Technical Disclosure Commons

Defensive Publications Series

December 15, 2017

Systems And Methods For Performing Geo-User Content Provision Experiments

Penelope Chu

Wiesner Vos

Henry Jackson

Ashish Saxena

James Koehler

SYSTEMS AND METHODS FOR PERFORMING GEO-USER CONTENT PROVISION EXPERIMENTS

BACKGROUND

The ability to quantify the effectiveness of online content provision can provide a variety of benefits for many parties. Geography-based content provision experiments have been used in an effort to measure the impact of online content but often suffer from a lack of statistical power. Significant amounts of noise may be involved in the measurement process when treatment and control groups include users outside of the target audience.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of an analysis system and associated environment, according to an illustrative implementation.

FIG. 2 is a flow chart of a process for performing geo-user content provision experiments, according to an illustrative implementation.

FIG. 3A is a graph of a dataset resulting from a geo-user experiment, according to an illustrative implementation.

FIG. 3B is a graph of the dataset from FIG. 3A aggregated for a geo-based regression analysis, according to an illustrative implementation.

FIG. 3C is a graph of the dataset from FIG. 3A aggregated for a time-based regression analysis, according to an illustrative implementation.

FIG. 4 is an illustration of an interface including results from a geo-user experiment, according to an illustrative implementation.

FIG. 5 is a block diagram of a computing system, according to an illustrative implementation.

DETAILED DESCRIPTION

Referring generally to the FIGURES, systems and methods for performing geo-user content provision experiments are shown, according to various illustrative implementations. Grouping users by geographic regions ("geos") may provide statistically similar pools of users when conducting an experiment. An activity log comprising a plurality of user actions may be filtered according to one or more characteristics of a set of one or more content items (e.g., a content campaign) in order to determine a subset of users that express some degree of likelihood of being exposed to the content items. One or more geos and associated users of the subset of users may be designated as part of a treatment group, and similarly one or more geos and associated users of the subset of users may be designated as part of a control group. A change in spend level for the associated content item(s) may be introduced for the treatment group. A response metric may be tracked among treatment and control groups both before the change in spend level occurs and after the change in spend level occurs in order to estimate an impact of the change in spend level on the response metric.

For situations in which the systems discussed herein collect and/or utilize personal information about users, or may make use of personal information, the users may be provided with an opportunity to control whether programs or features that may collect personal information (e.g., information about a user's social network, social actions or activities, a user's

preferences, a user's current location, etc.), 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 anonymized in one or more ways before it is stored or used, so that personally identifiable information is removed when generating parameters (e.g., demographic parameters). For example, a user's identity may be anonymized so that no personally identifiable 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, the user may have control over how information is collected about him or her and used by a content server. Further, the individual user information itself is not surfaced to the content provider, so the content provider cannot discern the interactions associated with particular users.

Referring now to FIG. 1, a block diagram of an analysis system 120 and associated environment 100 is shown, according to an illustrative implementation. User devices 104 (e.g., smartphones, tablets, computers, etc.) may be used by one or more users to perform various actions and/or access various types of resources over a network 102. Network 102 may be any kind of electronic communications network (e.g., the Internet). Throughout this disclosure, the term "user" may refer to any individual user that access resources via network 102 from a user device 104. Accessing resources may include, for example, navigating webpages (e.g., via an internet browser), using applications (e.g., social media, news, navigation, etc.), and performing search queries (e.g., searching for shoes, cars, etc.). In some implementations, one or more user devices 104 have enabled location services which can be tracked over network 102. Locations services may use GPS or other technologies to determine a location of a device 104.

