



[12] 发明专利申请公开说明书

[21] 申请号 03822800.9

[43] 公开日 2005 年 10 月 19 日

[11] 公开号 CN 1685337A

[22] 申请日 2003.9.24 [21] 申请号 03822800.9

[30] 优先权

[32] 2002. 9. 24 [33] US [31] 60/413,536

[32] 2002. 12. 6 [33] US [31] 10/314,427

[32] 2003. 3. 14 [33] US [31] 10/389,688

[86] 国际申请 PCT/US2003/030234 2003. 9. 24

[87] 国际公布 WO2004/029827 英 2004. 4. 8

[85] 进入国家阶段日期 2005. 3. 24

[71] 申请人 GOOGLE 公司

地址 美国加利福尼亚

[72] 发明人 杰弗里·A·迪安

乔治斯·R·哈瑞克

保罗·布楔特

[74] 专利代理机构 中国国际贸易促进委员会专利商
标事务所

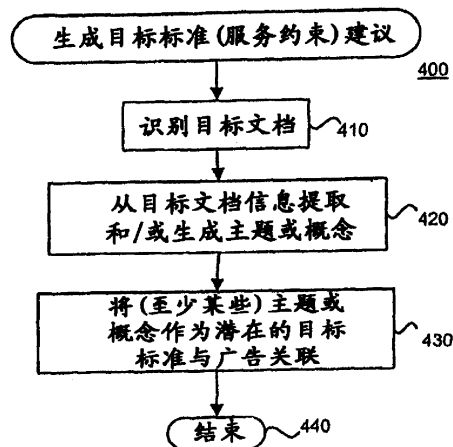
代理人 付建军

权利要求书 3 页 说明书 17 页 附图 5 页

[54] 发明名称 建议和/或提供广告供应约束信息

[57] 摘要

识别广告的目标信息(也被称为广告“供应约束”)或候选目标信息(1)。 通过从广告信息提取主题或概念(420), 和/或基于这种信息生成主题或概念(400), 可以识别目标信息(410), 所述广告信息比如其中链接了广告的 Web 页(或使广告或广告商感兴趣的某些其他 Web 页)中的信息。 主题或概念可以是与相关 Web 页、群集等等关联的相关查询。



1. 一种用于为广告生成供应约束的方法，该方法包括：
 - a) 识别目标文档；
 - 5 b) 从与目标文档关联的信息生成概念；以及
 - c) 将至少某些概念作为供应约束或作为候选供应约束与所述广告关联。
2. 根据权利要求 1 所述的方法，其中，目标文档是 Web 页面。
- 10 3. 根据权利要求 1 所述的方法，其中，基于广告中的链接来识别目标文档。
4. 根据权利要求 1 所述的方法，其中，使用用户输入来识别目标文档。
5. 根据权利要求 1 所述的方法，其中，基于与目标文档关联的概念来识别目标文档。
- 15 6. 根据权利要求 1 所述的方法，其中，目标文档是与广告关联的文档。
7. 根据权利要求 6 所述的方法，其中，广告用目标文档来呈现。
- 20 8. 根据权利要求 1 所述的方法，其中，与目标文档关联的信息包括从目标文档中提取的文字。
9. 根据权利要求 1 所述的方法，其中，与目标文档关联的信息包括与目标文档链接的其他文档。
10. 根据权利要求 1 所述的方法，其中，与目标文档关联的信息包括已返回了指向目标文档的链接的搜索引擎查询。
- 25 11. 根据权利要求 1 所述的方法，其中，与目标文档关联的信息包括将会返回指向目标文档的链接的搜索引擎查询。
12. 根据权利要求 1 所述的方法，其中，供应约束包括关键字。
13. 根据权利要求 1 所述的方法，其中，供应约束包括地理位

置信息。

14. 根据权利要求 1 所述的方法，其中，供应约束包括时间信息。

15. 根据权利要求 1 所述的方法，进一步包括：

- 5 d) 向用户发送供应约束，以便呈现；
e) 确定用户关于呈现供应约束的意图；以及
f) 响应确定的用户的意图，接受或拒绝至少某些供应约束。

16. 一种用于为广告生成供应约束的设备，该设备包括：

- a) 用于接受目标文档的输入；
10 b) 用于从与目标文档关联的信息生成概念的装置；以及
c) 用于将至少某些概念作为供应约束或作为候选供应约束与
与前述广告关联的装置。

17. 根据权利要求 16 所述的设备，其中，目标文档是 Web 页面。

15 18. 根据权利要求 16 所述的设备，其中，基于广告中的链接来识别目标文档。

19. 根据权利要求 16 所述的设备，其中，使用用户输入来识别目标文档。

20 20. 根据权利要求 16 所述的设备，其中，与目标文档关联的信息包括从目标文档中提取的文字。

21. 根据权利要求 16 所述的设备，其中，与目标文档关联的信息包括与目标文档链接的其他文档。

22. 根据权利要求 16 所述的设备，其中，与目标文档关联的信息包括已返回了指向目标文档的链接的搜索引擎查询。

25 23. 根据权利要求 16 所述的设备，其中，与目标文档关联的信息包括将会返回指向目标文档的链接的搜索引擎查询。

24. 根据权利要求 16 所述的设备，其中，供应约束包括关键字。

25 25. 根据权利要求 16 所述的设备，其中，供应约束包括地理位置信息。

26. 根据权利要求 16 所述的设备, 其中, 供应约束包括时间信息。

27. 根据权利要求 16 所述的设备, 进一步包括:

d) 用于执行下列操作的用户界面:

- 5 - 向用户发送供应约束, 以便呈现,
- 确定用户关于呈现供应约束的意图, 以及
- 响应确定的用户的意图, 接受或拒绝至少某些供应约束。

28. 根据权利要求 16 所述的设备, 其中, 基于与目标文档关联的概念来识别目标文档。

10 29. 根据权利要求 16 所述的设备, 其中, 目标文档是与广告关联的文档。

30. 根据权利要求 29 所述的设备, 其中, 所述广告用目标文档来呈现。

15 31. 一种存储机器可执行的指令的机器可读的介质, 当指令被计算机执行时, 实现根据权利要求 1 所述的方法。

32. 一种用于为广告生成内容的方法, 该方法包括:

