



- (51) **International Patent Classification:**
G06Q 30/00 (2012.01)
- (21) **International Application Number:**
PCT/CN2013/083146
- (22) **International Filing Date:**
9 September 2013 (09.09.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (71) **Applicant:** **GOOGLE INC.** [US/US]; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US).
- (72) **Inventor; and**
- (71) **Applicant (for DM only):** **SHENG, Yong** [CN/US]; 333 8th Street South, Kirkland, Washington 98033 (US).
- (72) **Inventors:** **ASUNCION, Arthur**; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US). **KRAFCIK, Jonathan Michael**; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US). **ZHU, Bogong**; Tsinghua Science Park Bldg. 6, No. 1 Zhongguancun East Road, Haidian District, Beijing 100084 (CN). **JOHNSON, Nicholas**; 1600 Amphitheatre Parkway, Mountain View, California 94043 (US). **HUANG, Yan**; 8832 NE 148th Place, Kenmore, Washington 98082 (US).
- (74) **Agent:** **KING & WOOD MALLESONS**; 20th Floor, East Tower, World Financial Center, No.1 Dongsanhuan Zhonglu, Chaoyang District, Beijing 100020 (CN).
- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) **Title:** IDENTIFYING GROUPS FOR CONTENT ITEM DISTRIBUTION

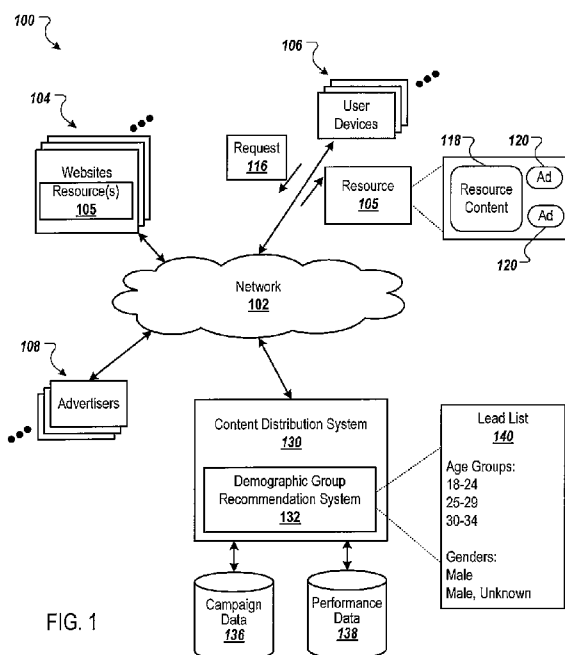


FIG. 1

(57) **Abstract:** Methods, systems, and apparatus, including computer programs encoded on a computer storage medium, for identifying groups for which to provide its members with content items. In one aspect, a method includes identifying, for a particular user action performed in response to presentation of a content item and based on performance data, an overall cost-per-action measure specifying a cost for obtaining the particular user action. The performance data can be grouped into demographic groups. The demographic groups can be arranged in a sequence. Subsets of demographic groups that each includes one or more contiguous demographic groups are identified. An improvement measure that specifies a difference between the overall cost-per-action measure for the particular user action and a subset cost-per-action measure for the particular user interaction can be determined for at least some of the subsets. A particular subset can be selected for providing the content item based on the improvement measures.



Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

IDENTIFYING GROUPS FOR CONTENT ITEM DISTRIBUTION

BACKGROUND

The Internet enables access to a wide variety of resources. For example, video, audio, web pages directed to particular subject matter, news articles, images, and other resources are accessible over the Internet. The wide variety of resources that are accessible over the Internet has enabled opportunities for content distributors to provide content items with resources that are requested by users. Content items are units of content (e.g., individual files or a set of files) that are presented in/with resources (e.g., web pages), for example, in response to a content item request that is initiated by code included in, or associated with, the resource. An advertisement is an example of a content item that advertisers can provide for presentation with particular resources, such as web pages and search results pages.

SUMMARY

This specification describes technologies relating to content item distribution.

In general, one innovative aspect of the subject matter described in this specification can be embodied in methods that include the actions of obtaining performance data for a content item campaign; identifying, for a particular user action performed in response to presentation of a content item and based on the performance data, an overall cost-per-action measure specifying a cost for obtaining the particular user action; grouping the performance data into demographic groups using demographic-based criteria for each demographic group and demographic data for users that received content items of the content item campaign; arranging the demographic groups in a sequence using the demographic-based criteria for the demographic groups; identifying subsets of the demographic groups that each include one or more contiguous demographic groups from the sequence of demographic groups; for each of at least some of the subsets: determining, based on the performance data for the subset, an improvement measure specifying a difference between the overall cost-per-action measure for the particular user action and a subset cost-per-action measure for the particular user interaction, the subset cost-per-action specifying a cost for obtaining the particular user interaction from a user in the demographic groups of the subset; and selecting a particular subset of demographic groups for providing the content item based at least on the determined improvement measures. Other

embodiments of this aspect include corresponding systems, apparatus, and computer programs, configured to perform the actions of the methods, encoded on computer storage devices.

These and other embodiments can each optionally include one or more of the following features. The demographic-based criteria for each demographic group can specify an age range for users that have received a content item of the content item campaign. Arranging the demographic groups in a sequence can include arranging the demographic groups in an ordinal sequence based on the age range for each demographic group. Each subset of demographic groups can include two or more demographic groups that are adjacent in the ordinal sequence.

The demographic-based criteria for each demographic group can specify a predicted gender for users that have received a content item of the campaign and a confidence that the users are the predicted gender. Arranging the demographic groups in a sequence can include arranging the demographic groups in a sequence based on the predicted gender. The particular user action can be a content item selection or a conversion.

Selecting a particular subset of demographic groups for providing the content item can include: identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold; identifying a conversion share for each candidate subset, the conversion share for a particular candidate subset being based on (a) a number of conversions resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of conversions for content items of the content item campaign; and selecting the particular subset of demographic groups based on the conversion shares for the candidate subsets.

Selecting a particular subset of demographic groups for providing the content item can include: identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold; identifying a selection share for each candidate subset, the selection share for a particular candidate subset being based on (a) a number of selections of content items of the content item campaign resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of selections for content items of the content item campaign; and selecting

the particular subset of demographic groups based on the selection shares for the candidate subsets.

5 Selecting a particular subset of demographic groups for providing the content item can include: identifying, for each subset of demographic groups, a spend share using (a) an amount of money spent in response to providing content item of the content item campaign to users identified as being members of a demographic group of the subset of demographic groups and (b) a total amount of money spent in response to providing content items of the content item campaign; identifying, as candidate subsets, each subset of demographic groups that has a spend share that is less than a spend share
10 threshold; and selecting the particular subset of demographic groups from the candidate subsets based on the determined improvements for the candidate subsets.

 Selecting a particular subset of demographic groups for providing the content item can include: identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold; identifying, for
15 each candidate subset, a spend share using (a) an amount of money spent in response to providing content item of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total amount of money spent in response to providing content items of the content item campaign; filtering from the candidate subsets each candidate subset that has a spend share that
20 meets a spend share threshold; identifying a conversion share for each remaining candidate subset, the conversion share for a particular candidate subset being based on (a) a number of conversions resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of conversions for content items of the content
25 item campaign; and selecting, from the remaining candidate subsets, the particular subset of demographic groups based on the conversion shares for the remaining candidate subsets.

 Particular embodiments of the subject matter described in this specification can be implemented so as to realize none, one, or more of the following advantages.
30 Performance measures, such as cost-per-click measures and/or cost-per-conversion measures, realized by distributing a content item to members of a subset of demographic groups (e.g., a contiguous subset of demographic groups) can be improved relative to distributing the content item to the population in general or other

demographic groups outside of the subset. The performance measures can be improved without limiting the audience for the content items by selecting subsets having at least a threshold selection share or at least a threshold conversion share for the content item.

The details of one or more embodiments of the subject matter described in this specification are set forth in the accompanying drawings and the description below. Other features, aspects, and advantages of the subject matter will become apparent from the description, the drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of an example environment in which a content distribution system distributes content.

FIG. 2 is a flow chart of an example process for identifying a subset of contiguous demographic groups for recommendation.

FIG. 3 is a flow chart of an example process for filtering candidate subsets of demographic groups.

FIG. 4 depicts a graph of improvement measures and action shares for subsets of demographic groups.

FIG. 5 is block diagram of an example computer system.

Like reference numbers and designations in the various drawings indicate like elements.

DETAILED DESCRIPTION

A system can identify and recommend a set of demographic groups (e.g., one, two, or more different demographic groups) for providing content items (e.g., advertisements, videos, images, etc.) that improves cost per action measures (e.g., cost per click or cost per conversion measures), while minimizing restrictions on audience size (content item inventory). For example, a content item provider may be interested in identifying age group(s) (e.g., from multiple different predefined age groups) to provide content items such that the cost-per-action for a particular user action (e.g., content item selection or conversion) is reduced while still providing the content item to a sizeable audience. The system can organize the age groups into a sequence (e.g., an ordinal sequence) and compute performance measures and optionally other metrics for each contiguous subset of two or more age groups. Based on the computed performance measures and/or other metrics, the system can identify a particular subset

of the different age groups, and suggest that content items be distributed based on the subset. For example, the subset can be included in distribution parameters for the content items and/or bids associated with the distribution parameters can be modified in a manner that causes the content items to be selected more often to be provided to users
5 within the particular subset of age groups relative to other age groups. A similar process can be used for other demographic groups, such as gender groups, income-based groups, location-based groups, etc.

In some implementations, the system can identify candidate subsets of contiguous demographic groups for a content item campaign based on an improvement
10 measure for each subset. The improvement measure for a subset of demographic groups may be based on an improvement (or a difference) in a cost-per-action measure (e.g., cost-per-click measure or cost-per-conversion measure) for the subset relative to an overall cost-per-action measure for the campaign. For example, the cost-per-action for a particular age group (or other demographic group) may be 15% less than the
15 overall cost-per-action for the entire campaign. The system may identify, as candidate subsets, those subsets that have an improvement measure that meets an improvement threshold (e.g., 5%, 10%, 20%, or another number).