In some implementations, a content management system 110 is connected to network 102. Content management system 110 is shown to be connected to a content database 112 which stores content items 114. Content management system 110 may include one or more servers, processors, and computer-readable instructions that can be configured to communicate over network 102. Resources accessed by users over network 102 may include content interfaces for displaying content. For example, a golf website may have one or more content interfaces that contain one or more content items 114 (e.g., a content item related to a new putter). Content management system 110 may determine content to be published in or more content interfaces associated with one or more resources. In some implementations, content management system 110 is configured to conduct a content auction among third-party content providers for each content interface. An auction winner may be determined based on bid amounts and/or a quality score (i.e., a measure of how likely a user is to click or select the content item). Content management system 110 may also allow content providers to create content campaigns. Each campaign can include any number of parameters such as a minimum and maximum bid amount, a target bid amount, a budget (e.g., daily, weekly, monthly, total), etc. One or more content provider devices 106 may be used to access content management system 110 in order to adjust campaign parameters, add or remove content items, and otherwise configure the presentation of content items managed by system 110.

External resources 108 may be any type of resource that can be accessed via network 102 other than those shown in FIG. 1. External resources relevant to the present disclosure may include data aggregating systems, location data sources, content provider data sources, etc. In

some implementations, analysis system 120 is configured to retrieve offline sales metrics such as physical store visits, in-store sales, etc.

Analysis system 120 may be used by content providers in an effort to quantify the impact of their content items. System 120 may include one or more servers, one or more processors, and computer-readable instructions. System 120 may be configured to communicate over network 102 via a variety of architectures (e.g., client/server, peer-to-peer, etc.). System 120 can be configured to provide a variety of interfaces for setting up experiments (e.g., via a content provider device 106), monitoring progress of experiments, and viewing results of experiments.

System 120 is shown to include an experiment database 121 which may be configured to store a variety of information relevant to content provision experiments performed by system 120. Information may be received from user devices 104, content provider devices 106, external resources 108, and/or content management system 110, for example. In some implementations, a content provider submits a request to perform a geo-user experiment to system 120 and information about the request (e.g., content items, campaign identification, desired change in spend level, geographic areas to target, etc.) may be stored in database 121. In addition, system 120 may be configured to retrieve data via network 102 (e.g., user activity data, content campaign data, etc.) which may be stored in database 121.

System 120 is shown to include one or more processors 122 along with machine-readable instructions 123 that may include one or more modules. Each module may be implemented using hardware, software (e.g., computer-readable instructions), or a combination thereof. In the illustrated implementation, the modules include a content module 124, a user activity module

125, a geography module 126, an analysis module 127, and an interface module 128. It should be understood that various implementations may include more, fewer, or different modules than illustrated in FIG. 1, and all such modifications are contemplated within the scope of the present disclosure.

Content module 124 may be configured to communicate with content management system 110 via network 102 in order to determine a set of one or more content items associated with a content provider to be analyzed during a geo-user experiment. In addition, content module 124 may be configured to determine one or more characteristics associated with the one or more content items. Characteristics may include associated keywords used in a search query, website views, video views, content views, content clicks, etc. For example, content module 124 may be configured to determine (e.g., via a campaign ID or other identifier) content items associated with a content campaign for a new restaurant. In this example, content module 124 may also determine that the set of content items is presented based on a set of target keywords (e.g., restaurant, new restaurant, restaurant in (geographic location), etc.). Content module 124 may also be configured to initiate a change in spend level associated with a set of content items for analysis during a geo-user experiment.

User activity module 125 may be configured to retrieve and analyze user activity data comprising actions performed by users over network 102. In some implementations, user activity module 125 retrieves user activity data and creates an activity log with one or more log entries. The activity log can span over any specified time period (e.g., past month, past week, etc.) and can be specific to users based on any constraints (e.g., users in France, users in Los

Angeles, users of a particular mobile operating system in Boston, etc.). When conducting a geo-user experiment, an activity log may be filtered (i.e., certain entries removed) by user activity module 125 according to one or more characteristics of the set of content items being analyzed. User activity module 125 may be configured to use the filtered activity log in order to determine a subset of users (i.e., a subset of the users associated with the original activity log). The subset of users may be users that have a likelihood of being exposed to the content items being analyzed. In addition, user activity module 125 may be configured to retrieve user activity data related to a response metric being analyzed during an experiment.