- a) 识别目标文档;
- b) 从与目标文档关联的信息生成概念; 以及
- c) 将至少某些概念作为内容或作为候选内容与所述广告关联。

20

建议和/或提供广告供应约束信息

5 相关申请

本申请要求以 2002 年 9 月 24 日提出的 Jeffrey Dean、Georges Harik、Paul Bucheit 作为发明人的标题为“METHODS AND APPARATUS FOR SERVING RELEVANT ADVERTISEMENTS”的美国临时申请序列号 No. 60/413,536(在此进行了引用), 2002 年 12 月 6 日提出的 Jeffrey Dean、Georges Harik、Paul Bucheit 作为发明人的标题为“METHODS AND APPARATUS FOR SERVING RELEVANT ADVERTISEMENTS”的美国专利申请序列号 No. 10/314,427(在此进行了引用), 以及 2003 年 2 月 26 日提出的 Darrell Anderson、Paul Bucheit、Alex Carobus、Claire Cui、Jeffrey A. Dean、Georges R. Hank、Deepak Jindal 和 Narayanan Shivakumar 作为发明人的标题为“SERVING ADVERTISEMENTS BASED ON CONTENT”的美国专利申请序列号 No. 10/____,____(在此进行了引用)作为优先权基础。

20

技术领域

本发明涉及诸如在线广告之类的广告。具体来说, 本发明涉及帮助在线广告商确定他们的广告的目标。

25

背景技术

使用诸如电视、广播、报纸杂志之类的传统媒体的广告已为大家所熟知。广告商使用这种媒体借助于他们的广告(“ads”)影响到广大的受众。为影响到更敏感的受众, 广告商使用了人口统计研究。例如, 广告商可能使用诸如足球比赛之类的转播事件, 向年轻的男性受

众做啤酒和动作影片广告。然而，即使利用人口统计研究和有关各种媒体渠道的典型受众的完全合理的假设，广告商也认识到，他们的大部分广告预算只是浪费，因为目标受众对他们接收到的广告不感兴趣。

诸如因特网之类的交互媒体具有更好地找到广告的目标的潜力。

5 例如，某些网站提供了基于查找信息的用户输入的查询关键字的信息搜索功能。此用户查询可以被用来作为使用户感兴趣的信息类型的指针。通过将用户查询与广告商指定的关键字列表进行比较，可以向这些搜索服务用户提供某种形式的有针对性的广告。这样的系统的一个示例是 Google, Inc. 提供的 Adwords 系统。

10 尽管诸如 Adwords 之类的系统给广告商提供了更好地确定广告目标的能力，但是，广告商仍希望这些系统能够帮助他们可以更有效和/或更轻松地确定他们的广告的目标。

因此，提供用于帮助广告商更有效和/或更轻松地确定他们的广告的目标的方法、设备和数据结构是有用的。

15

发明内容

本发明通过识别广告的目标信息（也被称为广告“供应约束”或“目标标准”）来满足此需求及其他需求。通过从诸如广告创意、其中链接了广告的 Web 页中的信息等广告信息，提取主题或概念，和/或基于这种信息生成主题或概念，可以识别目标信息。通过从下列与广告的传送相关的其他信息提取主题或概念，和/或基于这种信息生成主题或概念，可以识别目标信息，所述与广告的传送相关的其他信息比如：其中可能要显示广告的文档（例如，Web 页），广告商对在其中做广告感兴趣或可能感兴趣的文档（例如，Web 页），在其上面有竞争性的、类似的或相辅的产品或服务的广告或链接了这样的内容的文档（例如，Web 页），在其上面供应或描述了竞争性的、类似的或相辅的产品或服务的文档（例如，Web 页），涉及广告的地理目标或时间目标的信息，类似于上述任何文档或信息来源的文档（例如，Web 页）等等或，一般而言，使广告商感兴趣的任何其他文档（例如，Web

20

25

页)。

在本发明的一个实施例中，主题或概念可以是与相关的 Web 页关联的字或相关的查询。 在本发明的另一个实施例中，主题或概念可以是群集。 如此，在本发明的一个实施例中，可以基于从一个或多个相关的文档（例如，Web 页）中提取的主题或概念群集，向关键字触发的广告的广告商提供关键字建议。 或者，或此外，可以向概念、主题和/或群集触发的广告的广告商提供概念、主题和/或群集建议。

本发明的其他方面旨在供应具有涉及前述的方面的功能的计算机系统 and 机器可读的介质。

10

附图说明

此说明书收入的并构成此说明书的一部分的附图说明了本发明的实施例，与说明书一起，用于说明本发明。 在图形中，

图 1 是说明可以在其中实现本发明的环境的图表；

15

图 2 是从功能上说明根据本发明的广告系统的图表；

图 3 是说明可以用来实现本发明的设备的图表；

图 4 是根据本发明，为广告商识别目标标准（例如候选目标标准）的示例方法的流程图；

图 5 是根据本发明的原理，可以用来执行各种操作或存储各种信息的示例设备；

20

图 6 是关于 Web 页是目标文档的示例。

具体实施方式

下面对本发明的详细描述参考了附图。 详细描述不对本发明作出限制。

25

本发明涉及从目标文档中确定（例如，候选）目标标准的方法、设备和数据结构。 然后，可以作为建议的目标标准向广告商提供目标标准。 在一种实现方式中，目标文档是 Web 页，广告是能够在该 Web 页上呈现，或与该 Web 页相关联的电子文件。 通过分析 Web

页的内容，可以生成对应于该 Web 页的概念或主题集，如列表。有各种各样的技术可以用来执行此任务，其中一个是通过计算 Web 页的检索词向量并从该向量中选择上面 N 个检索词。然后，可以作为与广告关联的目标标准（例如，为广告指定的关键字）提供概念或主题列表。

那些精通本技术的人将认识到，根据本发明的许多其他实现方式也是可以的。

A. 典型的环境和体系结构

图 1 是说明可以在其中使用本发明的典型环境的图表。典型环境包括广告商 110、广告系统 120、广告消费者 130、广告目标 140。

广告商 110 可以是直接销售被广告的商品或服务的一方（例如，Amazon.com）或代表广告商操作的代理。广告商 110 所希望的广告可以以各种各样的形式存在。例如，广告可以是标准的印刷广告、在线广告、文字广告，以及图像广告、音频广告、视频广告、音频/可视广告，或任何其他类型的感觉消息，或前述的所希望的形式组合。

广告系统 120 与广告商 110 和广告消费者 130 两者连接。它可以执行各种各样的功能，如下文将要参考图 2 详细讲述的。本发明可以与这样的广告系统 120 一起使用。

广告消费者 130 是向广告系统 120 发出广告请求、从广告系统 120 获取广告，并向广告目标 140 供应广告的实体。通常，广告消费者是提供内容的实体，广告将与内容关联。在一种实现方式中，广告消费者 130 是搜索引擎，如网址为 www.google.com 的 Google, Inc. 所使用的搜索引擎。

广告目标 140 是最终接收广告的个人（或个人集合）。例如，在可视广告的情况下，广告目标 140 是观看广告的人。

图 2 是从功能上说明根据本发明的广告系统的图表。该系统包括广告活动条目和管理组件 210、工具组件 220、计费组件 230、一个或多个数据库 240、广告消费者接口组件 250、广告选择组件 260、

