

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06Q 30/00 (2006.01)

G06F 17/30 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200910093751.1

[43] 公开日 2010年3月17日

[11] 公开号 CN 101673385A

[22] 申请日 2009.9.28

[21] 申请号 200910093751.1

[71] 申请人 百度在线网络技术(北京)有限公司
地址 100080 北京市海淀区北四环西路58号
理想国际大厦12层

[72] 发明人 王健民 黄晶 姜轲

[74] 专利代理机构 北京同立钧成知识产权代理有限公司
代理人 刘芳

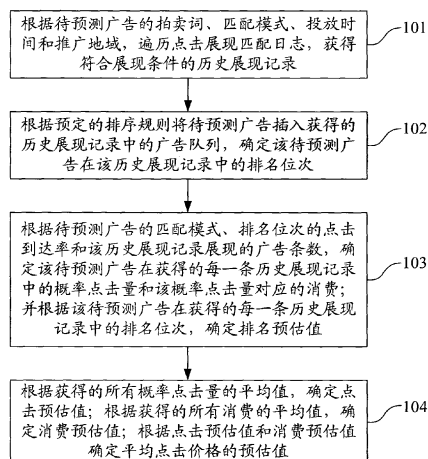
权利要求书7页 说明书15页 附图4页

[54] 发明名称

消费预估方法和装置

[57] 摘要

本发明实施例公开了一种消费预估方法和装置,所述消费预估方法包括:根据待预测广告的拍卖词、匹配模式、投放时间和推广地域,遍历点击展现匹配日志,获得符合展现条件的历史展现记录;根据预定的排序规则将待预测广告插入历史展现记录中的广告队列,确定待预测广告的排名位次;根据待预测广告的匹配模式、排名位次的点击到达率和该历史展现记录展现的广告条数,确定待预测广告的概率点击量和对应的消费,进而确定点击预估值和消费预估值;并根据待预测广告在获得的每一条历史展现记录中的排名位次,确定排名预估值。本发明实施例实现了充分利用历史数据和待预测广告的广告属性,进行消费预估,提高了消费预估的准确度,简化了消费预估的过程。



1、一种消费预估方法，其特征在于，包括：

根据待预测广告的拍卖词、匹配模式、投放时间和推广地域，遍历点击展现匹配日志，获得符合展现条件的历史展现记录；

根据预定的排序规则将所述待预测广告插入所述历史展现记录中的广告队列，确定所述待预测广告在所述历史展现记录中的排名位次；

根据所述待预测广告的匹配模式、所述排名位次的点击到达率和所述历史展现记录展现的广告条数，确定所述待预测广告在获得的每一条历史展现记录中的概率点击量和所述概率点击量对应的消费；并根据所述待预测广告在获得的每一条历史展现记录中的排名位次，确定排名预估值；

根据获得的所有概率点击量的平均值，确定点击预估值；根据获得的所有消费的平均值，确定消费预估值；根据所述点击预估值和所述消费预估值确定平均点击价格的预估值。

2、根据权利要求1所述的方法，其特征在于，还包括：

根据点击日志的标识，在展现日志中查找与所述点击日志的标识匹配的展现日志，作为所述点击展现匹配日志；所述点击展现匹配日志中的每条历史展现记录包括检索词、展现地域、展现时间、展现的广告条数、所述历史展现记录中每条广告的标识、质量得分、最高点击出价、排名位次和匹配模式。

3、根据权利要求1所述的方法，其特征在于，所述根据待预测广告的拍卖词、匹配模式、投放时间和推广地域，遍历点击展现匹配日志，获得符合展现条件的历史展现记录包括：

读取所述点击展现匹配日志中的一条历史展现记录；

当所述待预测广告的匹配模式为广泛匹配、所述历史展现记录的标识为所述待预测广告的拍卖词的标识或者所述历史展现记录的标识在所述待预测广告的拍卖词的标识对应的历史展现记录的标识列表中、所述历史展现记录

的展现地域在所述待预测广告的推广地域中且所述历史展现记录的展现时间在所述待预测广告的投放时间范围内时, 确定所述历史展现记录为符合展现条件的历史展现记录;

当所述待预测广告的匹配模式为精确匹配、所述历史展现记录的标识为所述待预测广告的拍卖词的标识、所述历史展现记录的展现地域在所述待预测广告的推广地域中且所述历史展现记录的展现时间在所述待预测广告的投放时间范围内时, 确定所述历史展现记录为符合展现条件的历史展现记录。

4、根据权利要求 1 所述的方法, 其特征在于, 所述根据预定的排序规则将所述待预测广告插入所述历史展现记录中的广告队列, 确定所述待预测广告在所述历史展现记录中的排名位次包括:

根据所述历史展现记录中展现的广告的质量得分和最高点击出价, 计算所述历史展现记录中展现的所有广告的质量得分和最高点击出价的乘积;

根据所述质量得分和最高点击出价的乘积对所述历史展现记录中展现的所有广告进行排序, 获得广告队列;

计算所述待预测广告的质量得分和最高点击出价的乘积, 根据所述待预测广告的质量得分和最高点击出价的乘积, 将所述待预测广告插入所述广告队列, 获得所述待预测广告在所述历史展现记录中的排名位次。

5、根据权利要求 4 所述的方法, 其特征在于, 还包括:

当所述广告队列中存在与所述待预测广告相同的广告时, 从所述广告队列中删除与所述待预测广告相同的广告, 并将排名位次在与所述待预测广告相同的广告的排名位次之后的广告顺次前移。

6、根据权利要求 1 所述的方法, 其特征在于, 所述根据所述待预测广告在获得的每一条历史展现记录中的排名位次, 确定排名预估值包括:

统计每个排名位次的出现次数, 确定出现次数最多的排名位次为所述排名预估值。

7、根据权利要求 1 所述的方法, 其特征在于, 所述根据获得的所有概率

点击量的平均值，确定点击预估值包括：

根据获得的所有概率点击量的平均值，确定点击参考值；

根据所述点击参考值和确定的修正因子，确定点击修正值；

根据所述点击修正值和确定的点击增长因子，确定所述点击预估值。

8、根据权利要求 7 所述的方法，其特征在于，所述确定点击增长因子的步骤包括：

计算历史一年中每天的点击总量的对数值，根据所述对数值确定历史预定单位时间内的点击次数的均值；

根据所述历史预定单位时间内的点击次数的均值，确定二次回归曲线函数；

根据所述二次回归曲线函数计算未来预定单位时间内的点击次数的均值；

计算所述未来预定单位时间内的点击次数的均值与所述历史预定单位时间内的点击次数的均值的比值，确定所述比值为所述点击增长因子。

9、根据权利要求 1 所述的方法，其特征在于，所述根据获得的所有消费的平均值，确定消费预估值包括：

根据获得的所有消费的平均值，确定消费参考值；

根据所述消费参考值和确定的修正因子，确定消费修正值；

根据所述消费修正值和确定的消费增长因子，确定所述消费预估值。

10、根据权利要求 9 所述的方法，其特征在于，所述确定消费增长因子的步骤包括：

计算历史一年中每天的消费总量的对数值，根据所述对数值确定历史预定单位时间内的消费均值；

根据所述历史预定单位时间内的消费均值，确定二次回归曲线函数；

根据所述二次回归曲线函数计算未来预定单位时间内的消费均值；

计算所述未来预定单位时间内的消费均值与所述历史预定单位时间

内的消费均值的比值，确定所述比值为所述消费增长因子。

11、根据权利要求 7 或 9 所述的方法，其特征在于，所述确定修正因子的步骤包括：

根据所述待预测广告在所述点击展现匹配日志的展现时间内的真实点击总量和所述待预测广告在所述点击展现匹配日志的所有历史展现记录中累计的点击量，确定所述修正因子。

12、一种消费预估装置，其特征在于，包括：

记录获得模块，用于根据待预测广告的拍卖词、匹配模式、投放时间和推广地域，遍历点击展现匹配日志，获得符合展现条件的历史展现记录；

排名确定模块，用于根据预定的排序规则将所述待预测广告插入所述记录获得模块获得的历史展现记录中的广告队列，确定所述待预测广告在所述历史展现记录中的排名位次；

概率点击量获得模块，用于根据所述待预测广告的匹配模式、所述排名确定模块确定的排名位次的点击到达率和所述历史展现记录展现的广告条数，确定所述待预测广告在获得的每一条历史展现记录中的概率点击量和所述概率点击量对应的消费；

排名预估值确定模块，用于根据所述待预测广告在获得的每一条历史展现记录中的排名位次，确定排名预估值；

点击预估值确定模块，用于根据所述概率点击量获得模块获得的所有概率点击量的平均值，确定点击预估值；

消费预估值确定模块，用于根据所述概率点击量获得模块获得的所有消费的平均值，确定消费预估值；

平均点击价格预估值确定模块，用于根据所述点击预估值确定模块确定的点击预估值和所述消费预估值确定模块确定的消费预估值确定平均点击价格的预估值。

13、根据权利要求 12 所述的消费预估装置，其特征在于，还包括：

日志获得模块，用于根据点击日志的标识，在展现日志中查找与所述点击日志的标识匹配的展现日志，作为所述点击展现匹配日志；所述点击展现匹配日志中的每条历史展现记录包括检索词、展现地域、展现时间、展现的广告条数、所述历史展现记录中每条广告的标识、质量得分、最高点击出价、排名位次和匹配模式。

14、根据权利要求 12 所述的消费预估装置，其特征在于，所述记录获得模块包括：

读取子模块，用于读取所述点击展现匹配日志中的一条历史展现记录；

确定子模块，用于当所述待预测广告的匹配模式为广泛匹配、所述读取子模块读取的历史展现记录的标识为所述待预测广告的拍卖词的标识或者所述读取子模块读取的历史展现记录的标识在所述待预测广告的拍卖词的标识对应的历史展现记录的标识列表中、所述读取子模块读取的历史展现记录的展现地域在所述待预测广告的推广地域中且所述读取子模块读取的历史展现记录的展现时间在所述待预测广告的投放时间范围内时，确定所述读取子模块读取的历史展现记录为符合展现条件的历史展现记录；或者，

所述确定子模块用于当所述待预测广告的匹配模式为精确匹配、所述读取子模块读取的历史展现记录的标识为所述待预测广告的拍卖词的标识、所述读取子模块读取的历史展现记录的展现地域在所述待预测广告的推广地域中且所述读取子模块读取的历史展现记录的展现时间在所述待预测广告的投放时间范围内时，确定所述读取子模块读取的历史展现记录为符合展现条件的历史展现记录。

15、根据权利要求 12 所述的消费预估装置，其特征在于，所述排名确定模块包括：

计算子模块，用于根据所述历史展现记录中展现的广告的质量得分和最高点击出价，计算所述历史展现记录中展现的所有广告的质量得分和最高点击出价的乘积；

排序子模块，用于根据所述计算子模块计算的质量得分和最高点击出价的乘积对所述历史展现记录中展现的所有广告进行排序，获得广告队列；

所述计算子模块还用于计算所述待预测广告的质量得分和最高点击出价的乘积；

所述排名确定模块还包括：插入子模块，用于根据所述计算子模块计算的待预测广告的质量得分和最高点击出价的乘积，将所述待预测广告插入所述广告队列，获得所述待预测广告在所述历史展现记录中的排名位次。

16、根据权利要求 15 所述的消费预估装置，其特征在于，所述排名确定模块还包括：

删除子模块，用于当所述广告队列中存在与所述待预测广告相同的广告时，从所述广告队列中删除与所述待预测广告相同的广告，并将排名位次在与所述待预测广告相同的广告的排名位次之后的广告顺次前移。

17、根据权利要求 12 所述的消费预估装置，其特征在于，所述排名预估值确定模块具体用于统计所述排名确定模块确定的每个排名位次的出现次数，确定出现次数最多的排名位次为所述排名预估值。

18、根据权利要求 12 所述的消费预估装置，其特征在于，还包括：

修正因子确定模块，用于根据所述待预测广告在所述点击展现匹配日志的展现时间内的真实点击总量和所述待预测广告在所述点击展现匹配日志的所有历史展现记录中累计的点击量，确定所述修正因子。

19、根据权利要求 12 所述消费预估装置，其特征在于，还包括：

点击增长因子确定模块，用于计算历史一年中每天的点击总量的对数值，根据所述对数值确定历史预定单位时间内的点击次数的均值；根据所述历史预定单位时间内的点击次数的均值，确定二次回归曲线函数；根据所述二次回归曲线函数计算未来预定单位时间内的点击次数的均值；计算所述未来预定单位时间内的点击次数的均值与所述历史预定单位时间内的点击次数的均值的比值，确定所述比值为所述点击增长因子。

20、根据权利要求 18 或 19 所述的消费预估装置，其特征在于，所述点击预估值确定模块具体用于根据所述概率点击量获得模块获得的所有概率点击量的平均值，确定点击参考值，根据所述点击参考值和所述修正因子确定模块确定的修正因子，确定点击修正值；并根据所述点击修正值和所述点击增长因子确定模块确定的点击增长因子，确定所述点击预估值。

21、根据权利要求 12 所述的消费预估装置，其特征在于，还包括：

消费增长因子确定模块，用于计算历史一年中每天的消费总量的对数值，根据所述对数值确定历史预定单位时间内的消费均值；根据所述历史预定单位时间内的消费均值，确定二次回归曲线函数；根据所述二次回归曲线函数计算未来预定单位时间内的消费均值；计算所述未来预定单位时间内的消费均值与所述历史预定单位时间内的消费均值的比值，确定所述比值为所述消费增长因子。

22、根据权利要求 18 或 21 所述的消费预估装置，其特征在于，所述消费预估值确定模块具体用于根据所述概率点击量获得模块获得的所有消费的平均值，确定消费参考值，根据所述消费参考值和所述修正因子确定模块确定的修正因子，确定消费修正值；并根据所述消费修正值和所述消费增长因子确定模块确定的消费增长因子，确定所述消费预估值。