If there are no candidate subsets that have an improvement measure that meets the improvement threshold, the system may not identify a subset that is recommended
20 for providing content items. If there is a single candidate subset that has an improvement measure that meets the improvement threshold, the system may recommend that candidate subset. If there are multiple candidate subsets that have an improvement measure that meets the improvement threshold, the system may select the candidate subset that has the highest (or a threshold) conversion or selection share
25 (described below), or evaluate other metrics, such as spend share for the candidate subsets.

For example, the system may filter from the candidate subsets, those subsets that have a spend share that exceeds a spend share threshold. The spend share for a candidate subset of demographic groups can be based on an amount spent related to
30 providing content items of a content item campaign to users identified as being members of a demographic group of the candidate subset and the total money spent for the campaign. For example, the spend share may be the ratio between the amount of money spent for the candidate subset and the total money spent for the campaign. This

filter can be used to ensure that providing content items to the demographic groups of the selected subset will make a practical difference. For example, recommending a demographic group whose spend share is already high (e.g., 99%) is unlikely to significantly improve performance measures for the campaign.

5 The system may also filter from the candidate subsets, those subsets that have a selection share or a conversion share that does not meet a threshold selection or conversion share threshold. The conversion share for a candidate subset of demographic groups can be based on a number of conversions that have resulted from providing content items of the campaign to users that have been identified as being
10 members of a demographic group of the candidate subset and a total number of conversions for the campaign. Similarly, the selection share for a candidate subset of demographic groups can be based on a number of content item selections (e.g., user selections) that have resulted from providing content items of the campaign to users that have been identified as being members of a demographic group of the candidate
15 subset and a total number of content item selections (e.g., by users) for the campaign. This filtering helps ensure that demographic groups with low conversion rates or low selection rates are not selected for recommendation.

For situations in which the systems described here collect information about users, or may make use of information about users, the users may be provided with an
20 opportunity to control whether programs or features collect user information (e.g., information about a user's social network, social actions or activities, profession, a user's preferences, or a user's current location), or to control whether and/or how to receive content from the content server that may be more relevant to the user. In addition, certain data may be treated in one or more ways before it is stored or used, so
25 that certain information about the user is removed. For example, a user's identity may be treated so that no identifying information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, users may have control over how information is collected about
30 them and used by a content server.

FIG. 1 is a block diagram of an example environment 100 in which a content distribution system 130 distributes content items to user devices 106. The example environment 100 includes a network 102 such as a local area network (LAN), wide area

network (WAN), the Internet, or a combination thereof. The network 102 connects websites 104, user devices 106, advertisers 108, and the content distribution system 130. The example environment 100 may include millions of websites 104, user devices 106, and advertisers 108.

5 A website 104 is one or more resources 105 associated with a domain name and hosted by one or more servers. An example website is a collection of web pages formatted in hypertext markup language (HTML) that can contain text, images, multimedia content, and programming elements, e.g., scripts. Each website 104 is maintained by a publisher, e.g., an entity that manages and/or owns the website 104.

10 A resource 105 is data provided by the website 104 over the network 102 and that is associated with a resource address. Resources include HTML pages, word processing documents, and portable document format (PDF) documents, images, video, and feed sources, to name only a few. The resources can include content 118, e.g., words, phrases, images and sounds that may include embedded information (such as
15 meta-information in hyperlinks) and/or embedded instructions (such as scripts).

 A user device 106 is an electronic device that is capable of requesting and receiving resources over the network 102. Example user devices 106 include personal computers, mobile communication devices, and other devices that can send and receive data over the network 102. A user device 106 typically includes a user application,
20 such as a web browser, to facilitate the sending and receiving of data over the network 102.

 A user device 106 can request resources 105 from a website 104. In turn, data representing the resource 105 can be provided to the user device 106 for presentation by the user device 106. The data representing the resource 105 can include resource
25 content 118 (e.g., text, images, videos, etc. of the resource 105) and content item slots 120 (e.g., advertisement slots). When a resource 105 having a content item slot 120 is requested by a user device 106, the content distribution system 130 receives a content item request 116 requesting content items to be provided with the resource content 118.

 A content item request 116 can include data regarding the content item slots 120
30 (e.g. size or type of content item slot), data regarding the resource 105 on which the content item will be presented (e.g., category or keywords found on the resource, data regarding publisher of resource, etc.), and/or other data. If the content items are to be

presented in content item slots 120 of a search results page, the content item request 116 may include keywords of a search query submitted to a search system.

The content distribution system 130 allows advertisers 108 or other content item providers to define campaign rules that take into account attributes of content item slots and resources on which content items (e.g., advertisements, videos, images, etc.) are to be presented. Example campaign rules include keyword rules, in which an advertiser 108 provides bids for keywords that are present in either search queries or resource content 118. A bid represents a value that an advertiser 108 is willing to pay in response to a presentation of the content item (e.g., cost-per-impression bid), an interaction with the content item (e.g., cost-per-click bid), or a conversion that occurs in response to providing the content item (e.g., cost-per-conversion bid). Content items that are associated with keywords having bids that result in a content item slot 120 being awarded in response to an auction (or another selection process) are selected for displaying in the content item slots 120.

When a user of a user device 106 selects an advertisement, the user device 106 generates a request for a landing page of the advertisement, which is typically a web page of the advertiser 108. For example, the advertisers 108 may each have respective web pages, some of which are landing pages for the advertisements of the advertisers 108.

The content distribution system 130 includes a data storage system that stores campaign data 136, and performance data 138. The campaign data 136 stores content items (e.g., advertisements), campaign information, bid values for content items, and/or budgeting information for advertisers 108. The performance data 138 stores data indicating the performance of the content items that are served. Such performance data can include, for example, a number of user selections (e.g., clicks) and click-through rates for content items, the number of impressions for content items, the number of conversions and conversion rates for content items, and/or spend data.

A click-through may occur, for example, when a user device, e.g., user device 106, interacts with or “clicks” on a content item, and the user interaction or click of a content item can be referred to as a user selection of the content item. The click-through rate for a content item can be a performance measure that is obtained by dividing the number of times users clicked on the content item or a link associated with the content item by a number of times the content item was delivered to a user device. For

example, if a content item is delivered 100 times, and the content item receives three clicks (or other user selections), the click-through rate for the content item is 3%.

A “conversion” occurs when a user, for example, consummates a transaction related to a previously served content item. What constitutes a conversion may vary from case to case and can be determined in a variety of ways. For example, a conversion may occur when a user of a user device 106 selects a content item, is referred to a web page of the content item provider (e.g., advertiser 108), such as an advertisement landing page, and consummate a purchase before leaving that web page. Other conversion types can also be used, and user actions that constitute the different types of conversions can be specified by content item providers. A conversion rate for a content item can, for example, be defined as the ratio of the number of conversions of the content item to the number of impressions of the content item (e.g., the number of times the content item has been rendered or delivered), or the ratio of the number of conversions to the number of user selections.

The spend amounts can include, for each user selection or conversion of a content item, an amount the content provider is charged for the user selection or conversion. In some implementations, the content item provider is charged an amount based on a cost-per-click bid or cost-per-conversion bid provided by the content item provider in response to a content item selection or conversion, respectively. In some implementations, the content item provider is charged an amount based on a next highest cost-per-click bid or cost-per-conversion bid in response to a content item selection or conversion, respectively. For example, the content item provider may be charged an amount above the next highest cost-per-click or cost-per-conversion bid (e.g., the content item provider’s bid is the highest, but is charged an amount above the second highest bid). The cost-per-click and cost-per-conversion amounts charged (or to be charged) to the content item provider can be stored for each content item selection and for each conversion.

The spend data can also include total amounts spent for content items and/or total amounts spent for content item campaigns that can be used to distribute multiple different content items based on various campaign rules. The total amount spent for a content item may include costs for each type of bid. For example, the total amount spent for a content item may include costs charged to the content item provider for

impressions, user selections, and conversions. The total cost for a campaign may be the sum of the total amount spent for each content item in the content item campaign.

The performance data 138 can also include, or be used to determine, cost-per-action measures. A cost-per-action is an amount paid when a content item causes a particular user action. Example cost-per-action measures include cost-per-click measures and a cost-per conversion measures. A cost-per-click measure for a content item can represent an average (or mean, median, or other measure of central tendency) amount that the content item provider is charged for user selections of the content item. For example, a content item provider may provide a cost-per-click bid that the content item provider is willing to pay an entity (e.g., an entity that operates the content distribution system 130) for each user selection of the content item. When the content item is provided and selected (e.g., by a user), the content item provider may be charged an amount based on the cost-per-click bid (e.g., the bid amount or amount above a next highest bid amount). In addition, the content item provider may provide a different cost-per-click bid for different contexts. For example, the content item provider may be willing to pay more for a user selection of a sports-related advertisement when the advertisement is provided with a sports-related web page than if the advertisement is provided with a web page directed to gardening. Thus, the cost for a user selection of the content item may vary for each user selection. The cost-per-click measure can represent an average cost-per-click for the content item for all impressions of the content item or for certain groups of impressions.

Similarly, a cost-per-conversion measure can represent an average (or mean, median, or other measure of central tendency) amount that the content item provider is charged for each conversion that is attributed to the content item. For example, a content item provider may provide a cost-per-conversion bid that the content item provider is willing to pay an entity (e.g., an entity that operates the content distribution system 130) for each conversion that occurs following user interaction with the content item. When a conversion occurs in response to providing the content item, the content item provider may be charged an amount based on the cost-per-conversion bid (e.g., the bid amount or amount above a next highest bid amount). Similar to costs paid for user selections, the amount paid for a conversion can vary for each conversion. The cost-per-conversion measure can represent an average cost-per-conversion for the content

item for all impressions of the content item or for certain groups of impressions (e.g., impressions for a particular demographic group).