广告排序组件 270、广告供应组件 280，以及统计信息引擎组件 290。如果本发明将与这样的广告系统一起使用，则它主要涉及工具组件 220。为帮助理解本发明，下面将说明广告系统的其他组件。此外，虽然图 2 显示了构成广告系统 120 的组件的特定方案，但是，那些精通本技术的人可以认识到，并非所有的组件都需要按如图所示的方式排列，并非需要所有的组件，并且可以添加其他组件，或替换所显示

5 的组件。

广告条目和管理组件 210 是广告商用来输入广告活动所需的信息并管理广告活动的组件。在一个实施例中，广告活动包含以某种方式相关联的一个或多个广告。例如，Ford Motor Company 可能搞一个零百分比理财的广告活动，该活动可能包含涉及该主题的一系列广告。广告商可以通过广告条目和管理组件 210 提供的其他内容包括下列各项：一个或多个广告创意（简称为“广告”），与这些创意关联的一个或多个关键字或主题集（可以用作广告的目标信息）、地理目标信息、广告的价值指示、开始日期、结束日期等等。广告条目和管理

15 组件 210 所需的数据或它所获取的数据驻留在其中一个数据库 240 中。在广告系统的一个实施例中，广告条目和管理组件 210 可以执行诸如下列文件中所描述的操作：(i) 2003 年 1 月 10 日提出的 Salar Kamangar and Amit Patel 作为发明人的标题为“PAUSING

20 ONE OR MORE ADS, ONE OR MORE AD GROUPS, AND/OR ON OR MORE AD CAMPAIGNS”的美国专利申请序列号 No. 10/340,867（在此进行了引用）；和/或 (ii) 2003 年 1 月 10 日提出的 Eric Veach 作为发明人的标题为“PRICING ACROSS KEYWORDS ASSOCIATED WITH ONE OR MORE ADVERTISEMENTS”的美国

25 专利申请序列号 No. 10/340,193（在此进行了引用）。

工具组件 220 包含用于帮助广告商 110 创建、监视和管理其广告活动的各种各样的工具。例如，工具组件 220 可以包含帮助广告商 110 估计广告为特定关键字或主题接收到的印痕的数量的工具。同样，工具组件 220 可以用来帮助广告商 110 为给定广告定义诸如

关键字列表之类的目标标准。也可以提供其他可能的工具。取决于工具的特性，可以使用一个或多个数据库 240 来收集或存储信息。在广告系统的一个实施例中，工具组件 220 可以执行诸如 2003 年 1 月 10 日提出的 Magnus Sandburg、Eric Veach、John Bauer 和 Zhe Qian 作为发明人的标题为“METHOD AND APPARATUS FOR ESTIMATING ELECTRONIC ADVERTISING INVENTORY”的美国临时申请序列号 No. 60/____,____ (在此进行了引用) 中所描述的操作。

计费组件 230 帮助执行计费相关的功能。例如，计费组件 230 为特定广告商 110 或广告活动生成发票。此外，计费组件 230 还可以被广告商 110 用来监视其各种广告活动所花费的资金总量。计费组件 230 所需的数据或它所获取的数据驻留在数据库 240 中。在广告系统的一个实施例中，计费组件 230 可以执行诸如下列文件中所描述的操作：(i) 2002 年 11 月 8 日提出的 Eric Veach 作为发明人的标题为“**AUTOMATED PRICE MAINTENANCE FOR USE WITH A SYSTEM IN WHICH ADVERTISEMENTS ARE RENDERED WITH RELATIVE PREFERENCE BASED ON PERFORMANCE INFORMATION AND PRICE INFORMATION**”的美国临时申请序列号 No. 60/424,792 (在此进行了引用)；(ii) 2003 年 1 月 10 日提出的 Eric Veach 和 Salar Kamangar 作为发明人的标题为“**AUTOMATED PRICE MAINTENANCE FOR USE WITH A SYSTEM IN WHICH ADVERTISEMENTS ARE RENDERED WITH RELATIVE PREFERENCES**”的美国专利申请序列号 No. 10/340,543 (在此进行了引用)；和/或 (iii) 2003 年 1 月 10 日提出的 Eric Veach 和 Salar Kamangar 作为发明人的标题为“**AUTOMATED PRICE MAINTENANCE FOR USE WITH A SYSTEM IN WHICH ADVERTISEMENTS ARE RENDERED WITH RELATIVE PREFERENCE BASED ON PERFORMANCE INFORMATION AND PRICE INFORMATION**”的美国专利申请序

列号 No. 10/340,542 (在此进行了引用)。

数据库 240 包含由广告系统 120 使用的各种各样的数据。除了上文参考广告条目和管理系统 210 所述的信息之外，数据库 240 还可以包含有关哪些广告被显示、显示它们的频率、它们被选择的次数、怎样的人选择了那些广告、广告的显示导致交易成功等等的统计信息。虽然数据库 240 在图 2 中是作为一个单位显示的，但是，那些精通本技术的普通人员将认识到，也可以使用多个数据库来收集和存储在广告系统 120 中所使用的信息。

广告消费者接口 250 是与广告消费者 130 连接以获取或发送信息的组件。例如，广告消费者 130 可以向广告消费者接口 250 发送索取一个或多个广告的请求。请求可以包括诸如请求广告的站点、可用来帮助选择广告的任何信息、请求的廣告的数量等等信息。作为响应，广告消费者接口 250 可以向广告消费者 130 提供一个或多个广告。此外，广告消费者 130 还可以通过广告消费者接口 250 将有关广告的性能的信息发回到广告系统。例如，这可以包括上文参考数据库 240 所描述的统计信息。广告消费者接口组件 250 所需的数据或它所获取的数据驻留在数据库 240 中。

广告选择组件 260 接收对指定数量的广告的请求，以及帮助选择合适的广告的信息。例如，此信息可以包括最终用户指定的搜索查询。或者，或此外，如下面比较详细地描述的，此信息可以包括涉及正在请求其中的广告的页的内容的数据。在广告系统的一个实施例中，广告选择组件 260 可以执行诸如下列文件中所描述的操作：(i) 2002 年 9 月 24 日提出的 Jeffrey Dean、Georges Harik 和 Paul Bucheit 作为发明人的标题为“METHODS AND APPARATUS FOR SERVING RELEVANT ADVERTISEMENTS”的美国临时申请序列号 No. 60/413,536 (在此进行了引用)；(ii) 2002 年 12 月 6 日提出的 Jeffrey Dean、Georges Harik 和 Paul Bucheit 作为发明人的标题为“METHODS AND APPARATUS FOR SERVING RELEVANT ADVERTISEMENTS”的美国专利申请序列号 No. 10/314,427 (在此进

行了引用)； 和/或 (iii) 2003 年 2 月 26 日提出的 Darrell Anderson、Paul Bucheit、Alex Carobus、Claire Cui、Jeffrey Dean、Georges Harik、Deepak Jindal 和 Narayanan Shivakumar 作为发明人的标题为“SERVING ADVERTISEMENTS BASED ON CONTENT”的美国专利申请序列号 No. 10/____,____ (在此进行了引用)。