消费预估方法和装置

技术领域

本发明实施例涉及网络广告技术领域，特别涉及一种消费预估方法和装置。

背景技术

消费预估是根据按照按点击付费（Cost Per Click；以下简称：CPC）模式计费的广告所购买的拍卖词，以及该广告展现的匹配模式、竞价、搁置信息、投放地域和预算等属性信息，预测该广告在在线推广系统中未来一周的平均点击次数、消费、平均点击价格（Average Click Price；以下简称：ACP）和排名情况。

现有的消费预估方法主要有两种：一种是基于瞬时系统的真实展现和竞价情况，虚拟N次展现，将待预测的广告进行插入排序，得到计费 and 排名。但是这种方式只能提供ACP和排名的预估，并且由于瞬时系统的真实展现情况，无法体现出不同时间的展现的差异，预估的准确程度较低。

现有的另一种消费预估方法是基于统计的方式，基于待预测广告的历史点击次数、消费、ACP和排名情况，以及其购买的拍卖词的历史点击次数、消费、ACP和排名情况，通过回归、时序或机器学习的方式得到预估结果。虽然这种方式可以提供所有的预估结果，但是由于这种方式完全采用统计的方式，使得最终的预估结果偏向平滑，尤其在待预测广告的属性出现较大改变时，预估结果的准确程度较低。另外，在采用机器学习的方式进行消费预估时，模型的建立以及模型的更换都是非常耗时的。

发明内容

本发明实施例提供一种消费预估方法和装置，以实现提高消费预估的准确度，简化消费预估的过程。

本发明实施例提供了一种消费预估方法，包括：

根据待预测广告的拍卖词、匹配模式、投放时间和推广地域，遍历点击展现匹配日志，获得符合展现条件的历史展现记录；

根据预定的排序规则将所述待预测广告插入所述历史展现记录中的广告队列，确定所述待预测广告在所述历史展现记录中的排名位次；

根据所述待预测广告的匹配模式、所述排名位次的点击到达率和所述历史展现记录展现的广告条数，确定所述待预测广告在获得的每一条历史展现记录中的概率点击量和所述概率点击量对应的消费；并根据所述待预测广告在获得的每一条历史展现记录中的排名位次，确定排名预估值；

根据获得的所有概率点击量的平均值，确定点击预估值；根据获得的所有消费的平均值，确定消费预估值；根据所述点击预估值和所述消费预估值确定平均点击价格的预估值。

本发明实施例还提供一种消费预估装置，包括：

记录获得模块，用于根据待预测广告的拍卖词、匹配模式、投放时间和推广地域，遍历点击展现匹配日志，获得符合展现条件的历史展现记录；

排名确定模块，用于根据预定的排序规则将所述待预测广告插入所述记录获得模块获得的历史展现记录中的广告队列，确定所述待预测广告在所述历史展现记录中的排名位次；

概率点击量获得模块，用于根据所述待预测广告的匹配模式、所述排名确定模块确定的排名位次的点击到达率和所述历史展现记录展现的广告条数，确定所述待预测广告在获得的每一条历史展现记录中的概率点击量和所述概率点击量对应的消费；

排名预估值确定模块，用于根据所述待预测广告在获得的每一条历史展

现记录中的排名位次，确定排名预估值；

点击预估值确定模块，用于根据所述概率点击量获得模块获得的所有概率点击量的平均值，确定点击预估值；

消费预估值确定模块，用于根据所述概率点击量获得模块获得的所有消费的平均值，确定消费预估值；

平均点击价格预估值确定模块，用于根据所述点击预估值确定模块确定的点击预估值和所述消费预估值确定模块确定的消费预估值确定平均点击价格的预估值。

本发明实施例利用历史一段时间的点击展现匹配日志建立广告展现点击过程的环境，然后根据待预测广告的广告属性查找符合该待预测广告的推广条件的有点击的历史展现记录，并根据预定的排序规则确定待预测广告在查找到的历史展现记录中的排名位次，进而可以根据待预测广告的广告属性和排名位次确定待预测广告的概率点击量，并对概率点击量进行修正，同时参考流量的变化，确定点击预估值和消费预估值，排名预估值即为出现次数最多的排名位次。从而实现了充分利用历史数据和待预测广告的广告属性，进行消费预估，提高了消费预估的准确度，并且简化了消费预估的过程。

附图说明

为了更清楚地说明本发明或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图1为本发明消费预估方法一个实施例的流程图；

图2为本发明获得符合展现条件的历史展现记录一个实施例的流程图；

图3为本发明消费预估装置一个实施例的结构示意图；

图4为本发明消费预估装置另一个实施例的结构示意图。

具体实施方式

下面将结合本发明中的附图，对本发明中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅是本发明一部分实施例，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做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。

图1为本发明消费预估方法一个实施例的流程图，如图1所示，该实施例包括：

步骤101，根据待预测广告的拍卖词、匹配模式、投放时间和推广地域，遍历点击展现匹配日志，获得符合展现条件的历史展现记录。

本实施例中，获得点击展现匹配日志的过程具体可以为：

根据点击日志的标识，在展现日志中查找与该点击日志的标识匹配的展现日志，作为点击展现匹配日志；该点击展现匹配日志中的每条历史展现记录均包括检索词、展现地域、展现时间、展现的广告条数、该历史展现记录中每条广告的标识、质量得分（Q）、最高点击出价（Bid）、排名位次和匹配模式等。

图2为本发明获得符合展现条件的历史展现记录一个实施例的流程图，如图2所示，本实施例包括：

步骤201，读取点击展现匹配日志中的一条历史展现记录。

步骤202，判断待预测广告的匹配模式是否为广泛匹配。如果该待预测广告的匹配模式是广泛匹配，则执行步骤203；如果该待预测广告的匹配模式不是广泛匹配，而是精确匹配，则执行步骤204。

步骤203，判断读取的这条历史展现记录的标识是否为待预测广告的拍卖词的标识，或者读取的这条历史展现记录的标识是否在待预测广告的拍卖词的标识对应的历史展现记录的标识列表中。如果读取的这条历史展现记录的标识是待预测广告的拍卖词的标识，或者读取的这条历史展现记录的标识

在待预测广告的拍卖词的标识对应的历史展现记录的标识列表中，则执行步骤 205；如果读取的这条历史展现记录的标识不是待预测广告的拍卖词的标识，并且读取的这条历史展现记录的标识不在待预测广告的拍卖词的标识对应的历史展现记录的标识列表中，则执行步骤 208。

本实施例中，为快速定位待预测广告可以展现的历史记录，需要为检索词设置标识，如果检索词在拍卖词列表中，则以拍卖词的标识（Word Identifier；以下简称：Wordid）作为检索词的标识；如果检索词不在拍卖词列表中，则以展现记录的标识（Searchid）作为检索词的标识。

本实施例中，获得待预测广告的拍卖词的标识对应的历史展现记录的标识列表具体可以为：遍历点击展现匹配日志，读取该点击展现匹配日志中的一条历史展现记录，当该历史展现记录中存在匹配模式为广泛匹配的广告时，获得该广告的拍卖词的标识对应的历史展现记录的标识；重发上述过程，直至将点击展现匹配日志中的历史展现记录读取完毕，获得广告的拍卖词的标识对应的历史展现记录的标识列表。本实施例将获得的广告的拍卖词的标识对应的历史展现记录的标识列表保存在数据库中，以备后续获得符合展现条件的历史展现记录时使用。