The stored performance data 138 can include overall performance data for each content item and/or each content item campaign. For example, the performance data
5 138 may include an overall click-through rate, an overall number of impressions, and/or an overall conversion rate for each content item and for each content item campaign. The performance data 138 can also include an overall cost-per-click measure and overall cost-per-conversion measure for each content item and for each content item campaign.

10 In addition, the performance data 138 for each content item and/or each content item campaign can be delineated based on demographic groups. For example, demographic data for a user that selects a content item or takes an action that constitutes a conversion for the content item can be identified or estimated and stored with (or with a reference to) data regarding the user selection or conversion. Using the
15 demographic data, the content distribution system 130 can organize the performance data 138 into demographic groups, such as a set of predefined demographic groups. Each demographic group can be based on a particular demographic criterion, such as age, gender, income, geographic location, or some other data that can be used to segment users.

20 Example age-based demographic groups may each include a range of ages. For example, there may be a group for users that are 18-24, another group for users that are 25-30, another group for users that are 31-35, and so on. Groups based on income or other numerical-based criteria can be divided into groups similarly.

Non-numerical demographics, such as gender, can be based on a confidence in
25 the demographic data for the users. For example, there may be a male group for performance data for which there is a high confidence (e.g., greater than a confidence threshold) that the user that selected the content item or completed a conversion event is a male. Similarly, there may be a female group for performance data for which there is a high confidence (e.g., greater than a confidence threshold) that the user that
30 selected the content item or completed a conversion event is a female. There may also be a “male, unknown” group for performance data for which the user is predicted to be male, but the confidence is lower (e.g., lower than a confidence threshold). Similarly, there may be a “female, unknown” group for performance data for which the user is

predicted to be female, but the confidence is lower (e.g., lower than a confidence threshold).

In some implementations, the demographic groups may be predefined for all content items and/or content item campaigns, for sets of content items and/or content item campaigns, or for each content item and/or content item campaign. For example, the demographic groups may be defined prior to the start of a content item campaign. In this example, the performance data collected for the campaign may be organized into the predefined demographic groups based on the criteria for the groups (e.g., age range, gender, etc.) and demographic data for the users that interact with the content items.

10 The content distribution system 130 also includes a demographic group recommendation system 132. The demographic group recommendation system 132 can identify one or more subsets of demographic groups to recommend as an audience for a content item and/or for a content item campaign. As described in more detail below, the demographic group recommendation system 132 can identify the groups based on performance data for the groups. Once identified, the demographic group recommendation system 132 can generate a lead list 140 that specifies the recommended subset(s) of demographic groups. The example lead list 140 specifies the age groups 18-24, 25-29, and 30-34. The lead list 140 also identifies gender categories male and female, unknown. The demographic group recommendation system 20 132 can provide the lead list 140 to the content item provider. In turn, the content item provider may provide more content items to the recommended groups in order to improve performance measures for the content items. In an auction-based system, the content distribution system 130 may modify bids for the content item provider's content items such that the content items are provided more often to the recommended demographic groups. For example, the content distribution system 130 may increase the bids for the content items in response to a prediction or determination that the user receiving the content item is a member of a recommended demographic group. Consider, for example, the lead list 140. In this example, the content distribution system 130 may increase the content item's bid for a user identified as being between 25 18 and 34 years old as these ages have been identified for recommendation. In addition, the bids for providing the content item to users identified as belonging to another age group (e.g., 65+), may be lower as these age groups have not been identified for recommendation.

In some implementations, the demographic group recommendation system 132 identifies subsets of contiguous demographic groups for recommendation. For example, a recommendation that identifies the age groups “18-24,” “40-45,” and “65-70” may not be as useful to a content item provider as a recommendation that identifies
5 contiguous groups as it may be more difficult to adjust a campaign to increase the distribution to these separated groups. Thus, the demographic recommendation system 132 can identify, for recommendation, subsets of contiguous demographic groups based on performance measures for the subsets, as these contiguous demographic groups may be more likely to respond to similar content items and the resulting cost-per-action may
10 be better (e.g., lower) relative to the population as a whole or other demographic groups.

FIG. 2 is a flow chart of an example process 200 for identifying a subset of contiguous demographic groups for recommendation. Operations of the process 200 can be implemented, for example, by a data processing apparatus, such as the demographic group recommendation system 132 of FIG. 1. The process 200 can also
15 be implemented as instructions stored on computer storage medium, and execution of the instructions by a data processing apparatus can cause the data processing apparatus to perform the operations of this process 200.

Performance data is obtained for a content item campaign (202). The content item campaign may be associated with a content item provider and can include one or
20 more content items, such as advertisements, that are provided to users. The performance data for the content item campaign can include data such as, a number of user selections of the content item(s) of the campaign, a number of conversions attributed to the content item(s) of the campaign, cost-per-click measures, cost-per-conversion measures, campaign and content item spend data, and optionally other
25 performance data.

An overall cost-per-action measure for a particular user action is identified for the content item campaign (204). The particular user action may be an interaction with a content item of the campaign (e.g., a user selection of the content item), or a conversion that occurs in response to providing a content item of the campaign (e.g., a
30 purchase of a product). Thus, the overall cost-per-action measure may be an overall cost-per-click measure or an overall cost-per-conversion measure. In some implementations, as described in more detail below, both example types of cost-per-action measures can be used.

The overall cost-per-action measure may be based on the costs for the particular action for each of the one or more content items of the campaign. For example, the overall cost-per-action measure may be an average (or other measure of central tendency, such as a median) cost for obtaining the particular user action across all content items of the campaign.

The performance data is grouped into demographic groups (206). The performance data may be grouped into predefined demographic groups based on demographic-based criteria for each demographic group and demographic data for users that received content items of the content item campaign. For example, the performance data may be grouped into predefined age-based groups based on age range criteria (e.g., ages 18-24) for each demographic group. In this example, performance data regarding a content item when the content item is presented to a user that is identified or predicted to be between the ages of 18 and 24 (e.g., 20) may be grouped in the demographic group for users in the 18-24 age range. Similarly, performance data regarding the same content item when the content item is presented to a user that is identified or predicted to be between the ages of 25 and 29 (e.g., 26) may be grouped in the demographic group for users in the 25-29 age range.

In some implementations, the performance data is grouped into demographic groups based on a single type of demographic criteria. For example, the groups can be based on one of age, gender, or income, but not a combination of two or more criteria. That is, there may not be groups based on age and gender, in some implementations. For such implementations, the process 200 may be executed individually for each type of demographic data. In some implementations, the performance data can be grouped based on a combination of two or more demographic criteria.

The demographic groups are arranged in a sequence using the demographic-based criteria for the demographic groups (208). For number-based demographic groups, such as age or income groups, the demographic groups may be arranged in an ordinal sequence. For example, age-based groups may be arranged in an ascending sequence from the group having a particular range (e.g., 18-24) to a group having a different range (e.g., 65+). Similarly, income-based groups may be arranged in an ascending sequence from the group having the lowest income range (e.g., \$0 - \$15K per year) to the group having the highest income range (\$200K+ per year).

For demographic-based groups that are not based on numbers, the demographic groups may be arranged in a sequence that is appropriate for the demographic-based criteria. For gender-based groups, the groups may be arranged based on predicted gender. For example, the gender-based demographic groups may be arranged in the following sequence: Male -> Female -> Male, Unknown -> Female, Unknown -> Male, Female. Alternatively, the gender-based groups may be arranged in the following sequence: Male -> Male, Unknown -> Male, Female -> Female, Unknown -> Female. In these example sequences, the group “Male” may be for performance data for users identified as being male with high confidence (e.g., greater than a confidence threshold); the group “Female” may be for performance data for users identified as being female with high confidence (e.g., greater than a confidence threshold) the group “Male, Unknown” may be for performance data for users identified as being male with low confidence (e.g., less than a confidence threshold); the group “Female, Unknown” may be for performance data for users identified as being female with low confidence (e.g., less than a confidence threshold); and the user “Male, Female” may be for performance data for users in which the gender could not be identified, or could only be identified with little confidence (e.g., less than a threshold).

Subsets of demographic groups are identified (210). Each identified subset can include two or more contiguous (or adjacent) demographic groups from the sequence. For example, consider age-based groups. Each identified subset can include at least two contiguous age groups. For example, a subset may include the age group “18-24” and the age group “25-30” as the two contiguous age groups. Another subset may include the age groups “18-24,” “25-30,” and “31-35.”

In some implementations, each identified subset may only include contiguous groups. For example, an identified subset may include the age groups “18-24,” “25-30,” and “31-35,” but a subset of “18-24,” “25-30,” and “35-40” may not be eligible for recommendation as the age groups “25-30,” and “35-40” are not contiguous. Each possible subset of demographic groups that include only contiguous demographic groups may be identified, in some implementations.

In some implementations, subsets of demographic groups that include non-contiguous demographic groups may be eligible for recommendation. For example, subsets that are the complements of a sequence of contiguous demographic groups may be eligible. To illustrate, consider the subset that includes the age groups “25-34,” “35-

44,” and “45-54.” A subset of age groups that is the complement of this group may be the age groups “18-24,” “55-64,” and “65+.” Although the age groups of the complement subset are not contiguous, the complement subset may be eligible because the age groups of the original set (25-34, 35-44, and 45-54) are contiguous. In some
 5 implementations, subsets of contiguous age groups and their complements may be eligible for recommendation and thus, identified.

For each of at least some of the identified subsets of demographic groups, an improvement measure is determined (212). The improvement measure for a subset specifies a difference between the overall cost-per-action measure for the particular
 10 user action (e.g., selection or conversion) and a subset cost-per-action measure for the particular user action. Whereas the overall cost-per-action measure specifies an overall cost for obtaining the particular user action for the content item campaign (e.g., irrespective of the demographics of the users that performed the particular action), the subset cost-per-action measure can specify a cost for obtaining the particular user
 15 action from users identified in the demographic groups of the subset.