广告排序组件 270 从广告选择组件 260 接收相关广告的列表并确定应该将它们供应给最终用户的首选顺序(或诸如大小、位置、容量等等某些其他首选项属性。例如,可以基于与每一个广告关联的价值指示来将相关的广告排序。可以将这些排序的广告提供给广告供应组件 280。在广告系统的一个实施例中,广告排序组件 270 可以执行诸如下列文件中所描述的操作: (i) 2001 年 9 月 6 日提出的 Jane Manning、Salar Kamangar 和 Eric Veach 作为发明人的标题为“METHODS AND APPARATUS FOR ORDERING ADVERTISEMENTS”的美国临时申请序列号 No. 60/317,847(在此进行了引用); (ii) 2002 年 3 月 29 日提出的 Jane Manning、Salar Kamangar、Eric Veach 和 Lawrence Page 作为发明人的标题为“METHODS AND APPARATUS FOR ORDERING ADVERTISEMENTS BASED ON PERFORMANCE INFORMATION”的美国专利申请序列号 No. 10/112,656 (在此进行了引用); 和/或 (iii) 2002 年 3 月 29 日提出的 Salar Kamangar、Ross Koningstein 和 Eric Veach 作为发明人的标题为“METHODS AND APPARATUS FOR ORDERING ADVERTISEMENTS BASED ON PERFORMANCE INFORMATION AND PRICE INFORMATION”的美国专利申请序列号 No. 10/112,654 (在此进行了引用)。

广告供应组件 280 从广告排序组件 270 接收广告的排序列表,并将该列表格式化为适于供应给广告消费者 130 的方式。例如,这可以涉及将广告呈现为超文本标记语言 (HTML), 呈现为专有数据格

式等等。

统计信息引擎 290 包含与广告的选择和性能有关的信息。例如，统计信息引擎 290 可以记录广告消费者 130 作为广告请求的一部分提供的信息，广告选择组件 260 为该请求选择的广告、广告排序
5 组件 270 选择的顺序，以及广告供应组件 280 呈现的广告。此外，统计信息引擎 290 还可以记录有关一旦广告被提供给广告消费者 130 对广告发生了什么情况的信息。这包括诸如在什么位置提供了广告、向广告作出了什么响应、广告具有什么效果等等信息。

在广告系统的一个实施例中，可以使用优化操作（未显示）来优
10 化广告插播统计、收入或两者。这样的操作（上文所描述的其他组件中的某一个组件的操作）可以包括下列文件所描述的操作：(i) 2003 年 1 月 10 日提出的 Eric Veach 作为发明人的标题为“SETTING
MINIMUM PRICES IN AN ADVERTISING SYSTEM”的美国临时申
请序列号 No. No. 60/439,174（在此进行了引用）；(ii) 2002 年 1 月
15 24 日提出的 Jane Manning、Salar Kamangar、Eric Veach 和
Lawrence Page 作为发明人的标题为“DETERMINING A MINIMUM
PRICE”的美国专利申请序列号 No. 10/350,898（在此进行了引用）；
和/或 (iii) 2003 年 1 月 10 日提出的 Eric Veach 作为发明人的标
题为“GOVERNING THE SERVING OF ADVERTISEMENTS
20 BASED ON A COST TARGET”的美国专利申请序列号 No.
10/340,553（在此进行了引用）。

诸如上文参考图 1 和 2 所描述的典型系统或任何其他系统中
使用的那些在线广告可能具有各种特点。这样的特点可以由应用程序
和/或广告商来指定。这些特点在下面被称为“广告特点”。例如，在
25 文字广告的情况下，广告特点可以包括标题行、广告文字、可执行代
码、嵌入的链接等等。在图像广告的情况下，广告特点另外还可以包
括图像等等。取决于在线广告的类型，广告特点可以包括下列各项中
的一个或多个：文字、链接、音频文件、视频文件、图像文件、可执
行代码、嵌入的信息等等。

当供应在线广告时,可以使用一个或多个参数来描述如何、何时、和/或在哪儿供应广告。这些参数在下面被称为“供应参数”。供应参数可以包括下列各项中的一个或多个:在其上供应广告的页面的特点(包括该页上的信息)(包括被确定与页关联的一个或多个主题或概念,位于该页或该页内的信息或内容,有关诸如该页宿主(例如,AOL、Yahoo 等等)之类的信息,通过诸如流量、新鲜度、到该页或从该页发出的链接的数量和质量等等度量的页面的重要性,该页在目录结构内的位置等等),与广告的供应关联的搜索查询或搜索结果、用户特征(例如,他们的地理位置、他们使用的语言,所使用的浏览器的类型、前一页的视图、以前的行为)、启动向其作出响应而供应广告的请求的宿主或分支站点(例如,America Online、Google、Yahoo),广告在其中供应了该广告的页面上的绝对位置、广告相对于所供应的其他广告的位置(空间或时间),广告的大小,广告相对于其他广告的大小,广告的颜色,供应的其他广告的数量,供应的其他广告的类型,供应的日期,供应的星期,供应的年份等等。自然,还有其他可以在本发明的上下文中使用的供应参数。

虽然供应参数可能是非广告特点所固有的,但是,它们可能作为条件或约束与所述广告关联。当用作供应条件或约束时,这样的供应参数被简称为“供应约束”。供应约束还可以被称为“目标标准”。例如,在某些系统中,广告商也许能指定其广告只是在工作日供应,不低于某个位置,只向用户某个位置的用户供应等等。作为另一个示例,在某些系统下,广告商可以指定其广告只有在页面或搜索查询包括某些关键字或句子的情况下才供应。作为另一个示例,在某些系统中,广告商可以指定其广告只有在被供应的 Web 页面包括某些主题或概念或属于特定群集或某些其他类别的情况下才供应。

“广告信息”可以包括广告特点、广告供应约束、从广告特点或广告供应约束派生的信息(简称为“广告派生信息”),和/或涉及广告的信息(简称为“广告相关的信息”),以及这样的信息的扩展(例如,从广告相关的信息派生的信息)的任何组合。

“文档”在广义上可解释为包括任何机器可读的和机器可存储的工作成果。文档可以是文件、文件的组合、其中嵌入了指向其他文件的链接的一个或多个文件，等等；文件可以是任何类型，如文本、音频、图像、视频等等。将呈现给最终用户的文档的某些部分可以被视为文档的“内容”。文档中的广告插播可以通过嵌入的信息或指令来定义。在因特网的情况下，常见的文档是 Web 页面。Web 页面常常包括内容，并可以包括嵌入的信息（如元信息、超级链接等等）和/或嵌入的指令（如 Javascript 等等）。在很多情况下，文档具有唯一的、可寻址的存储位置，因此可以由此可寻址的位置唯一地标识。通用资源定位器 (URL) 是用于访问因特网上的信息的唯一地址。