步骤 204，判断读取的这条历史展现记录的标识是否为待预测广告的拍卖词的标识。

本实施例中，当待预测广告的匹配模式为精确匹配时，可以直接利用待预测广告的拍卖词的标识来定位可以展现的历史展现记录。

如果读取的历史展现记录的标识是待预测广告的拍卖词的标识，则执行步骤 205；如果读取的历史展现记录的标识不是待预测广告的拍卖词的标识，则执行步骤 208。

步骤 205，判断读取的这条历史展现记录的展现地域是否在待预测广告的推广地域中；如果在，则执行步骤 206；如果这条历史展现记录的展现地域不在待预测广告的推广地域中，则执行步骤 208。

步骤 206, 判断读取的这条历史展现记录的展现时间是否在待预测广告的投放时间范围内。如果在, 则执行步骤 207; 如果这条历史展现记录的展现时间不在待预测广告的投放时间范围内, 则执行步骤 208。

步骤 207, 确定读取的这条历史展现记录为符合展现条件的历史展现记录。

步骤 208, 判断是否已将点击展现匹配日志中的历史展现记录读取完毕。如果是, 则结束本流程; 如果尚未将点击展现匹配日志中的历史展现记录读取完毕, 则返回执行步骤 201。

步骤 102, 根据预定的排序规则将待预测广告插入获得的历史展现记录中的广告队列, 确定该待预测广告在该历史展现记录中的排名位次。具体可以为:

根据历史展现记录中展现的广告的质量得分和最高点击出价, 计算该历史展现记录中展现的所有广告的质量得分和最高点击出价的乘积; 并根据质量得分和最高点击出价的乘积对该历史展现记录中展现的所有广告进行排序, 获得广告队列; 然后, 计算待预测广告的质量得分和最高点击出价的乘积, 根据待预测广告的质量得分和最高点击出价的乘积, 将该待预测广告插入上述广告队列, 获得该待预测广告在历史展现记录中的排名位次。具体地, 在根据质量得分和最高点击出价的乘积进行排序时, 如果存在质量得分和最高点击出价的乘积相同的广告, 则将标识较小的广告排在前面; 对于没有标识的新广告, 排在有标识的广告的后面。这只是对质量得分和最高点击出价的乘积相同的广告的一种排序方式, 本发明实施例并不仅限于此, 也可采用其他方式对质量得分和最高点击出价的乘积相同的广告进行排序, 不影响本发明实施例的实现。

优选地, 在将待预测广告插入广告队列之前, 可以先判断在广告队列中是否存在与待预测广告相同的广告, 如果在广告队列中存在与待预测广告相同的广告, 则从该广告队列中删除与待预测广告相同的广告, 并将排名位次

在与待预测广告相同的广告的排名位次之后的广告顺次前移。如果广告队列中不存在与待预测广告相同的广告，则根据待预测广告的质量得分和最高点击出价的乘积，将待预测广告插入广告队列。

步骤 103，根据待预测广告的匹配模式、排名位次的点击到达率和该历史展现记录展现的广告条数，确定该待预测广告在获得的每一条历史展现记录中的概率点击量和该概率点击量对应的消费；并根据该待预测广告在获得的每一条历史展现记录中的排名位次，确定排名预估值。

本实施例中，根据待预测广告的匹配模式、排名位次的点击到达率和该历史展现记录展现的广告条数，可以确定该待预测广告在获得的每一条历史展现记录中的概率点击量和该概率点击量对应的消费，计算方式如下：

对于待预测广告 i ，在历史展现记录 s 中的概率点击量和该概率点击量对应的消费为：

$$\text{click}[i,s] = \frac{\text{TYPE_PARA} \times \text{CTR}[r_{i,s}]}{\sum_{k=1}^{N_s} \text{TYPE_PARA} \times \text{CTR}[r_{k,s}]} \quad (1)$$

$$\text{consume}[i,s] = \text{click}[i,s] \times \text{price}[i,s] \quad (2)$$

其中， $\text{click}[i,s]$ 为待预测广告 i ，在历史展现记录 s 中的概率点击量； $\text{consume}[i,s]$ 为概率点击量 $\text{click}[i,s]$ 对应的消费。

式 (1) 中， TYPE_PARA 标识不同匹配模式的广告被点击的差异，为一常数，对于匹配模式为精确匹配的广告，可以取 $\text{TYPE_PARA} = 1$ ；对于匹配模式为广泛匹配的广告，可以取 $\text{TYPE_PARA} = 0.6$ 。 CTR 是点击到达率 (Click Through Rate) 的简称， $r_{i,s}$ 表示待预测广告 i ，在历史展现记录 s 中的排名位次， $\text{CTR}[r_{i,s}]$ 表示待预测广告 i ，在历史展现记录 s 中的排名位次对应的点击到达率； $r_{k,s}$ 表示广告 k 在历史展现记录 s 中的排名位次， $\text{CTR}[r_{k,s}]$ 表示广告 k ，在历史展现记录 s 中的排名位次对应的点击到达率。其中，点击到达率是已确定好的数值，历史展现记录 s 中不同的排名位次对应不同的点击到达率，例如：排名位次为 1 时，点击到达率为 95%；排名位次为 2 时，点击到达率为 90% 等。该点击到达率可以保存在数据库或一个文件中，本发明

实施例不限定点击到达率的保存方式。

式(2)中, $\text{price}[i,s]$ 标识的是广告 i 在历史展现记录 s 上的计费价格。下面以通用二位计费 (General Second Price; 以下简称: GSP) 方式为例, 介绍 $\text{price}[i,s]$ 的计算方式。在 GSP 方式中, 若待预测广告在当前的竞价情况下的排名位次是最后一名, 则

$\text{price}[i,s] =$ 待预测广告 i 的最低展现价格;

若待预测广告在当前竞价情况下的排名位次不是最后一名, 则

$$\text{price}[i,s] = \max \left(\frac{Q[r_{i,s}+1,s] \times \text{bid}[r_{i,s}+1,s]}{Q[r_{i,s},s]}, \text{待预测广告 } i \text{ 的最低展现价格} \right)。$$

以上介绍了一种 $\text{price}[i,s]$ 的计算方式, 但本发明实施例并不局限于此, $\text{price}[i,s]$ 还可采用其他计算方式, 本发明实施例对 $\text{price}[i,s]$ 的计算方式不作限定。

本实施例中, 统计该待预测广告在获得的每一条历史展现记录中的排名位次的出现次数, 确定出现次数最多的排名位次为排名预估值。例如: 假设共有 s 、 t 、 u 、 v 和 x 这 5 条历史展现记录, 待预测广告 i 在历史展现记录 s 、 t 、 u 和 v 中的排名位次为 6, 在历史展现记录 x 中的排名位次为 5, 因此排名位次 6 的出现次数为 4, 排名位次 5 的出现次数为 1, 则确定排名位次 6 为待预测广告的排名预估值。