For example, the subset cost-per-action measure may be the average (or other measure of central tendency) cost for obtaining the action for each occurrence of the action by a user identified as being a member of a demographic group of the subset. As an example, consider a subset that includes the age groups “18-24” and “25-30.” In
 20 this example, people between 25-30 years old select a content item of the campaign and the content item provider is charged \$1 for each selection on average. In addition, people between 18-24 years old select a content item of the campaign and the content item provider is charged \$2 for each selection on average. The subset cost-per-action measure (in this example cost-per-selection) may be \$1.50 (assuming these are the only
 25 two selections by users in the age group of the subset and the average cost is used).

The improvement measure for a subset of demographic groups is based on a difference between the subset cost-per-action measure and the overall cost-per-action measure for the content item campaign. For example, the improvement measure for a subset may be determined using Relationship (1) below:

30

$$\text{Relationship: } \textit{Improvement measure} = \frac{\textit{Subset cost-per-action measure}}{\textit{Overall cost-per-action measure}} - 1$$

(1)

In this example, negative improvement measures represent an improvement in the subset cost-per-action measure for the subset relative to the overall cost-per-action measure for the content item campaign. Further, negative improvement measures having a higher absolute value represent more improvement over the overall cost-per-action measure than a negative improvement measure that has a lower absolute value. Positive improvement measures represent an increase in cost-per-action for the subset relative to the campaign.

The subsets can be filtered based on one or more filter criteria (214). This filtering is optional and can be used to filter subsets that do not result in sufficient improvement in the cost-per-action measure or that have another insufficient performance measure, such as an insufficient spend share, conversion share, or selection share. The filtering can also be based on a confidence measure associated with each improvement measure. For example, the confidence measure for an improvement measure may indicate a level confidence that the improvement measure is accurate and can be determined based on the amount of data available for the improvement measure. Subsets with an improvement measure that does not meet a confidence threshold may be filtered from consideration. An example process for filtering subsets of demographic groups is illustrated in FIG. 3 and described below.

A particular subset of demographic groups is selected (216). The particular subset may be selected in a variety of ways. For example, the improvement measure may be used to filter the subsets to those having a sufficient improvement measure (e.g., those having a negative improvement measure whose absolute value meets an improvement threshold). If there are no subsets that have an improvement measure that meets the improvement threshold, a particular subset may not be selected for recommendation or to provide content items. If there is a single subset that has an improvement measure that meets the improvement threshold, that subset may be recommended. If there are multiple subsets that have an improvement measure that meets the improvement threshold, the subset that has the highest (or threshold) action share (e.g., highest conversion share or highest selection share) may be selected.

The conversion share for a subset of demographic groups can be based on a number of conversions resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the subset

and a total number of conversions for content items of the campaign. For example, the conversion share for the subset may be a ratio between the number of conversions for members of the subset and the total number of conversions for content items of the campaign.

5 Similarly, the selection share for a subset of demographic groups can be based on a number of selections of content items of the content item campaign made by users identified as being members of a demographic group of the subset and a total number of selections of content items of the campaign. For example, the selection share for the subset may be a ratio between the number of user selections for members of the subset
10 and the total number of user selections of content items of the campaign.

 In some implementations, if the cost-per-action measure is based on a cost per-selection, then the remaining subset (after filtering) that has the highest selection share may be selected. Similarly, if the cost-per-action measure is based on a cost per-conversion, then the remaining subset that has the highest conversion share may be
15 selected.

 The demographic groups of the selected subset may be recommended to the content item provider so that the content provider may focus or increase its distribution of content items to the recommended demographic groups. In some implementations, the content distribution system 130 may modify bids for the content item campaign
20 such that members of the demographic groups of the selected subset receive more content items of the campaign than unselected demographic groups.

 In some implementations, the process 200 is executed using a cost-per-conversion measure as the cost-per action measure. For example, if the total number of conversions is less than a conversion threshold, then the process 200 can be executed
25 using a cost-per-click measure as the cost-per-action measure. In addition, or in the alternative, if there are no subsets that have an improvement measure that meets the improvement threshold using the cost-per-conversion measure as the cost-per-action measure, then the process 200 can be re-executed using the cost-per-click measure as the cost-per-action measure.

30 For example, the demographic group recommendation system 132 may first execute the process 200 using a cost-per-conversion measure as the cost-per-action measure. If the demographic group recommendation system 132 does not identify a contiguous subset of demographic groups that results in a sufficient improvement

measure (e.g., a negative improvement measure whose absolute value meets an improvement threshold), the demographic group recommendation system 132 may re-execute the process 200 using a cost-per-click measure as the cost-per-action measure. If this execution of the process 200 does not result in the identification of a subset of demographic groups that meet the improvement threshold, the demographic group recommendation system 132 may not recommend any demographic groups.

As described above, the demographic group recommendation system 132 can execute the process 200 for each type of demographic group to identify a subset of contiguous demographic groups of that type for recommendation. For example, the demographic group recommendation system 132 can execute the process 200 to identify a subset of age groups to recommend, and execute the process 200 again to identify subsets of genders to recommend.

In some implementations, the demographic group recommendation system 132 can identify groups based on multiple demographic criteria. For example, the demographic group recommendation system 132 may identify a group based on gender, age, and geographic location. To do this, the demographic group system 132 may identify the possible contiguous subsets for each type of demographic. Continuing the previous example, the demographic group recommendation system 132 may identify the possible subsets of contiguous age groups, the possible subsets of contiguous gender groups, and the possible subsets of contiguous location groups. The demographic group recommendation system 132 can then go through the cross-product space of gender x age x location and determine the cost-per-action improvement measure and other performance measures (e.g., conversion share, selection share, and/or spend share) for each eligible meta-group. The demographic group recommendation system 132 can then filter the meta-groups using spend share, selection share, and/or conversion share and select from any remaining groups as described above with reference to FIG 2.

FIG. 3 is a flow chart of an example process 300 for filtering candidate subsets of demographic groups. Operations of the process 300 can be implemented, for example, by a data processing apparatus, such as the demographic group recommendation system 132 of FIG. 1. The process 300 can also be implemented as instructions stored on computer storage medium, and execution of the instructions by a

data processing apparatus can cause the data processing apparatus to perform the operations of this process 300.

Subsets of demographic groups are identified for a content item campaign (302). Each subset of demographic groups may include contiguous demographic groups. For each subset of demographic groups, an improvement measure is identified. For example, the improvement measure may be based on an improvement in a cost-per-click measure or an improvement in a cost-per-conversion measure relative to an overall measure for a content item campaign. The operations of block 302 may be the same or similar to the operations of blocks 202-212 of FIG. 2.

Subsets of demographic groups that have an improvement measure that meets an improvement threshold are identified as candidate subsets (304). For example, the candidate subsets may be those that have at least a threshold improvement in a cost-per-action measure relative to an overall cost-per-action for the campaign. The improvement measure for each subset can be determined, for example, using Relationship (1) above. For each subset having a negative improvement measure (and thus an improvement in cost-per-action), the absolute value of the improvement measure can be compared to an improvement threshold. The subsets having an improvement measure whose absolute value meets the improvement threshold may be identified as candidate subsets.

A conversion share, selection share, and/or spend share are identified for each candidate subset (306). Example conversion shares and selection shares are described above with reference to FIG. 2. The spend share for a candidate subset of demographic groups can be based on an amount of money spent related to providing content items to users identified as being members of a demographic group of the candidate subset and the total money spent for the campaign. For example, the spend share may be the ratio between the amount of money spent related to providing content items to users identified as being members of a demographic group of the candidate subset and the total money spent for the campaign.

Candidate subsets that have a conversion share that is less than a conversion share threshold are filtered from the candidate subsets (308). For example, the conversion share for each candidate subset may be compared to the conversion share threshold. The candidate subsets that have a conversion share that does not meet the conversion share threshold may be removed from the candidate subsets.

Candidate subsets that have a selection share that is less than a selection share threshold are filtered from the candidate subsets (310). For example, the selection share for each candidate subset may be compared to the selection share threshold. The candidate subsets that have a selection share that does not meet the selection share threshold may be removed from the candidate subsets.

Candidate subsets that have a spend share that exceeds a spend share threshold are filtered from the candidate subsets (312). For example, the spend share for each candidate subset may be compared to the spend share threshold. The candidate subsets that have a spend share that exceeds the spend share threshold may be removed from the candidate subsets.

A subset of demographic groups can be selected from the remaining candidates, if any, as described above with reference to FIG. 2. Although three filtering operations are illustrated in FIG. 3, the process 300 can include less than all three or other filtering operations. For example, if the improvement measure is based on a cost-per-conversion improvement, the candidates may not be filtered based on selection share. Similarly, if the improvement measure is based on a cost-per-selection improvement, the candidates may not be filtered based on conversion share.

FIG. 4 depicts a graph 400 of improvement measures and action shares for subsets of demographic groups. The example graph 400 includes seven subsets 401-407 and their respective improvement measures plotted against their respective action shares. The action shares may be a selection share or a conversion share. Also depicted in the graph are an improvement threshold 420 and an action share threshold 430 (e.g., a selection share threshold if the action share is a selection share or conversion share threshold if the action share is a conversion share).

Each subset 401-407 represents a subset of demographic groups for a particular content item campaign. The improvement measure and action share for each subset is determined based on performance data for the demographic groups included in the subset with respect to content items of the campaign, as described above.