“文档信息”可以包括文档中包括的任何信息，从文档中包括的信息派生的信息（简称为“文档派生信息”），和/或与信息相关的文档（简称为“文档相关的信息”），以及这样的信息的扩展（例如，从相关的信息派生的信息）。文档派生信息的一个示例是基于文档的文本内容的类别。文档相关的信息的示例包括具有指向即时文档的链接的其他文档中的文档信息，以及即时文档链接到其中的其他文档中的文档信息。

文档中的内容可以呈现在“内容呈现应用程序或设备”中。内容呈现应用程序的示例包括因特网浏览器（例如，Explorer 或 Netscape），媒体播放器（例如，MP3 播放器、Realnetworks 流式音频文件播放器等等）、查看器（例如，Adobe Acrobat pdf 阅读器）等等。

图 3 是显示可以在其中实现本发明的体系结构的图表。体系结构包括多个客户端设备 302、服务器设备 310，以及网络 301，该网络可以是因特网。每一个客户端设备 302 都包括连接到处理器 308 的计算机可读的介质 309，如随机存取存储器。处理器 308 执行存储器 309 中存储的程序指令。客户端设备 302 还可以包括许多其他外部或内部设备，如，但不限于，鼠标、CD-ROM、键盘和显示器。如此，正如那些精通本技术的人所理解的，客户端设备可以是

个人计算机、个人数字助理、移动电话、内容播放器等等。

通过客户端设备 302, 请求者 305 可以通过网络 301 彼此进行通信, 并与连接到网络 301 的其他系统和设备 (如服务器设备 310) 进行通信。请求者 305 可以是广告商 110、广告消费者 130 或广告目标 140。

类似于客户端设备 302, 服务器设备 310 可以包括连接到计算机可读的存储器 312 的处理器 311。服务器设备 310 还可以另外包括辅助存储元素, 如数据库 240。

客户端处理器 308 和服务器处理器 311 可以是许多著名的微处理器中的任何一个, 如位于 California, Santa Clara 的 Intel Corporation 的处理器。一般而言, 客户端设备 302 可以是连接到网络的并且与应用程序进行交互的任何类型的计算平台, 如数字助理或“智能”蜂窝电话或寻呼机。服务器 310, 虽然是作为单个计算机系统来描述的, 但是, 也可以作为计算机处理器的网络来实施。

存储器 312 可以包含许多程序, 如上文参考图 2 所描述的组件。

B. 典型的方法

图 4 是根据本发明的用于基于目标文档确定目标标准 (例如, 候选目标标准) 的典型方法 400 的流程图。为了在诸如图 2 中所描述的广告系统的背景下说明本发明, 可以理解, 这里所描述的方法 400 作为工具组件 220 的组成部分来发生, 虽然那些精通本技术的人可以认识到, 它不必在该组件 220 中或单独在该组件 220 中发生。

典型的方法 400 不受流程图中显示的顺序限制。方法 400 识别一个或多个目标文档。(方框 410) 目标文档可以是与广告链接的 Web 页面, 广告商有兴趣在其上显示广告的 Web 页, 或如前所述的任何其他文档。或者, 或此外, 目标文档可以基于广告信息以运算方式来确定。或者, 或此外, 目标文档可以由广告用户来指定。

概念或主题从与识别的目标文档关联的文档信息来提取和/或生成。(方框 420) 然后, 将提取的和/或生成的至少某些概念或主题

作为目标标准和/或潜在的目标标准与争论中的广告关联。（方框430）在一个实施例中，潜在的目标标准只不过是建议，但是这样的潜在的目标标准也可以用于自动地确定广告的目标。在一个实施例中，潜在的目标标准被用作目标标准，除非删除（例如，被广告商删除）。

5 在另一个实施例中，潜在的目标标准不被用作目标标准，除非被选择为包含（例如，被广告商选择为包含）。在又一个实施例中，第一潜在的目标标准集只用作目标标准，如果它（或该集的组分）被选择为包含，而使用另一个潜在的目标标准集，除非它（或该集的组分）被删除。虽然未显示，主题或概念还可以被用来生成或建议广告信息，

10 如，广告创意中的文本。

提取和/或生成概念或主题

识别对应于目标文档的主题的一种方式是通过分析目标文档内的某些或所有文本。可以确定，出现在目标文档的标题中的每一个检索词（字或片语）中的一个或多个对应于目标文档的主题。或者，可以

15 确定，出现在目标文档的正文中的每一个检索词中的一个或多个对应于目标文档的主题。在简单的情况下，目标文档内的每一个检索词都可以被确定为主题。一个稍微复杂的方法将是将检索词标识为主题，如果它出现在目标文档 N 次以上，如 $N = 2$ （实际上，每当对文本内的检索词进行分析时，都可以使用这样的基于阈值的方法）。甚至

20 可以执行更加复杂的分析，如通过使用目标文档的检索词向量，该向量给每一个检索词都分配了权重。例如，可以给频繁地出现在目标文档中的检索词分配一个比不是那么频繁出现的那些检索词相对较高的权重。

此外，可以调整权重，以给予不是那么频繁地出现在集合中的检索词较高的权重，如文档所属的集合或文档的通用集合。在给文本内的检索词分配了权重或分数的任何情况下，那些所产生的分数可以用来确定哪些检索词将被确定为目标文档的主题。例如，可以确定，只有分数靠前的检索词才能构成目标文档的主题。或者，或此外，可以

25 确定，上面 Z 个检索词（或其子集）将构成目标文档的主题， Z 是

某个定义的数字。或者，或此外，可以确定，具有超过 Y 的分数的检索词（或其子集）将构成目标文档的主题， Y 是某个定义的数字。如此，可以基于绝对和/或相对标准来确定主题。

5 或者，或除了使用目标文档内的文本或其他信息之外，还可以使用与目标文档关联的元-信息。在 Web 页面的情况下，这可以是锚定文字。可以使用一个或多个这样的简要描述来（象征性地）修改目标文档，方式是，用简要描述补充或替换某些或所有其内容。因此，可以从目标文档的标题和目标文档的简要描述的组合来标识主题。

10 或者，或除了从这些引用的简要描述之外，还可以使用引用本身。例如，从另一个文档到目标文档的引用可以用作两个文档相类似的指示。或者，或此外，从目标文档到另一个文档的引用可以用作两个文档相类似的指示。在此基础上，可以使用其他文档的内容（或元信息）来（象征性地）修改目标文档，方式是，通过用其他文档的内容来补充或替换其内容。然后，可以使用上文所描述的技术来对被修改的目标文档的内容进行分析，以识别一个或多个主题。