步骤 104, 根据获得的所有概率点击量的平均值, 确定点击预估值; 根据获得的所有消费的平均值, 确定消费预估值; 根据点击预估值和消费预估值确定平均点击价格的预估值。

其中, 根据获得的所有概率点击量的平均值, 确定点击预估值具体可以为:

(a) 根据获得的所有概率点击量的平均值, 确定点击参考值; 具体地, 在对获得的每一条历史展现记录都计算完毕时, 累计所有的概率点击量, 计算所有概率点击量的平均值, 作为点击参考值, 如式(3)所示,

$$\text{点击参考值} = \frac{\sum \text{click}[i,s]}{T} \quad (3)$$

式(3)中, T 为总时间长度, 当获得的历史展现记录为历史一周的展现记录时, $T = 7$; 当获得的历史展现记录为历史一月的展现记录时, $T = 28、29、30$ 或 31 , 以此类推。

(b) 根据该点击参考值和确定的修正因子, 确定点击修正值; 具体地, 由于概率点击量与真实点击量存在一定的偏差, 因此需要根据待预测广告在点击展现匹配日志的展现时间内的真实点击总量和待预测广告在该点击展现匹配日志的所有历史展现记录中累计的点击量, 确定修正因子, 如式(4)所示,

$$\text{修正因子} = \frac{\text{Sum}_{\text{click}}}{\text{click}[i]} \quad (4)$$

式(4)中, $\text{Sum}_{\text{click}}$ 为待预测广告在点击展现匹配日志的展现时间内的真实点击总量, 其中, 点击展现匹配日志的展现时间可以为历史一周或历史一个月; $\text{click}[i]$ 为利用式(1)获得的该点击展现匹配日志中所有历史展现记录的 $\text{click}[i,s]$ 的累加值。

然后, 利用修正因子对点击参考值进行修正, 获得点击修正值, 如式(5)所示,

$$\text{点击修正值} = \text{点击参考值} \times \text{修正因子} \quad (5)$$

(c) 根据点击修正值和确定的点击增长因子, 确定点击预估值; 具体地, 确定点击增长因子的步骤可以为: 计算历史一年中每天的点击总量的对数值, 根据该对数值确定历史预定单位时间内的点击次数的均值; 根据历史预定单位时间内的点击次数的均值, 确定二次回归曲线函数; 根据二次回归曲线函数计算未来预定单位时间内的点击次数的均值; 计算未来预定单位时间内的点击次数的均值与历史预定单位时间内的点击次数的均值的比值, 确定该比值为点击增长因子。以预定单位时间为一周为例, 点击增长因子为未来一周内的点击次数的均值与历史一周内的点击次数的均值的比值, 如式(6)所示,

$$\text{点击增长因子} = \text{未来一周内的点击次数的均值} / \text{历史一周内的点击}$$

次数的均值 (6)

然后, 根据点击修正值和点击增长因子, 确定点击预估值, 如式 (7) 所示,

$$\text{点击预估值} = \text{点击修正值} \times \text{点击增长因子} \quad (7)$$

本实施例中, 根据获得的所有消费的平均值, 确定消费预估值具体可以为:

(a) 根据获得的所有消费的平均值, 确定消费参考值; 具体地, 在对获得的每一条历史展现记录都计算完毕时, 累计获得的所有消费, 计算所有消费的平均值, 作为消费参考值, 如式 (8) 所示,

$$\text{消费参考值} = \frac{\sum \text{consume}[i,s]}{T} \quad (8)$$

式 (8) 中, T 为总时间长度, 当获得的历史展现记录为历史一周的展现记录时, $T=7$; 当获得的历史展现记录为历史一月的展现记录时, $T=28$ 、29、30 或 31, 以此类推。

(b) 根据消费参考值和确定的修正因子, 确定消费修正值; 具体地, 修正因子可以通过式 (4) 所示的公式确定; 然后利用修正因子对消费参考值进行修正, 获得消费修正值, 如式 (9) 所示,

$$\text{消费修正值} = \text{消费参考值} \times \text{修正因子} \quad (9)$$

(c) 根据消费修正值和确定的消费增长因子, 确定消费预估值; 具体地, 确定消费增长因子的步骤可以为: 计算历史一年中每天的消费总量的对数值, 根据该对数值确定历史预定单位时间内的消费均值; 根据历史预定单位时间内的消费均值, 确定二次回归曲线函数; 根据该二次回归曲线函数计算未来预定单位时间内的消费均值; 计算未来预定单位时间内的消费均值与历史预定单位时间内的消费均值的比值, 确定该比值为消费增长因子。以预定单位时间为一周为例, 消费增长因子为未来一周内的消费均值与历史一周内的消费均值的比值, 如式 (10) 所示,

$$\text{消费增长因子} = \text{未来一周内的消费均值} / \text{历史一周内的消费均值} \quad (10)$$

然后,根据消费修正值和消费增长因子,确定消费预估值,如式(11)所示,

$$\text{消费预估值} = \text{消费修正值} \times \text{消费增长因子} \quad (11)$$

获得点击预估值和消费预估值之后,可以根据点击预估值和消费预估值确定平均点击价格(Average Click Price;以下简称:ACP)的预估值,如式(12)所示,

$$\text{ACP 预估值} = \text{消费预估值} / \text{点击预估值} \quad (12)$$

至此,本发明实施例的消费预估过程完成,获得了消费预估结果,包括:点击预估值、消费预估值、ACP预估值和排名预估值这四个预估值。

上述实施例利用历史一段时间的点击展现匹配日志建立广告展现点击过程的环境,然后根据待预测广告的广告属性查找符合该待预测广告的推广条件的有点击的历史展现记录,并根据预定的排序规则确定待预测广告在查找到的历史展现记录中的排名位次,进而可以根据待预测广告的广告属性和排名位次确定待预测广告的概率点击量,并对概率点击量进行修正,同时参考流量的变化,确定点击预估值和消费预估值,排名预估值即为出现次数最多的排名位次。从而实现了充分利用历史数据和待预测广告的广告属性,进行消费预估,提高了消费预估的准确度,并且简化了消费预估的过程。

本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于一计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤;而前述的存储介质包括:ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

图3为本发明消费预估装置一个实施例的结构示意图,本实施例的消费预估装置可以作为服务器、或服务器的一部分,实现本发明图1所示实施例的流程。如图3所示,该消费预估装置可以包括:记录获得模块31、排名确定模块32、概率点击量获得模块33、排名预估值确定模块34、点击预估值确定模块35、消费预估值确定模块36和平均点击价格预估值确定模块37。