In this example, subsets of demographic groups that have an improvement measure that is to the left of (e.g., less than) the improvement threshold and an action share that is above (e.g., greater than) than the action share threshold are eligible to be selected for recommendation. As the improvement threshold is to the left of the graph's origin, the subsets that are eligible are those that have an improvement measure

that is negative and has an absolute value that is greater than (or equal to) the absolute value of the improvement threshold. For example, the improvement measures for the subsets 401-407 may be determined using Relationship (1) shown above. Thus, a negative improvement measure represents a percent reduction in a cost-per-action measure relative to the overall cost-per-action measure for the campaign. Similarly, positive improvement measures represent a percent increase in a cost-per-action measure relative to the overall cost-per-action measure for the campaign.

In this example, the subsets 404 and 405 are ineligible for selection because they both have an improvement measure to the right of the improvement threshold 420. Thus, the percent improvement for each of these subsets 404 and 405 does not meet the improvement threshold 420. In this example, the subset 407 is not eligible for selection because it has an action share that is less than the action share threshold 430. In this example, the subset 407 is not eligible because both its improvement measure and its action share fail to meet the respective thresholds. Thus, the subsets 404-407 may be filtered from a set of candidate subsets using the process 300 of FIG. 3.

The remaining subsets 401-403 are eligible as each has an improvement measure that meets the improvement threshold 420 and an action share that meets the action share threshold 430. The demographic group recommendation system 132 may select one or more of these subsets for recommendation to the content item provider associated with the content item campaign. In some implementations, the demographic group recommendation system 132 would select the subset 401 as the subset 401 has the highest action share (e.g., highest conversion share or highest selection share). In some implementations, the demographic group recommendation system 132 would select the subset 402 as the subset 402 has the largest percent improvement in the cost-per-action measure. The demographic groups selection system 132 may also consider spend share, a combination of performance measures (e.g., a combination of action share, improvement measure, and/or spend share) and/or other performance metrics to select between the subsets 401-403.

FIG. 5 is a block diagram of an example computer system 500 that can be used to perform operations described above. The system 500 includes a processor 510, a memory 520, a storage device 530, and an input/output device 540. Each of the components 510, 520, 530, and 540 can be interconnected, for example, using a system bus 550. The processor 510 is capable of processing instructions for execution within

the system 500. In one implementation, the processor 510 is a single-threaded processor. In another implementation, the processor 510 is a multi-threaded processor. The processor 510 is capable of processing instructions stored in the memory 520 or on the storage device 530.

5 The memory 520 stores information within the system 500. In one implementation, the memory 520 is a computer-readable medium. In one implementation, the memory 520 is a volatile memory unit. In another implementation, the memory 520 is a non-volatile memory unit.

 The storage device 530 is capable of providing mass storage for the system 500.
10 In one implementation, the storage device 530 is a computer-readable medium. In various different implementations, the storage device 530 can include, for example, a hard disk device, an optical disk device, a storage device that is shared over a network by multiple computing devices (e.g., a cloud storage device), or some other large capacity storage device.

15 The input/output device 540 provides input/output operations for the system 500. In one implementation, the input/output device 540 can include one or more of a network interface devices, e.g., an Ethernet card, a serial communication device, e.g., and RS-232 port, and/or a wireless interface device, e.g., and 802.11 card. In another implementation, the input/output device can include driver devices configured to
20 receive input data and send output data to other input/output devices, e.g., keyboard, printer and display devices 560. Other implementations, however, can also be used, such as mobile computing devices, mobile communication devices, set-top box television client devices, etc.

 Although an example processing system has been described in FIG. 5,
25 implementations of the subject matter and the functional operations described in this specification can be implemented in other types of digital electronic circuitry, or in computer software, firmware, or hardware, including the structures disclosed in this specification and their structural equivalents, or in combinations of one or more of them.

 Embodiments of the subject matter and the operations described in this
30 specification can be implemented in digital electronic circuitry, or in computer software, firmware, or hardware, including the structures disclosed in this specification and their structural equivalents, or in combinations of one or more of them. Embodiments of the subject matter described in this specification can be implemented as one or more

computer programs, i.e., one or more modules of computer program instructions, encoded on computer storage medium for execution by, or to control the operation of, data processing apparatus. Alternatively or in addition, the program instructions can be encoded on an artificially-generated propagated signal, e.g., a machine-generated
5 electrical, optical, or electromagnetic signal, that is generated to encode information for transmission to suitable receiver apparatus for execution by a data processing apparatus. A computer storage medium can be, or be included in, a computer-readable storage device, a computer-readable storage substrate, a random or serial access memory array or device, or a combination of one or more of them. Moreover, while a computer
10 storage medium is not a propagated signal, a computer storage medium can be a source or destination of computer program instructions encoded in an artificially-generated propagated signal. The computer storage medium can also be, or be included in, one or more separate physical components or media (e.g., multiple CDs, disks, or other storage devices).

15 The operations described in this specification can be implemented as operations performed by a data processing apparatus on data stored on one or more computer-readable storage devices or received from other sources.

The term “data processing apparatus” encompasses all kinds of apparatus, devices, and machines for processing data, including by way of example a
20 programmable processor, a computer, a system on a chip, or multiple ones, or combinations, of the foregoing. The apparatus can include special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application-specific integrated circuit). The apparatus can also include, in addition to hardware, code that creates an execution environment for the computer program in
25 question, e.g., code that constitutes processor firmware, a protocol stack, a database management system, an operating system, a cross-platform runtime environment, a virtual machine, or a combination of one or more of them. The apparatus and execution environment can realize various different computing model infrastructures, such as web services, distributed computing and grid computing infrastructures.

30 A computer program (also known as a program, software, software application, script, or code) can be written in any form of programming language, including compiled or interpreted languages, declarative or procedural languages, and it can be deployed in any form, including as a stand-alone program or as a module, component,

subroutine, object, or other unit suitable for use in a computing environment. A computer program may, but need not, correspond to a file in a file system. A program can be stored in a portion of a file that holds other programs or data (e.g., one or more scripts stored in a markup language document), in a single file dedicated to the program
5 in question, or in multiple coordinated files (e.g., files that store one or more modules, sub-programs, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network.

The processes and logic flows described in this specification can be performed
10 by one or more programmable processors executing one or more computer programs to perform actions by operating on input data and generating output. The processes and logic flows can also be performed by, and apparatus can also be implemented as, special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application-specific integrated circuit).

Processors suitable for the execution of a computer program include, by way of
15 example, both general and special purpose microprocessors, and any one or more processors of any kind of digital computer. Generally, a processor will receive instructions and data from a read-only memory or a random access memory or both. The essential elements of a computer are a processor for performing actions in
20 accordance with instructions and one or more memory devices for storing instructions and data. Generally, a computer will also include, or be operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, e.g., magnetic, magneto-optical disks, or optical disks. However, a computer need not have such devices. Moreover, a computer can be embedded in another device, e.g., a
25 mobile telephone, a personal digital assistant (PDA), a mobile audio or video player, a game console, a Global Positioning System (GPS) receiver, or a portable storage device (e.g., a universal serial bus (USB) flash drive), to name just a few. Devices suitable for storing computer program instructions and data include all forms of non-volatile memory, media and memory devices, including by way of example semiconductor
30 memory devices, e.g., EPROM, EEPROM, and flash memory devices; magnetic disks, e.g., internal hard disks or removable disks; magneto-optical disks; and CD-ROM and DVD-ROM disks. The processor and the memory can be supplemented by, or incorporated in, special purpose logic circuitry.

To provide for interaction with a user, embodiments of the subject matter described in this specification can be implemented on a computer having a display device, e.g., a CRT (cathode ray tube) or LCD (liquid crystal display) monitor, for displaying information to the user and a keyboard and a pointing device, e.g., a mouse
5 or a trackball, by which the user can provide input to the computer. Other kinds of devices can be used to provide for interaction with a user as well; for example, feedback provided to the user can be any form of sensory feedback, e.g., visual feedback, auditory feedback, or tactile feedback; and input from the user can be received in any form, including acoustic, speech, or tactile input. In addition, a
10 computer can interact with a user by sending documents to and receiving documents from a device that is used by the user; for example, by sending web pages to a web browser on a user's client device in response to requests received from the web browser.

Embodiments of the subject matter described in this specification can be implemented in a computing system that includes a back-end component, e.g., as a data
15 server, or that includes a middleware component, e.g., an application server, or that includes a front-end component, e.g., a client computer having a graphical user interface or a Web browser through which a user can interact with an implementation of the subject matter described in this specification, or any combination of one or more such back-end, middleware, or front-end components. The components of the system
20 can be interconnected by any form or medium of digital data communication, e.g., a communication network. Examples of communication networks include a local area network ("LAN") and a wide area network ("WAN"), an inter-network (e.g., the Internet), and peer-to-peer networks (e.g., ad hoc peer-to-peer networks).

The computing system can include clients and servers. A client and server are
25 generally remote from each other and typically interact through a communication network. The relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other. In some embodiments, a server transmits data (e.g., an HTML page) to a client device (e.g., for purposes of displaying data to and receiving user input from a user
30 interacting with the client device). Data generated at the client device (e.g., a result of the user interaction) can be received from the client device at the server.

While this specification contains many specific implementation details, these should not be construed as limitations on the scope of any inventions or of what may be

claimed, but rather as descriptions of features specific to particular embodiments of particular inventions. Certain features that are described in this specification in the context of separate embodiments can also be implemented in combination in a single embodiment. Conversely, various features that are described in the context of a single
5 embodiment can also be implemented in multiple embodiments separately or in any suitable subcombination. Moreover, although features may be described above as acting in certain combinations and even initially claimed as such, one or more features from a claimed combination can in some cases be excised from the combination, and the claimed combination may be directed to a subcombination or variation of a
10 subcombination.

Similarly, while operations are depicted in the drawings in a particular order, this should not be understood as requiring that such operations be performed in the particular order shown or in sequential order, or that all illustrated operations be performed, to achieve desirable results. In certain circumstances, multitasking and
15 parallel processing may be advantageous. Moreover, the separation of various system components in the embodiments described above should not be understood as requiring such separation in all embodiments, and it should be understood that the described program components and systems can generally be integrated together in a single software product or packaged into multiple software products.