15 或者，或除了使用与目标文档关联的内容（也许包括元数据）之外，还可以使用其他技术识别目标文档的一个或多个主题。例如，可以确定，产生了对目标文档的引用的上面 N 个查询（或其子集）将构成目标文档的主题， N 是某个定义的数字。这些查询可以是，搜索引擎中的文字查询，这些查询产生了链接到目标文档或 Web 页面的结果。或者，或此外，可以使用其他类似的文档（例如，与目标文档相同的集合中的，与目标文档相同类别中的等等）的内容来（象征性地）修改目标文档的内容。然后，可以使用上文所描述的任何技术来分析目标文档，以识别一个或多个主题。在 Web 页面的情况下，25 这可以是存储在与目标 Web 页面在同一个主机上的相关页面的子目录内的其他 Web 页面。或者，或此外，可以使用用于将目标文档分类为一个或多个主题或类别集的任何技术。甚至还可以使用访问了目标文档（或目标 Web 页面）的一个或多个用户的搜索查询历史来识别目标文档的主题，在理论上，对在时间上靠近该搜索查询历史的目

标文档的访问表示，用户认为概念是相关的。

使用上文所描述的各种技术中的一个或多个，或其他技术，可以识别目标文档的一个或多个主题。一旦识别了这些主题，就可以使用各种各样的技术来确定与这些被识别的主题相关的其他主题。例如，
5 可以使用同义词库来确定与识别的主题密切相关的，或在概念上类似于识别的主题的其他主题（例如，同义词）。

为了清楚起见，前述的对“修改”目标文档的引用是象征性的，帮助理解在字面上不在目标文档内的其他信息的使用。那些精通本技术的人将认识到，不必实际修改目标文档，以利用此其他信息。

10 2002年10月3日提出的标题为“**Methods and Apparatus for Probabilistic Hierarchical Inferential Learner**”的美国临时申请序列号 No. 60/416,144（在此进行了引用）描述了根据本发明的原理确定可以使用的信息的一个或多个概念或主题的典型方式。

C. 典型设备

15 图5是可以实现上文讨论的一个或多个操作的机器500的高级别的方框图。机器500基本上包括一个或多个处理器510、一个或多个输入/输出接口单元530、一个或多个存储设备520，以及用于在连接的元件之间进行信息通信的一个或多个系统总线和/或网络540。一个或多个输入设备532和一个或多个输出设备534可以与
20 一个或多个输入/输出接口530连接。

一个或多个处理器510可以执行机器可执行的指令（例如，在位于 California, Palo Alto 的 Sun Microsystems 所提供的 Solaris 操作系统、诸如位于 North Carolina, Durham 的 Red Hat, Inc. 之类的许多供应商所提供的 Linux 操作系统上运行的 C 或 C++），
25 以实现本发明的一个或多个方面。至少一部分机器可执行的指令可以存储（临时或永久地）在一个或多个存储设备520上，和/或可以通过一个或多个输入接口单元530从外部源接收。

在一个实施例中，机器500可以是一个或多个常规个人计算机。在此情况下，处理单元510可以是一个或多个微处理器。总线540

可以包括系统总线。存储设备 520 可以包括诸如只读存储器 (ROM) 和/或随机存取存储器 (RAM) 之类的系统存储器。存储设备 520 还可以包括用于从硬盘中读取和向硬盘中写入的硬盘驱动器,用于从(例如,可移动)磁盘读取和向其中写入的磁盘驱动器,以及从诸如光盘或其他(磁性)光学介质之类的可移动(磁性)光盘读取和向其中写入的光盘驱动器。

用户可以通过诸如键盘和指示设备(例如,鼠标)向个人计算机中输入命令和信息。也可以包括(或作为备选)诸如麦克风、游戏杆、游戏板、碟形卫星天线、扫描仪等等其他输入设备。这些及其他输入设备常常通过连接到系统总线 540 的合适的接口 530 连接到处理单元 510。输出设备 534 可以包括监视器或其他类型的显示设备,这些设备也可以通过合适的接口连接到系统总线 540。除了监视器之外(或代替),个人计算机还可以包括其他(外围)输出设备(未显示),如扬声器和打印机。

15 D. 典型实施例的说明性操作

图 6 说明了在具有指向广告商的 Web 页面的链接的文字广告的情况下,在建议关键字及其他目标标准的背景下使用的本发明的典型实施例的操作。如图 6 所示,基于文字的广告 610 包括标题文字 620、正文文字 630,以及指向所希望的 Web 页面 660 的链接 640。一个或多个目标标准或供应约束 650 可以与基于文字的广告 610 关联。

如此示例中进一步说明的,从基于文字的广告 610 链接的 Web 页面 660 包括到其他 Web 页面 662 和 664 的其他链接(用虚弧线表示)。相关的查询 675,如返回了指向 Web 页面 660 的链接的过去的查询,和/或将会返回指向 Web 页面 660 的链接的潜在查询都可用(例如,从搜索引擎)。

在此示例中,从广告本身 610 识别了目标文档 660、662、664 (回想一下,方框 410)。在此示例中,从某些或所有信息 670,和/或广告 610 本身中的信息,提取和/或生成主题或概念(回想一下方

框 420。) 最后, 某些或所有这些概念或主题可以作为候选目标标准或供应约束 680 来提供(回想一下方框 430。), 如建议的关键字、地理位置供应约束、和/或时间供应约束。

E. 结束语

- 5 前面对本发明的优选实施例的描述提供了说明和描述, 但不是详尽的说明或将本发明的范围限于所说明的形式。根据上述原理, 可以进行各种修改, 也可以通过本发明的实践来进行修改。