Geography module 126 may be configured to designate geographic regions as control or treatment groups. A geographic region of interest (e.g., the United States) can be partitioned into a set of smaller geographic areas, or "geos". These geos can provide comparable sets of users for testing during a geo-user experiment. Details of how geos are chosen are beyond the scope of this disclosure, however geos generally are large enough (e.g., at least larger than a postal code) to ensure content serving accuracy and the ability to monitor the desired response metric at the geo level. In the United States, for example, one possible set of geos is the standard 210 designated marketing areas (DMAs). After a set of two or more geos for the geo-user experiment is identified, geography module 126 can designate each geo into a control group or a treatment group. A change in spend level will only be implemented for geos in the treatment group, whereas geos in the control group will remain unchanged. The designation of geos into control or treatment groups can be implemented in a variety of ways, including randomization or designation by a content provider.

Analysis module 127 can be configured to analyze results of a geo-user experiment. In some implementations, analysis module 127 retrieves data via network 102 related to a response metric being analyzed during the experiment. For example, if the response metric being measured is physical store visits, analysis module 127 can be configured to retrieve store visit data over network 102. A variety of response metrics can be tracked during a geo-user experiment. In some implementations, the response metric is an offline sales metrics such as physical store visits. Store visits may be determined using location information from one or more user devices 104. In some implementations, the response metric may include user interactions in a mapping interface, which may be indicative of an intention to visit a physical location or store. User interactions with the mapping interface may include, for example, searching for store locations within the control group or treatment group, requesting directions to a location of a store within the control group or the treatment group, and/or navigating to a location of a store within the control group or the treatment group. Online sales metrics such as conversion data or any other user-specific action that can be measured and defined as a conversion event (e.g., online sales, provision of requested data via an online form, etc.) can also be used. Data can be retrieved from content management system 110, user devices 104 (e.g., through the use of cookies or other identifiers), content provider devices 106, and/or external resources 108, for example. Analysis module 127 can be configured to use a variety of statistical modeling approaches in order to calculate an impact of the modification of a spend level associated with the geo-user experiment. In some implementations, analysis module 127 is configured to use a time-based regression analysis to calculate an impact of a modification of a spend level. In other implementations, analysis module 127 is configured to use a geo-based

regression analysis to calculate an impact of a modification of a spend level. More detail regarding time-based regression analysis and geo-based regression analysis is described below.

Interface module 128 can be configured to generate graphical interface data for presentation to a content provider. The graphical interface data can be based on a calculated impact of a modification of a spend level and can be generated for presentation to a content provider device 106, for example. In some implementations, the graphical interface data includes a recommended spend level associated with a set of one or more content items analyzed during a geo-user experiment. More detail about graphical interface data is described below.

Referring now to FIG. 2, a flow chart of a process 200 for performing geo-user experiments is shown, according to some implementations. Process 200 can be performed by analysis system 120, for example. Process 200 can be used to conduct geo-user experiments that provide advanced estimates of an impact of online content on various response metrics associated with a content provider.

Process 200 is shown to include receiving a request to perform a geo-user experiment (step 202). The request can be submitted by a content provider via a content provider device 106, for example. A variety of information to be analyzed may be included in the request, such as which content item(s) or campaign(s) is to be analyzed, an amount of spend change (e.g., maximum acceptable amount of spend change acceptable to content provider), and a desired response metric. As described above, many different response metrics are possible, such as offline (e.g., store visit) data or online conversion events (e.g., clicks, views, online sales, etc). In some implementations, content providers wish to analyze the effect of content provided over

network 102 on an offline sales metric. In such implementations, system 120 can be configured to analyze the effects of a change in spend level on physical store visits or expressed interest in visiting a physical store (e.g., searching for store location and/or requesting directions to a store). Physical store visits can be determined by retrieving data from location-enabled user devices 104 and/or retrieving data from external resources 108. Information included with the request to perform a geo-user experiment may be stored in experiment database 121.