20 Thus, particular embodiments of the subject matter have been described. Other embodiments are within the scope of the following claims. In some cases, the actions recited in the claims can be performed in a different order and still achieve desirable results. In addition, the processes depicted in the accompanying figures do not necessarily require the particular order shown, or sequential order, to achieve desirable
25 results. In certain implementations, multitasking and parallel processing may be advantageous.

WHAT IS CLAIMED IS:

1. A method performed by data processing apparatus, the method comprising:
5 obtaining performance data for a content item campaign;
identifying, for a particular user action performed in response to presentation of a content item and based on the performance data, an overall cost-per-action measure specifying a cost for obtaining the particular user action;
grouping the performance data into demographic groups using demographic-
10 based criteria for each demographic group and demographic data for users that received content items of the content item campaign;
arranging the demographic groups in a sequence using the demographic-based criteria for the demographic groups;
identifying subsets of the demographic groups that each include one or more
15 contiguous demographic groups from the sequence of demographic groups;
for each of at least some of the subsets:
determining, based on the performance data for the subset, an improvement measure specifying a difference between the overall cost-per-action measure for the particular user action and a subset cost-per-action measure for the
20 particular user interaction, the subset cost-per-action specifying a cost for obtaining the particular user interaction from a user in the demographic groups of the subset; and
selecting a particular subset of demographic groups for providing the content item based at least on the determined improvement measures.
2. The method of claim 1, wherein:
25 the demographic-based criteria for each demographic group specifies an age range for users that have received a content item of the content item campaign;
arranging the demographic groups in a sequence comprises arranging the demographic groups in an ordinal sequence based on the age range for each demographic group; and
30 each subset of demographic groups include two or more demographic groups that are adjacent in the ordinal sequence.

3. The method of claim 1, wherein:
the demographic-based criteria for each demographic group specifies a predicted gender for users that have received a content item of the campaign and a confidence that the users are the predicted gender; and
- 5 arranging the demographic groups in a sequence comprises arranging the demographic groups in a sequence based on the predicted gender.
4. The method of claim 1, wherein the particular user action is a content item selection or a conversion.
5. The method of claim 1, wherein selecting a particular subset of demographic
- 10 groups for providing the content item comprises:
identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold;
identifying a conversion share for each candidate subset, the conversion share for a particular candidate subset being based on (a) a number of conversions resulting
- 15 from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of conversions for content items of the content item campaign; and
selecting the particular subset of demographic groups based on the conversion shares for the candidate subsets.
- 20 6. The method of claim 1, wherein selecting a particular subset of demographic groups for providing the content item comprises:
identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold;
identifying a selection share for each candidate subset, the selection share for a
- 25 particular candidate subset being based on (a) a number of selections of content items of the content item campaign resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of selections for content items of the content item campaign; and
- 30 selecting the particular subset of demographic groups based on the selection shares for the candidate subsets.

7. The method of claim 1, wherein selecting a particular subset of demographic groups for providing the content item comprises:

identifying, for each subset of demographic groups, a spend share using (a) an amount of money spent in response to providing content item of the content item

5 campaign to users identified as being members of a demographic group of the subset of demographic groups and (b) a total amount of money spent in response to providing content items of the content item campaign;

identifying, as candidate subsets, each subset of demographic groups that has a spend share that is less than a spend share threshold; and

10 selecting the particular subset of demographic groups from the candidate subsets based on the determined improvements for the candidate subsets.

8. The method of claim 1, wherein selecting a particular subset of demographic groups for providing the content item comprises:

15 identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold;

identifying, for each candidate subset, a spend share using (a) an amount of money spent in response to providing content item of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total amount of money spent in response to providing content items of the content
20 item campaign;

filtering from the candidate subsets each candidate subset that has a spend share that meets a spend share threshold;

identifying a conversion share for each remaining candidate subset, the conversion share for a particular candidate subset being based on (a) a number of
25 conversions resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of conversions for content items of the content item campaign; and

selecting, from the remaining candidate subsets, the particular subset of demographic groups based on the conversion shares for the remaining candidate
30 subsets.

9. A system, comprising:
a data processing apparatus;
a memory storage apparatus in data communication with the data processing apparatus, the memory storage apparatus storing instructions executable by the data processing apparatus and that upon such execution cause the data processing apparatus to perform operations comprising:
- 5 obtaining performance data for a content item campaign;
identifying, for a particular user action performed in response to presentation of a content item and based on the performance data, an overall cost-per-action measure specifying a cost for obtaining the particular user action;
- 10 grouping the performance data into demographic groups using demographic-based criteria for each demographic group and demographic data for users that received content items of the content item campaign;
arranging the demographic groups in a sequence using the demographic-based criteria for the demographic groups;
- 15 identifying subsets of the demographic groups that each include one or more contiguous demographic groups from the sequence of demographic groups;
for each of at least some of the subsets:
determining, based on the performance data for the subset, an improvement measure specifying a difference between the overall cost-per-action measure for the particular user action and a subset cost-per-action measure for the particular user interaction, the subset cost-per-action specifying a cost for obtaining the particular user interaction from a user in the demographic groups of the subset; and
selecting a particular subset of demographic groups for providing the content item based at least on the determined improvement measures.
- 20
25

10. The system of claim 9, wherein:
the demographic-based criteria for each demographic group specifies an age range for users that have received a content item of the content item campaign;
arranging the demographic groups in a sequence comprises arranging the
5 demographic groups in an ordinal sequence based on the age range for each demographic group; and
each subset of demographic groups include two or more demographic groups that are adjacent in the ordinal sequence.
11. The system of claim 9, wherein:
10 the demographic-based criteria for each demographic group specifies a predicted gender for users that have received a content item of the campaign and a confidence that the users are the predicted gender; and
arranging the demographic groups in a sequence comprises arranging the demographic groups in a sequence based on the predicted gender.
12. The system of claim 9, wherein the particular user action is a content item
15 selection or a conversion.
13. The system of claim 9, wherein selecting a particular subset of demographic groups for providing the content item comprises:
identifying, as candidate subsets, each subset of demographic groups that has an
20 improvement measure that meets an improvement threshold;
identifying a conversion share for each candidate subset, the conversion share for a particular candidate subset being based on (a) a number of conversions resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of
25 conversions for content items of the content item campaign; and
selecting the particular subset of demographic groups based on the conversion shares for the candidate subsets.

14. The system of claim 9, wherein selecting a particular subset of demographic groups for providing the content item comprises:

identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold;

5 identifying a selection share for each candidate subset, the selection share for a particular candidate subset being based on (a) a number of selections of content items of the content item campaign resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of selections for content items of the content
10 item campaign; and

selecting the particular subset of demographic groups based on the selection shares for the candidate subsets.

15. The system of claim 9, wherein selecting a particular subset of demographic groups for providing the content item comprises:

15 identifying, for each subset of demographic groups, a spend share using (a) an amount of money spent in response to providing content item of the content item campaign to users identified as being members of a demographic group of the subset of demographic groups and (b) a total amount of money spent in response to providing content items of the content item campaign;

20 identifying, as candidate subsets, each subset of demographic groups that has a spend share that is less than a spend share threshold; and

selecting the particular subset of demographic groups from the candidate subsets based on the determined improvements for the candidate subsets.

16. The system of claim 9, wherein selecting a particular subset of demographic groups for providing the content item comprises:

identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold;

5 identifying, for each candidate subset, a spend share using (a) an amount of money spent in response to providing content item of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total amount of money spent in response to providing content items of the content item campaign;

10 filtering from the candidate subsets each candidate subset that has a spend share that meets a spend share threshold;

identifying a conversion share for each remaining candidate subset, the conversion share for a particular candidate subset being based on (a) a number of conversions resulting from providing content items of the content item campaign to
15 users identified as being members of a demographic group of the candidate subset and (b) a total number of conversions for content items of the content item campaign; and

selecting, from the remaining candidate subsets, the particular subset of demographic groups based on the conversion shares for the remaining candidate subsets.

17. A computer storage medium encoded with a computer program, the program comprising instructions that when executed by a data processing apparatus cause the data processing apparatus to perform operations comprising:
- obtaining performance data for a content item campaign;
- 5 identifying, for a particular user action performed in response to presentation of a content item and based on the performance data, an overall cost-per-action measure specifying a cost for obtaining the particular user action;
- grouping the performance data into demographic groups using demographic-based criteria for each demographic group and demographic data for users that received
- 10 content items of the content item campaign;
- arranging the demographic groups in a sequence using the demographic-based criteria for the demographic groups;
- identifying subsets of the demographic groups that each include one or more contiguous demographic groups from the sequence of demographic groups;
- 15 for each of at least some of the subsets:
- determining, based on the performance data for the subset, an improvement measure specifying a difference between the overall cost-per-action measure for the particular user action and a subset cost-per-action measure for the particular user interaction, the subset cost-per-action specifying a cost for obtaining the
- 20 particular user interaction from a user in the demographic groups of the subset; and
- selecting a particular subset of demographic groups for providing the content item based at least on the determined improvement measures.

18. The computer storage medium of claim 17, wherein selecting a particular subset of demographic groups for providing the content item comprises:

identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold;

5 identifying a conversion share for each candidate subset, the conversion share for a particular candidate subset being based on (a) a number of conversions resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of conversions for content items of the content item campaign; and

10 selecting the particular subset of demographic groups based on the conversion shares for the candidate subsets.

19. The computer storage medium of claim 17, wherein selecting a particular subset of demographic groups for providing the content item comprises:

15 identifying, as candidate subsets, each subset of demographic groups that has an improvement measure that meets an improvement threshold;

identifying a selection share for each candidate subset, the selection share for a particular candidate subset being based on (a) a number of selections of content items of the content item campaign resulting from providing content items of the content item campaign to users identified as being members of a demographic group of the candidate subset and (b) a total number of selections for content items of the content item campaign; and

20

selecting the particular subset of demographic groups based on the selection shares for the candidate subsets.

