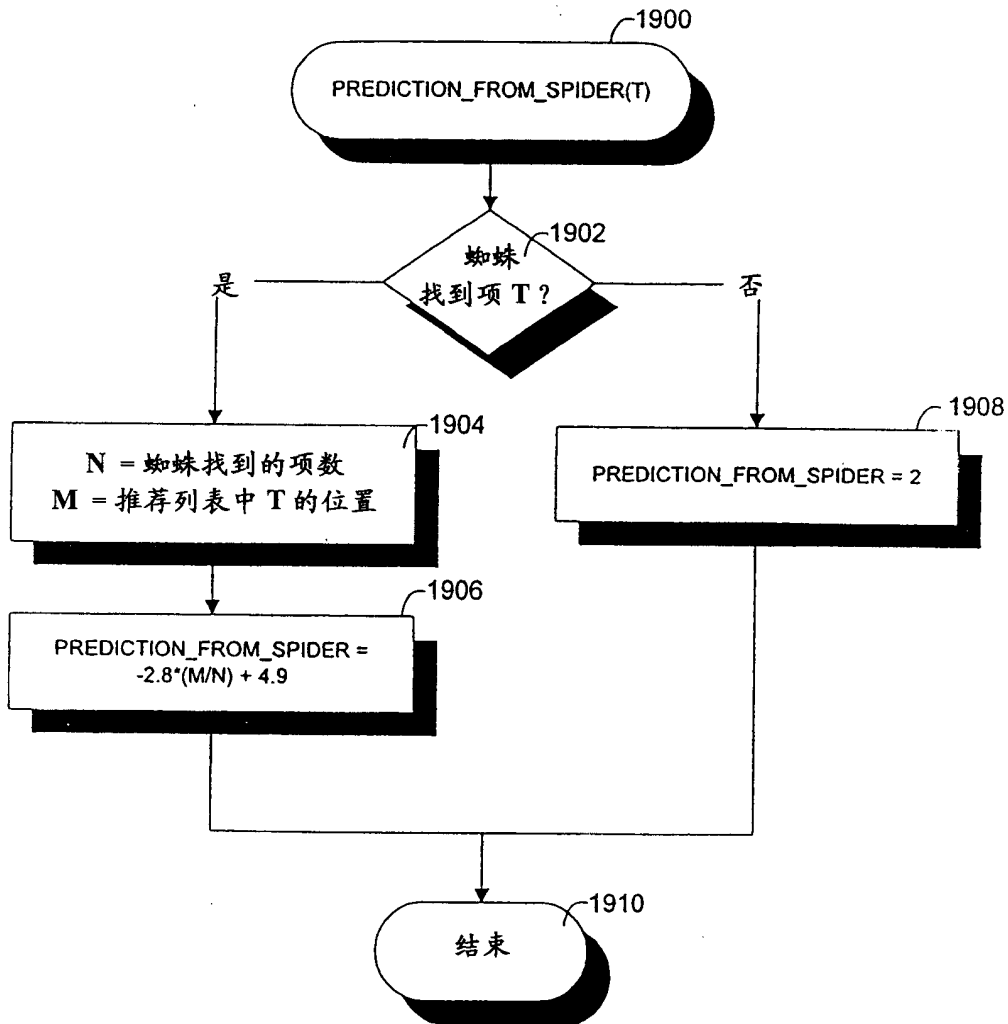


图 20