其中,记录获得模块 31 可以根据待预测广告的拍卖词、匹配模式、投放时间和推广地域,遍历点击展现匹配日志,获得符合展现条件的历史展现记录;排名确定模块 32 可以根据预定的排序规则将待预测广告插入记录获得模块 31 获得的历史展现记录中的广告队列,确定待预测广告在历史展现记录中的排名位次;概率点击量获得模块 33 可以根据待预测广告的匹配模式、排名确定模块 32 确定的排名位次的点击到达率和历史展现记录展现的广告条数,确定待预测广告在获得的每一条历史展现记录中的概率点击量和该概率点击量对应的消费;排名预估值确定模块 34 可以根据待预测广告在获得的每一条历史展现记录中的排名位次,确定排名预估值;点击预估值确定模块 35 可以根据概率点击量获得模块 33 获得的所有概率点击量的平均值,确定点击预估值;消费预估值确定模块 36 可以根据概率点击量获得模块 33 获得的所有消费的平均值,确定消费预估值;平均点击价格预估值确定模块 37 可以根据点击预估值确定模块 35 确定的点击预估值和消费预估值确定模块 36 确定的消费预估值确定平均点击价格的预估值。

上述消费预估装置实现了充分利用历史数据和待预测广告的广告属性,进行消费预估,提高了消费预估的准确度,并且简化了消费预估的过程。

图 4 为本发明消费预估装置另一个实施例的结构示意图,本实施例的消费预估装置可以作为服务器、或服务器的一部分,实现本发明图 1 或图 2 所示实施例的流程。如图 4 所示,该消费预估装置可以包括:记录获得模块 41、排名确定模块 42、概率点击量获得模块 43、排名预估值确定模块 44、点击预估值确定模块 45、消费预估值确定模块 46、平均点击价格预估值确定模块 47、日志获得模块 48、修正因子确定模块 49、点击增长因子确定模块 410 和消费增长因子确定模块 4011。

其中,记录获得模块 41 可以根据待预测广告的拍卖词、匹配模式、投放时间和推广地域,遍历点击展现匹配日志,获得符合展现条件的历史展现记录;排名确定模块 42 可以根据预定的排序规则将待预测广告插入记录获得模块 41 获得的历史展现记录中的广告队列,确定待预测广告在历史展现记录中

的排名位次；概率点击量获得模块 43 可以根据待预测广告的匹配模式、排名确定模块 42 确定的排名位次的点击到达率和历史展现记录展现的广告条数，确定待预测广告在获得的每一条历史展现记录中的概率点击量和该概率点击量对应的消费；排名预估值确定模块 44 可以根据待预测广告在获得的每一条历史展现记录中的排名位次，确定排名预估值；点击预估值确定模块 45 可以根据概率点击量获得模块 43 获得的所有概率点击量的平均值，确定点击预估值；消费预估值确定模块 46 可以根据概率点击量获得模块 43 获得的所有消费的平均值，确定消费预估值；平均点击价格预估值确定模块 47 可以根据点击预估值确定模块 45 确定的点击预估值和消费预估值确定模块 46 确定的消费预估值确定平均点击价格的预估值。

日志获得模块 48 可以根据点击日志的标识，在展现日志中查找与点击日志的标识匹配的展现日志，作为点击展现匹配日志；该点击展现匹配日志中的每条历史展现记录包括检索词、展现地域、展现时间、展现的广告条数、所述历史展现记录中每条广告的标识、质量得分、最高点击出价、排名位次和匹配模式。

记录获得模块 41 可以包括：读取子模块 411 和确定子模块 412，

读取子模块 411 可以读取点击展现匹配日志中的一条历史展现记录；

确定子模块 412 可以当待预测广告的匹配模式为广泛匹配、读取子模块 411 读取的历史展现记录的标识为待预测广告的拍卖词的标识或者读取子模块 411 读取的历史展现记录的标识在待预测广告的拍卖词的标识对应的历史展现记录的标识列表中、读取子模块 411 读取的历史展现记录的展现地域在待预测广告的推广地域中且读取子模块 411 读取的历史展现记录的展现时间在待预测广告的投放时间范围内时，确定读取子模块 411 读取的历史展现记录为符合展现条件的历史展现记录；或者，

确定子模块 412 可以当待预测广告的匹配模式为精确匹配、读取子模块 411 读取的历史展现记录的标识为待预测广告的拍卖词的标识、读取子模块 411 读取的历史展现记录的展现地域在待预测广告的推广地域中且读取子模

块 411 读取的历史展现记录的展现时间在待预测广告的投放时间范围内时, 确定读取子模块 411 读取的历史展现记录为符合展现条件的历史展现记录。

排名确定模块 42 可以包括: 计算子模块 421 和排序子模块 422; 其中, 计算子模块 421 可以根据历史展现记录中展现的广告的质量得分和最高点击出价, 计算历史展现记录中展现的所有广告的质量得分和最高点击出价的乘积; 排序子模块 422 可以根据计算子模块 421 计算的质量得分和最高点击出价的乘积对该历史展现记录中展现的所有广告进行排序, 获得广告队列;

计算子模块 421 还可以计算待预测广告的质量得分和最高点击出价的乘积;

排名确定模块 42 还可以包括: 插入子模块 423, 可以根据计算子模块 421 计算的待预测广告的质量得分和最高点击出价的乘积, 将待预测广告插入广告队列, 获得待预测广告在历史展现记录中的排名位次。

排名确定模块 42 还可以包括: 删除子模块 424, 可以在插入子模块 423 将待预测广告插入广告队列之前, 当广告队列中存在与待预测广告相同的广告时, 从广告队列中删除与待预测广告相同的广告, 并将排名位次在与待预测广告相同的广告的排名位次之后的广告顺次前移。

排名预估值确定模块 44 具体可以统计排名确定模块 42 确定的每个排名位次的出现次数, 确定出现次数最多的排名位次为排名预估值。

修正因子确定模块 49 可以根据待预测广告在点击展现匹配日志的展现时间内的真实点击总量和待预测广告在点击展现匹配日志的所有历史展现记录中累计的点击量, 确定修正因子。

点击增长因子确定模块 410 可以计算历史一年中每天的点击总量的对数值, 根据该对数值确定历史预定单位时间内的点击次数的均值; 根据该历史预定单位时间内的点击次数的均值, 确定二次回归曲线函数; 根据二次回归曲线函数计算未来预定单位时间内的点击次数的均值; 计算未来预定单位时间内的点击次数的均值与历史预定单位时间内的点击次数的均值的比值, 确定该比值为点击增长因子。点击预估值确定模块 45 具体可以

根据概率点击量获得模块 43 获得的所有概率点击量的平均值，确定点击参考值，根据该点击参考值和修正因子确定模块 49 确定的修正因子，确定点击修正值；并根据点击修正值和点击增长因子确定模块 410 确定的点击增长因子，确定点击预估值。

消费增长因子确定模块 4011 可以计算历史一年中每天的消费总量的对数值，根据该对数值确定历史预定单位时间内的消费均值；根据历史预定单位时间内的消费均值，确定二次回归曲线函数；根据二次回归曲线函数计算未来预定单位时间内的消费均值；计算未来预定单位时间内的消费均值与历史预定单位时间内的消费均值的比值，确定该比值为消费增长因子。消费预估值确定模块 46 具体可以根据概率点击量获得模块 43 获得的所有消费的平均值，确定消费参考值，根据消费参考值和修正因子确定模块 49 确定的修正因子，确定消费修正值；并根据消费修正值和消费增长因子确定模块 4011 确定的消费增长因子，确定消费预估值。