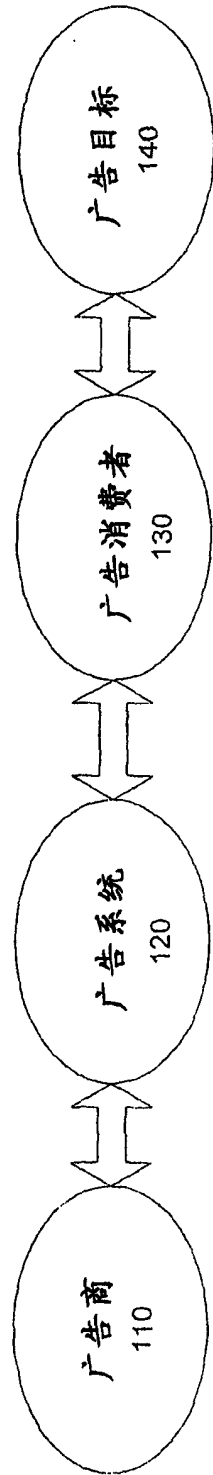


图1

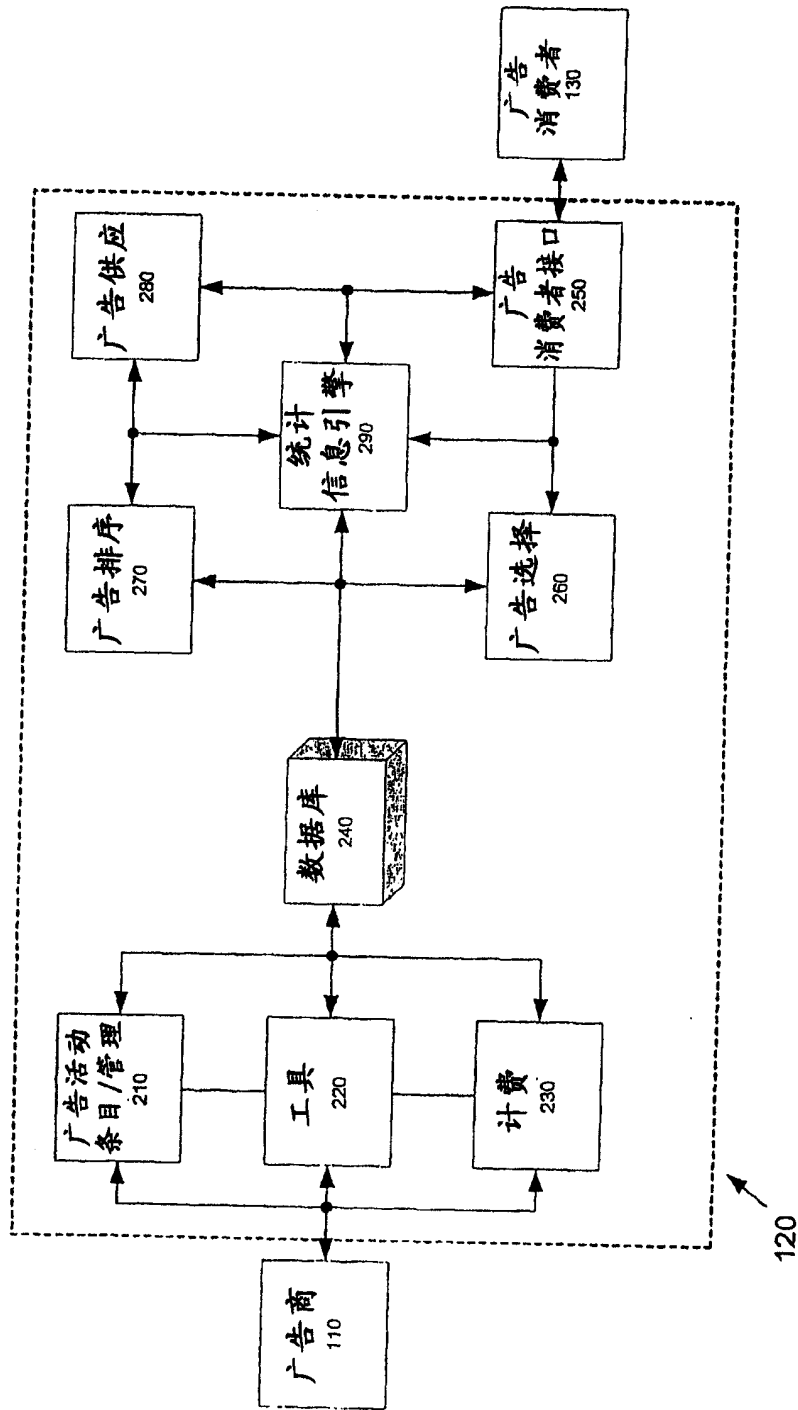


图2

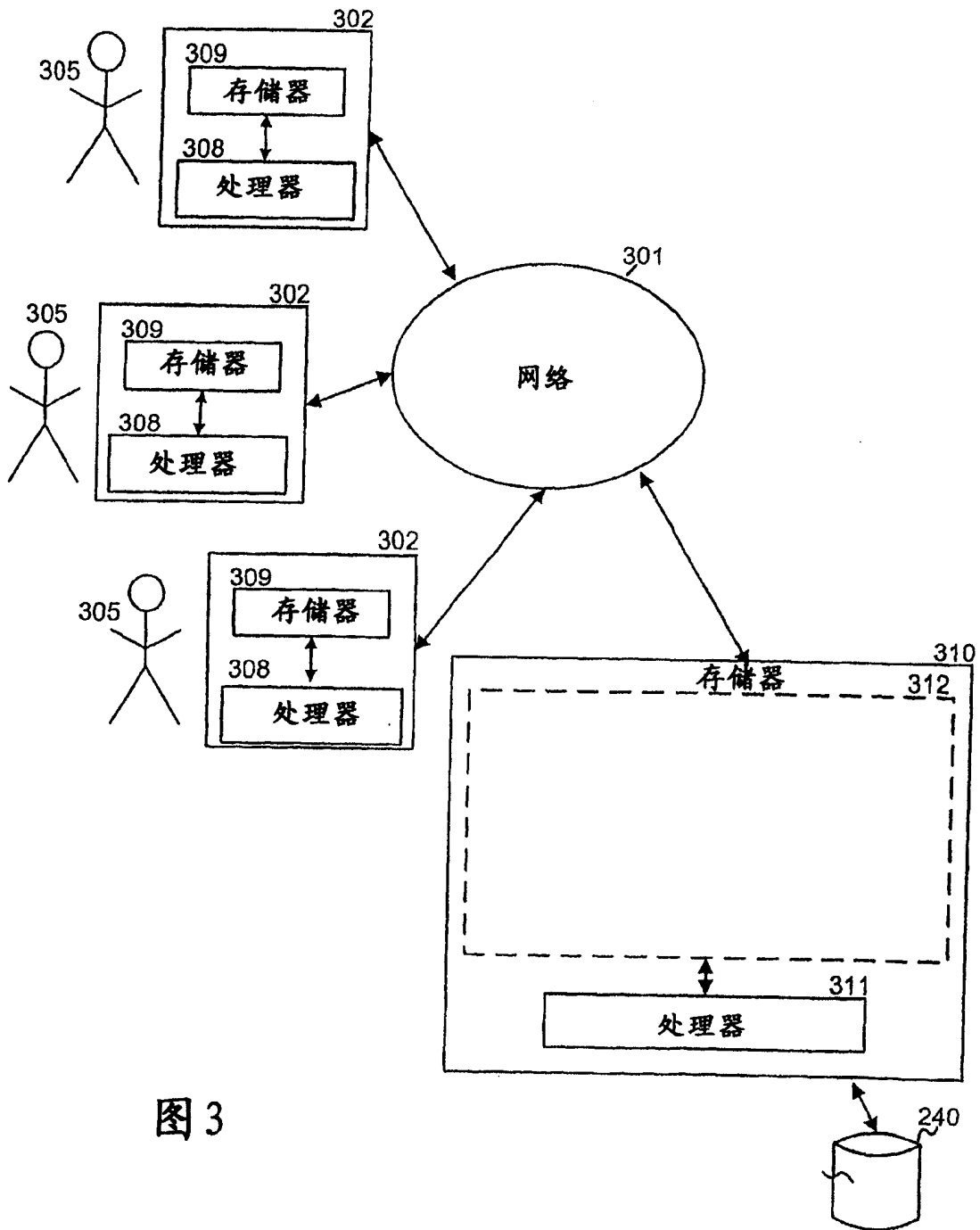


图 3