Process 200 is shown to include determining a set of one or more content items to be analyzed during the experiment and one or more characteristics associated with the set of content items (step 204). In some implementations, step 204 is performed by content module 124. Content module 124 can be configured to determine the set of one or more content items to be analyzed during the experiment by communicating with content management system 110 over network 102, for example. The set of one or more content items to be analyzed may be designated by the content provider when submitting the request to perform an experiment in step 202. In some implementations, the content items to be analyzed may be a particular content item or set of content items, all content items associated with a content campaign defined by content management system 110, or all content items associated with the content provider. Content module 124 can also determine one or more characteristics associated with the content items for use in determining users to be involved in the experiment. Characteristics may include, but are not limited to, keywords associated with the content items (e.g., target keywords used to determine to whom the content items are presented), an industry or vertical associated with the content items, a subject matter of the content items, defined audience characteristics for the content items, etc.

Process 200 is shown to include retrieving an activity log comprising a plurality of user actions and filtering the activity log to determine a subset of users using the one or more characteristics associated with the set of content items being analyzed (step 206). In some implementations, step 206 is performed by user activity module 125. The activity log may be formed from data retrieved by system 120 via network 102 and stored in database 121 and may include log entries representative of user interactions with content items. The activity log retrieved in step 206 may be filtered according to one or more characteristics so that any log entries not associated with the characteristics (e.g., associated with interactions with subject matter different from the subject matter of the content items being analyzed) are removed. The filtered activity log may be used to determine a subset of users to be included in the experiment (e.g., users that are most likely to be exposed to the set of content items).

Process 200 is shown to include retrieving data comprising actions performed by the subset of users over a first timeframe (step 208). In some implementations, step 208 is performed by user activity module 125. The first timeframe can be a pretest period of the experiment in which a spend level for the set of content items is equal for all users of the subset of users determined in step 206. The pretest period may be a few months or any other time period. The pretest period can provide a baseline for the subset of users that can be used when calculating an impact of a change in spend level introduced in future steps. In some implementations, data collected during a pretest period may determine the length of a test period, the magnitude of treatment intervention (e.g., change in spend level), or both. Geography module 126 can determine a geographic location of the subset of users observed during the pretest period.

Process 200 is shown to include designating one or more geographic regions and corresponding users of the subset of users as a control group and one or more geographic regions and corresponding users of the subset of users as a treatment group (step 210). In some implementations, step 210 is performed by geography module 126. As described above, a geographic area (e.g., a country, state, city, county, etc.) can be partitioned into a set of sub-regions, or "geos" which can provide statistically similar groups of users. In step 210, each geo being analyzed as part of the geo-user experiment can be designated as a control geo or a treatment geo. The collection of one or more geos designated as treatment geos can be referred to as the "treatment group" and the collection of one or more geos designated as control geos can be designated as the "control group". Various approaches can be taken when choosing which geos to designate as treatment and control geos, or which geos to include in the experiment at all. Some possibilities include randomization of geo designation (e.g., by geography module 126) and content provider designation of geos (e.g., in step 202). Once a control group and a treatment group have been determined, a change in spend level can be implemented and observed during a test period of the geo-user experiment.

Process 200 is shown to include modifying a spend level for presenting the set of content items to one or more user devices of the subset of users within the treatment group (step 212). In some implementations, step 212 is performed by content module 124. The spend level can be either increased or decreased. Once the spend level has been changed, the pretest period ends and the test period begins.

Process 200 is shown to include retrieving data comprising actions performed by the subset of users (i.e., both treatment and control groups) over a second timeframe after the spend level is modified for the treatment group (step 214). In some implementations, step 214 is performed by user activity module 125. Step 214 can be identical to step 208 with the only difference being the change in spend level introduced for the treatment group in step 212.

Process 200 is shown to include calculating an impact of the modification of the spend level for the treatment group and generating graphical interface data for presentation to a device of the content provider using the calculated impact (step 216). In some implementations, step 216 is performed by analysis module 127 and interface module 128. Step 216 involves statistical analysis and modeling of four main sets of user activity data: control group during pretest period, treatment group during pretest period, control group during test period, and treatment group during test period. More detail about the impact calculation and generation of graphical interface data is described below.