20. The computer storage medium of claim 17, wherein selecting a particular subset of demographic groups for providing the content item comprises:

identifying, for each subset of demographic groups, a spend share using (a) an amount of money spent in response to providing content item of the content item

5 campaign to users identified as being members of a demographic group of the subset of demographic groups and (b) a total amount of money spent in response to providing content items of the content item campaign;

identifying, as candidate subsets, each subset of demographic groups that has a spend share that is less than a spend share threshold; and

10 selecting the particular subset of demographic groups from the candidate subsets based on the determined improvements for the candidate subsets.

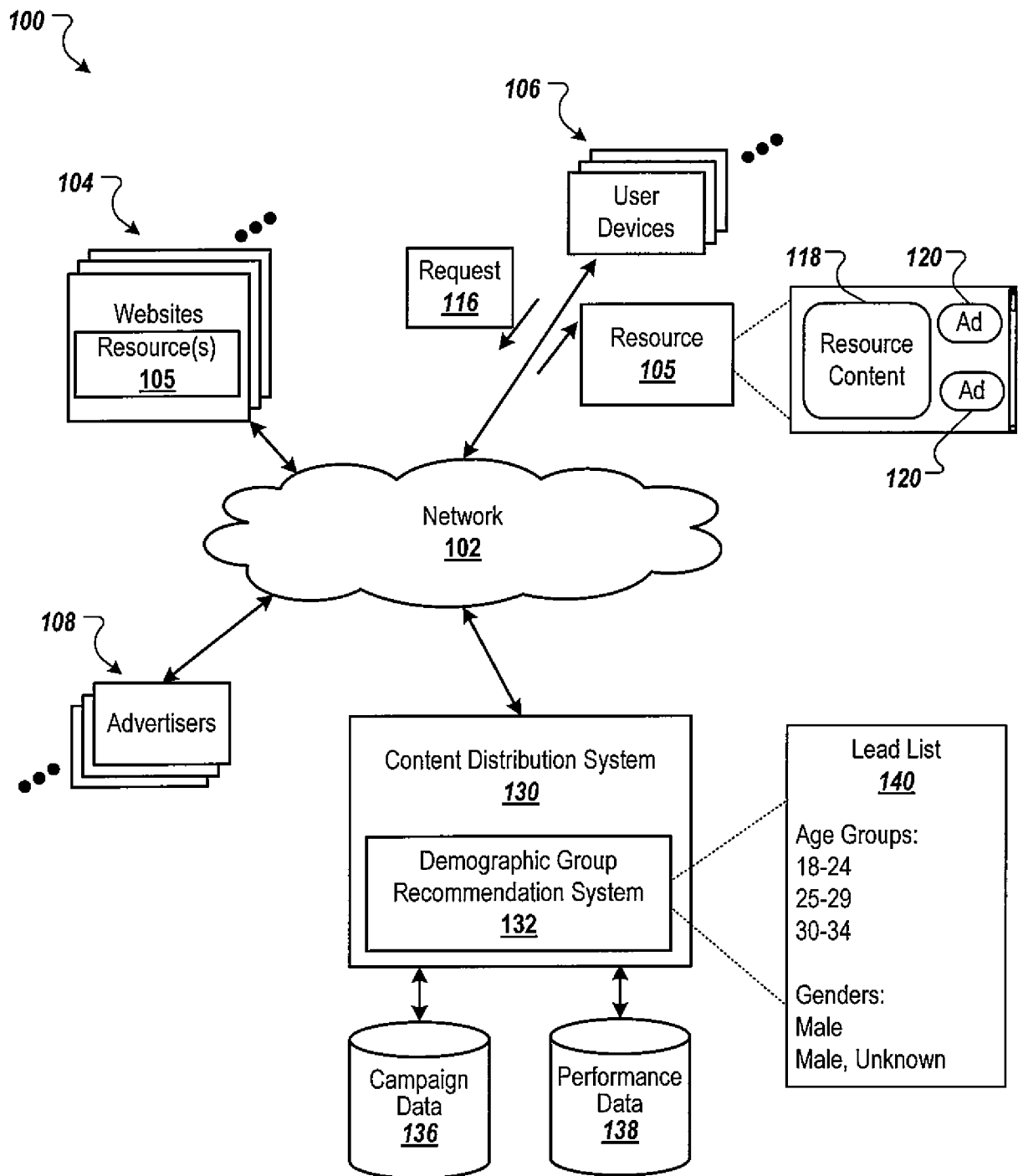


FIG. 1

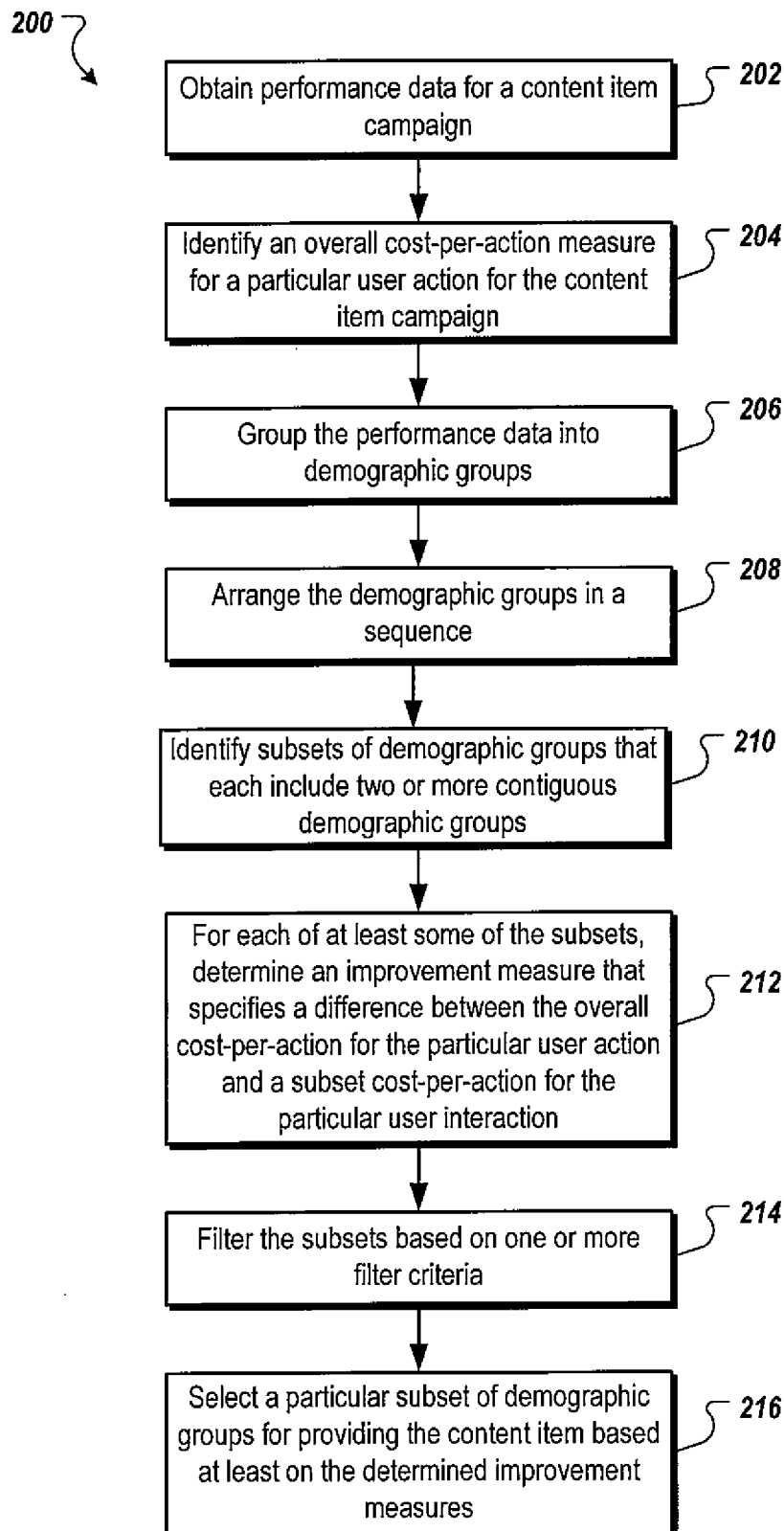


FIG. 2

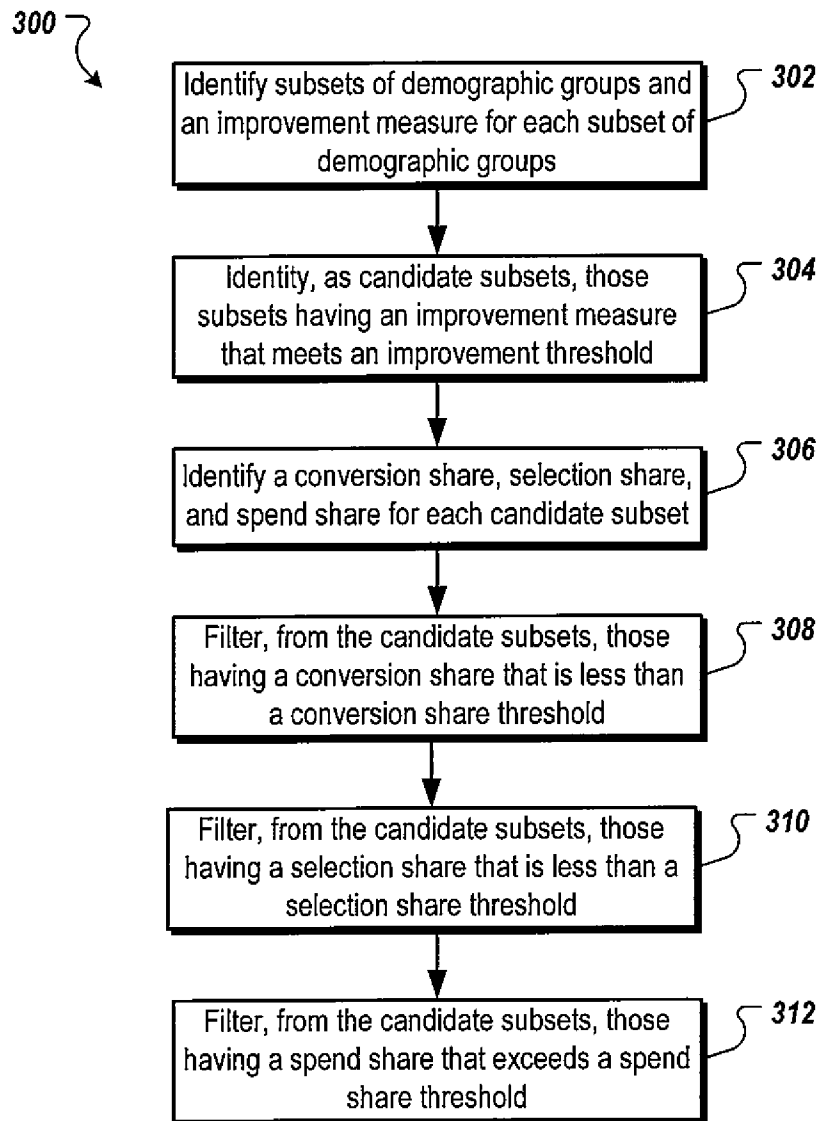


FIG. 3

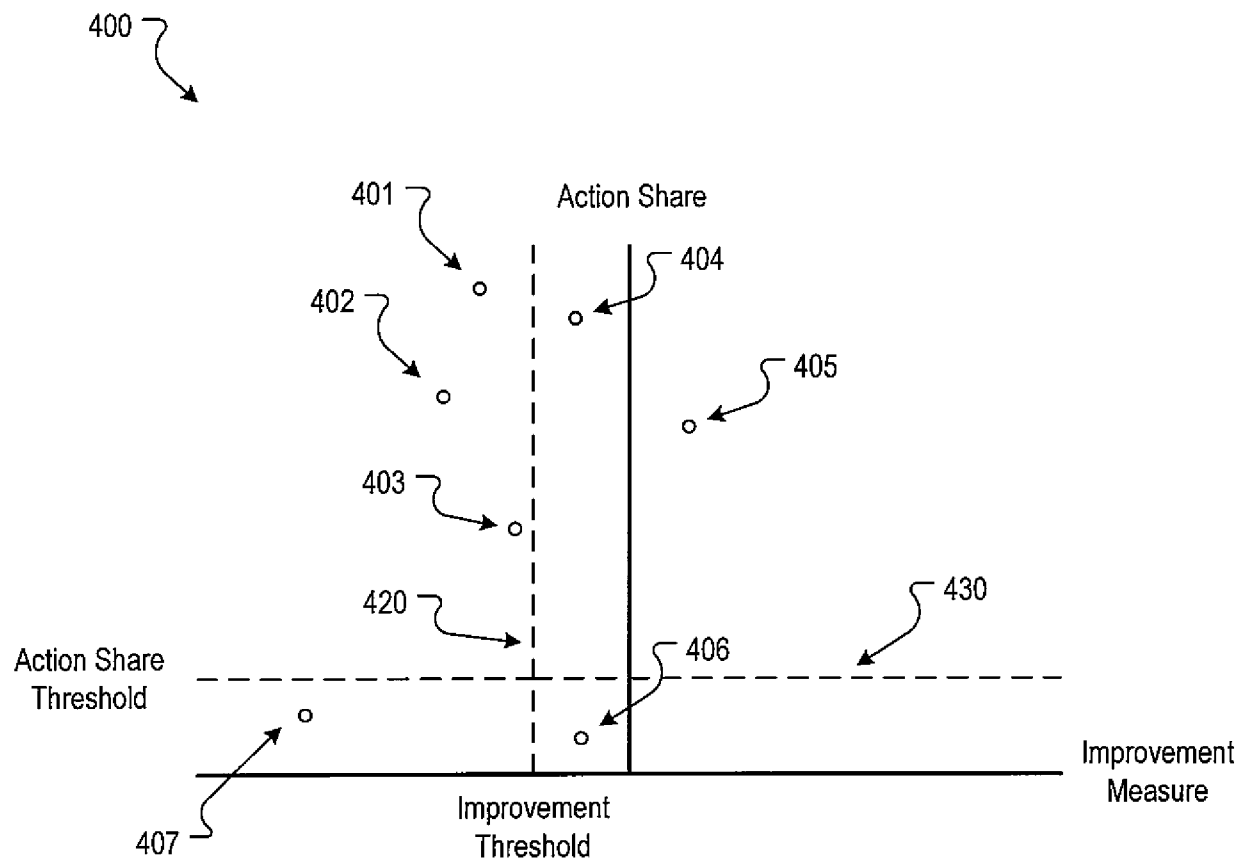


FIG. 4

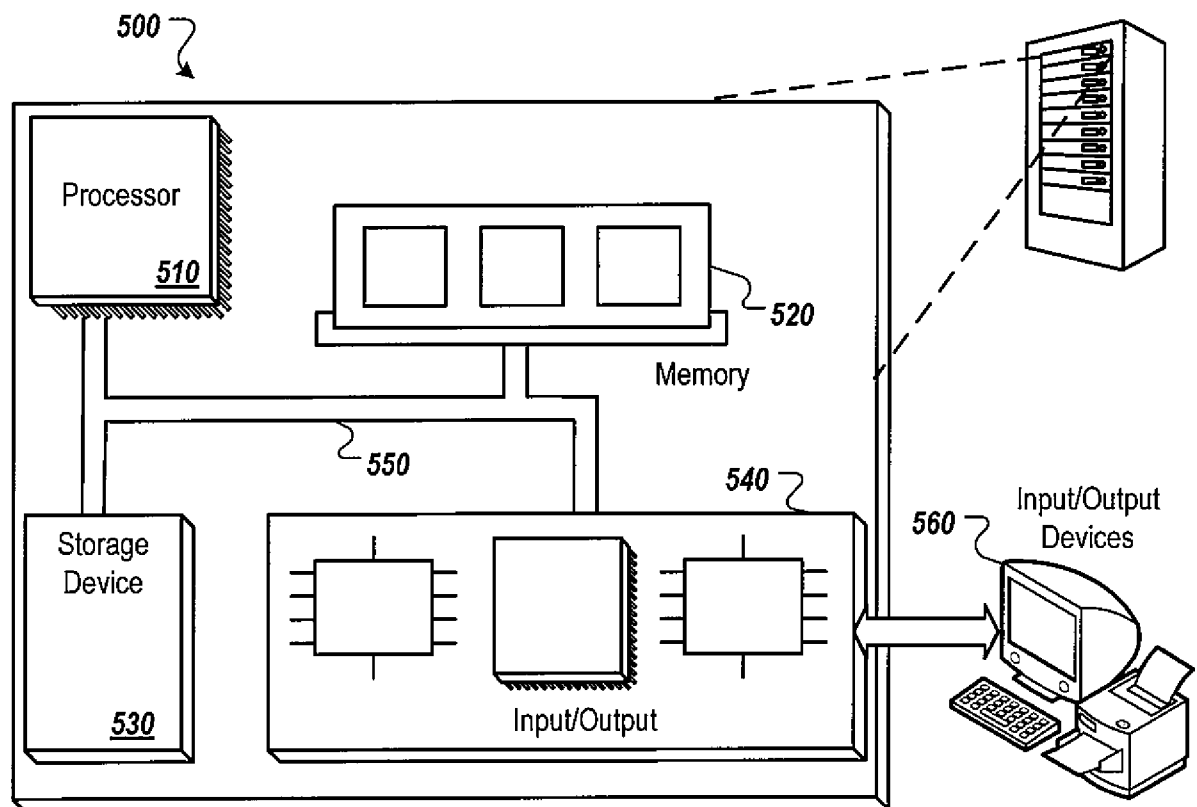


FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/083146

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 30/00(2012.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q; G06F; H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI;EPODOC;CNKI;IEEE;CNPAT: google, advertiser, advertisement, user, statistic, demographic, cost, measure, group, consumer, age

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2007075544A2 (APPLE, JEFF ET AL.) 05 July 2007 (2007-07-05) description, paragraphs [0063] to [0066]	1-20