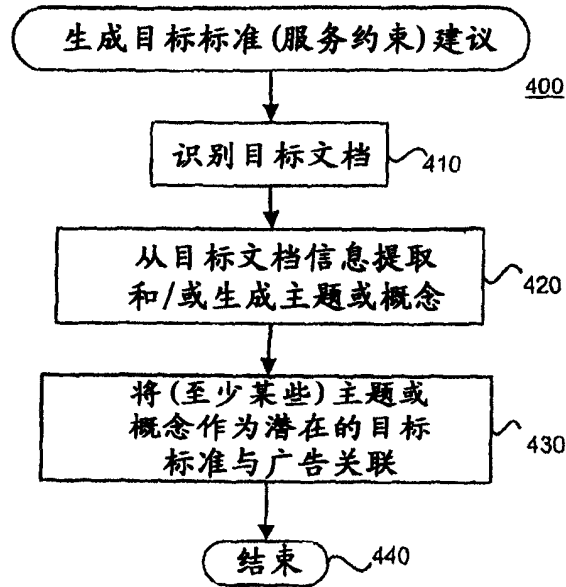


图 4

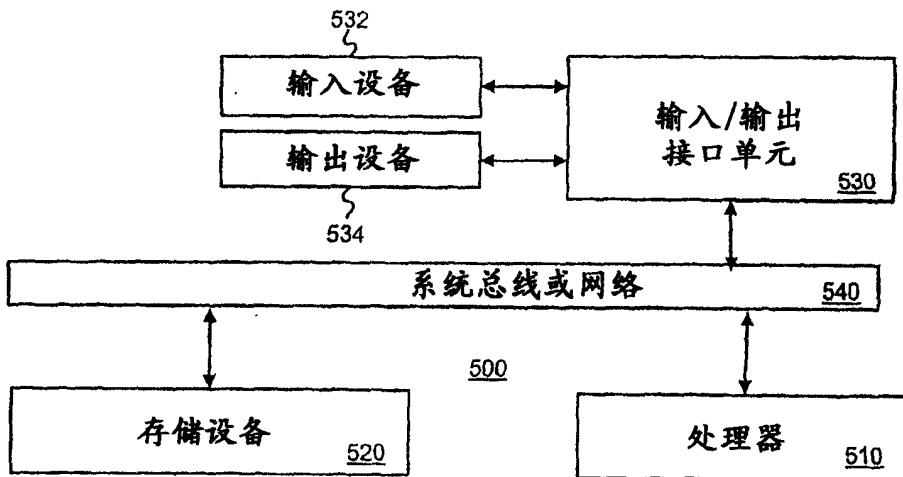


图 5

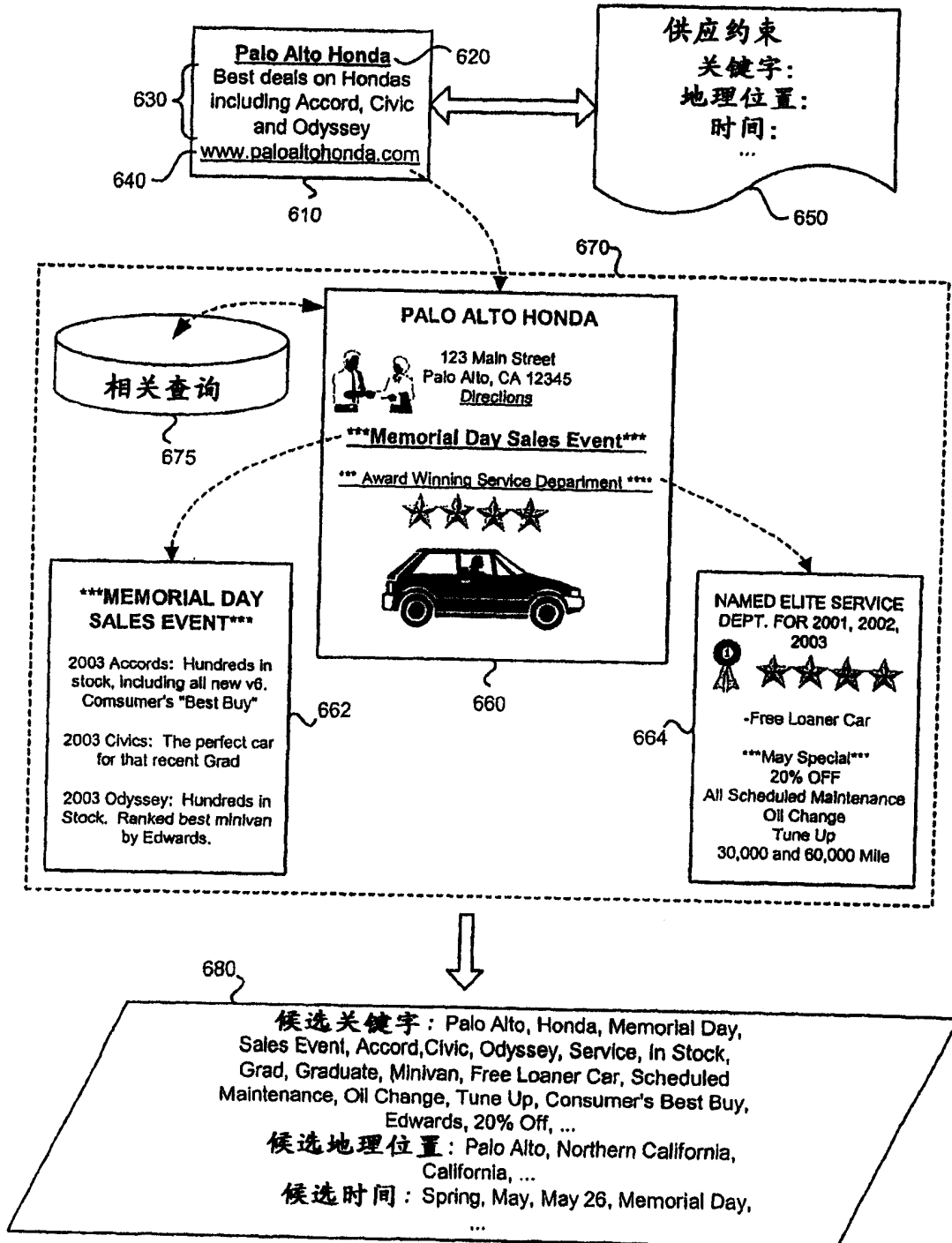


图6