上述消费预估装置，实现了充分利用历史数据和待预测广告的广告属性，进行消费预估，提高了消费预估的准确度，并且简化了消费预估的过程。

本领域技术人员可以理解附图只是一个优选实施例的示意图，附图中的模块或流程并不一定是实施本发明所必须的。

本领域技术人员可以理解实施例中的装置中的模块可以按照实施例描述进行分布于实施例的装置中，也可以进行相应变化位于不同于本实施例的一个或多个装置中。上述实施例的模块可以合并为一个模块，也可以进一步拆分成多个子模块。

最后应说明的是：以上实施例仅用以说明本发明的技术方案而非对其进行限制，尽管参照较佳实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对本发明的技术方案进行修改或者等同替换，而这些修改或者等同替换亦不能使修改后的技术方案脱离本发明技术方案的精神和范围。

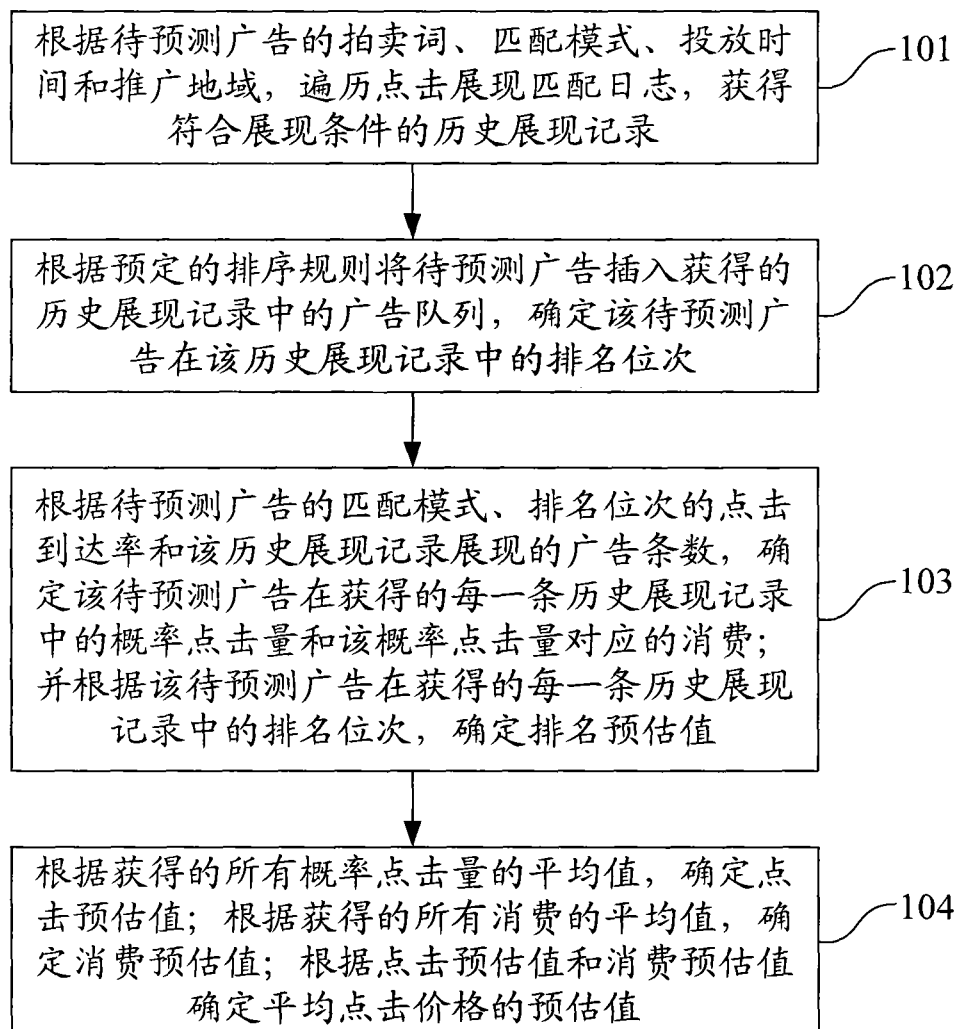


图 1

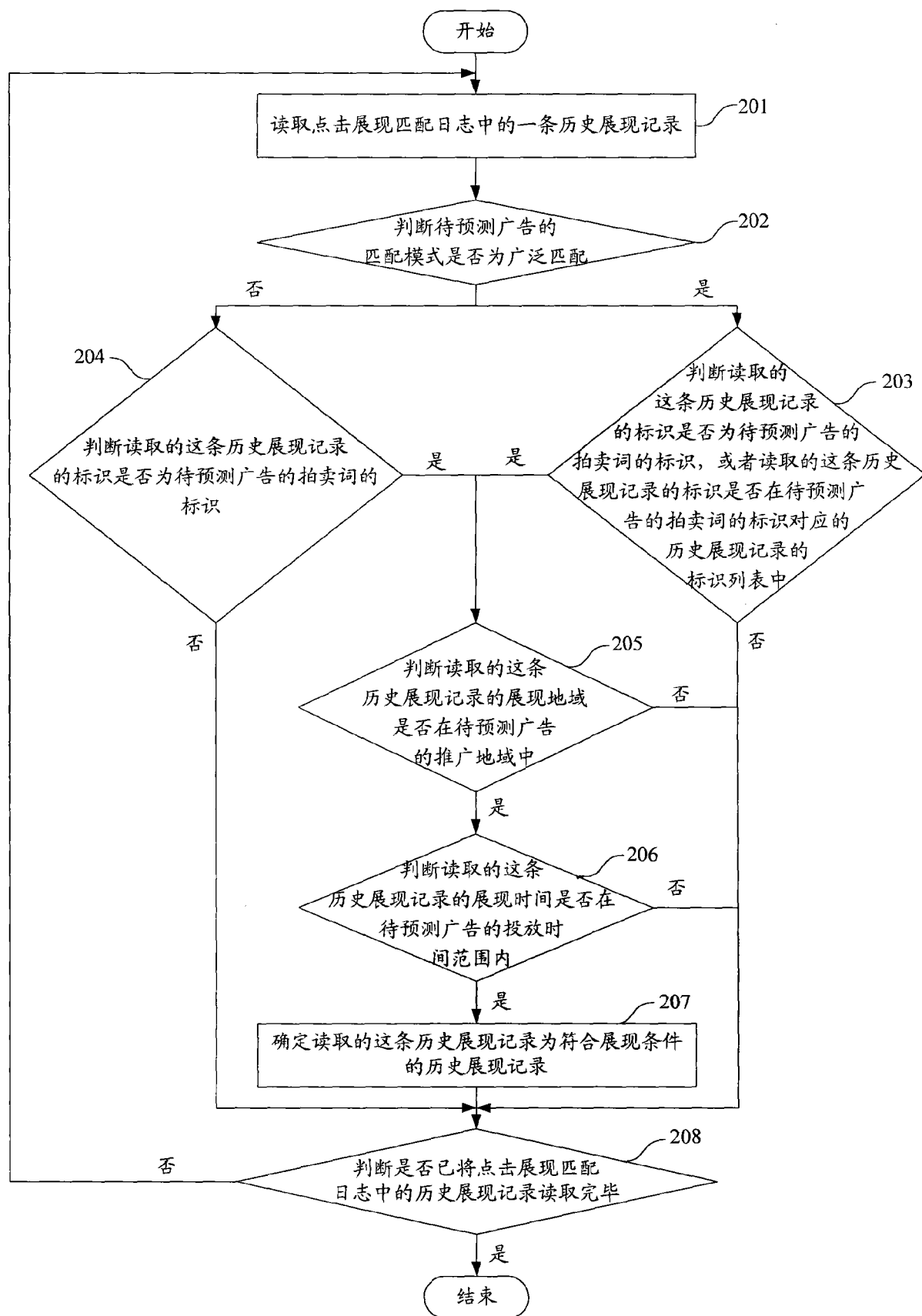


图 2

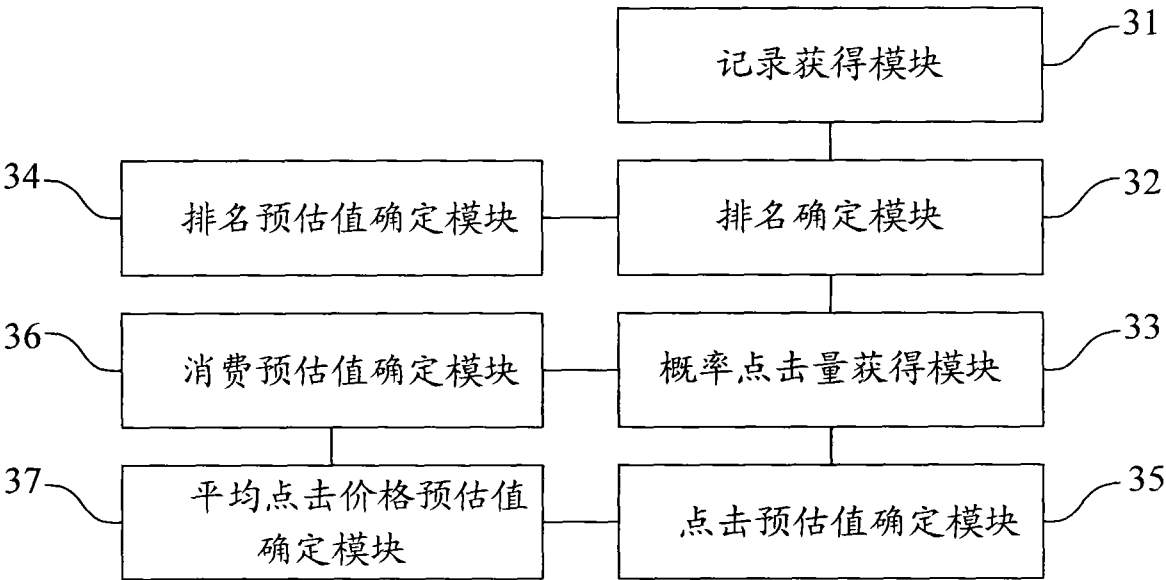


图 3

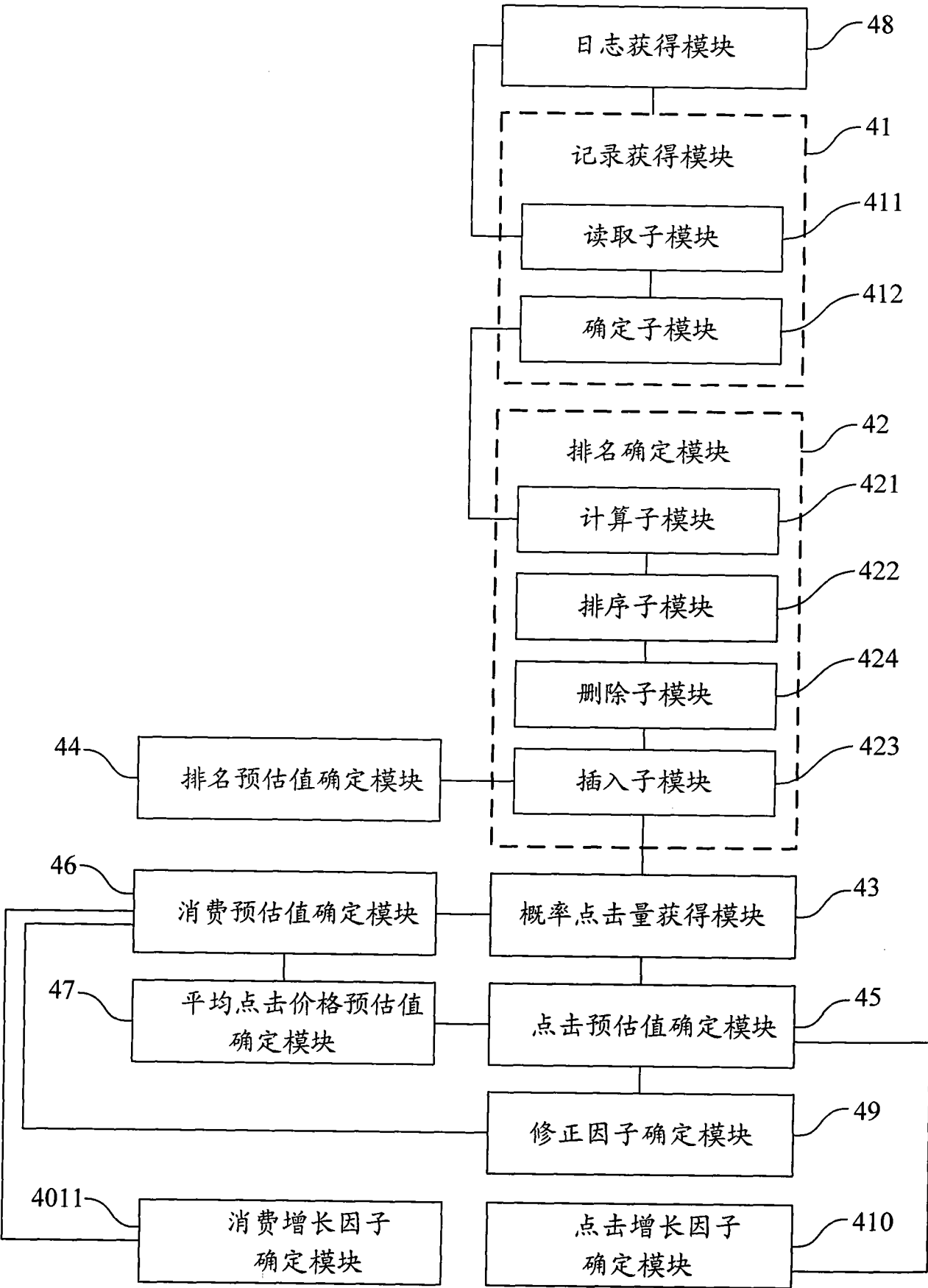


图 4