A	US 2004122735A1 (BANG TECHNOLOGIES, LLC) 24 June 2004 (2004-06-24) the whole document	1-20
A	CN 101520782A (TAO,PENG) 02 September 2009 (2009-09-02) the whole document	1-20

☐ Further documents are listed in the continuation of Box C.
 ☒ See patent family annex.

* Special categories of cited documents:

“A”	document defining the general state of the art which is not considered to be of particular relevance	“T”	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“E”	earlier application or patent but published on or after the international filing date	“X”	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“L”	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“Y”	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“O”	document referring to an oral disclosure, use, exhibition or other means	“&”	document member of the same patent family
“P”	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

12 February 2014

Date of mailing of the international search report

28 May 2014

Name and mailing address of the ISA/

STATE INTELLECTUAL PROPERTY OFFICE OF THE
P.R.CHINA(ISA/CN)
6,Xitucheng Rd., Jimen Bridge, Haidian District, Beijing
100088 China

Authorized officer

LI,Tianxing

Facsimile No. (86-10)62019451

Telephone No. (86-10)62413429

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2013/083146

Patent document cited in search report		Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
WO	2007075544A2	05 July 2007	US	2007143186A1	21 June 2007
US	2004122735A1	24 June 2004	US	2004267561 A1	30 December 2004
CN	101520782A	02 September 2009	None		