Referring now to FIG. 3A, an example dataset 300 resulting from a geo-user experiment is shown, according to some implementations. Each of graphs 301-308 corresponds to a response metric for a single geo. In this example, the response metric is physical store visits, however, as described above, a variety of response metrics can be tracked during a geo-user experiment. Dataset 300 results from a geo-user experiment with approximately a two month pretest period (Feb 1 – Apr 1) and a one month test period (Apr 1 – May 1). Graphs 301, 302, 305, and 307 depict data resulting from treatment group geos and graphs 303, 304, 306, and 308 depict data resulting from control group geos. From graphs 301-308 it can be seen that a change

in spend level occurring around April 1st for the treatment group results in an increase in store visits. This increase can be analyzed and compared to the change in spend level in order to calculate an impact of the spend change on the number of store visits (i.e., response metric).

Referring now to FIG. 3B, a graph of dataset 300 aggregated for a geo-based regression analysis is shown, according to some implementations. As shown, data aggregation for geo-based regression analysis results in two data points for each of geos 301-308: one representing the pretest period and one representing the test period. Different methods of aggregation are possible, but in the case the data points for the pretest and test periods are an observed average of the response metric for each geo over each time period. Geo-based regression analysis provides a straightforward approach to analyzing results from a geo-user experiment and can achieve a higher degree of accuracy as the number of geos involved in an experiment increases.

Referring now to FIG. 3C, a graph of dataset 300 aggregated for a time-based regression analysis is shown, according to some implementations. As shown, data aggregation for time-based regression analysis results in two data points per time interval: one for the treatment group and one for the control group. Any time interval may be used, however a new pair of data points is often calculated daily or weekly (in this case, daily). As shown in FIG. 3C, graph 310 represents the treatment group and graph 320 represents the control group. A noticeable increase can be seen in graph 310 during the test period.

Aggregated datasets can be used as inputs to one or more statistical models. In some implementations, the statistical model uses regression analysis techniques to estimate relationships among variables. For example, data collected during a pretest period can be used to

train the model and estimate unknown parameters in order to predict behavior of the treatment group during the test period. In some implementations, the statistical model is designed to estimate a return on ad spend (ROAS) associated with the change in spend level for the treatment group. Design considerations that may affect uncertainty and estimation of ROAS include magnitude of the change in spend level, pretest period length, test period length, the fraction of geos designated as control geos, and the number of geos. In general, ROAS may be defined as an observed change in response metric divided by a change in spend level (i.e., cost to content provider).

$$[0001] \quad ROAS = \frac{\Delta \text{ response metric}}{\Delta \text{ cost}}$$

In some implementations, a user-level model may be used to analyze data from a geo-user experiment. A user-level model can be configured to calculate an impact of a change in spend level without aggregation of data across users as described above with respect to time-based regression analysis and geo-based regression analysis. A user-level model can be configured to analyze user-level data at the geo-level during the analysis phase. Randomization occurring at the geo-level may be utilized by a user-level model.

Referring now to FIG. 4, an interface 400 for presenting results of a geo-user experiment to a content provider is shown, according to some implementations. Interface 400 may provide a summary of experiment results that includes experiment information 405, experiment results 410, and recommendations 415. Experiment information 405 may include various information about the geo-user experiment such as customer (e.g., content provider) identification, content campaign identification, geographic area for the experiment, and/or a time period indicating a

duration of the experiment. Experiment results section 410 may provide an indication of a change in spend level, a corresponding change in a response metric, and a calculated ROAS. In some implementations, various links are provided within interface 400 that can be selected in order to view further details about the experiment results. Interface 400 may also include one or more recommendations 415 that can be generated based on the results of the experiment. It should be noted that the interface 400 is meant to be illustrative only, and a wide variety of interfaces for presentation to a content provider can be generated by analysis system 120.

FIG. 5 illustrates a depiction of a computer system 500 that can be used, for example, to implement an illustrative user device 104, an illustrative content management system 110, an illustrative content provider device 106, an illustrative analysis system 120, and/or various other illustrative systems described in the present disclosure. The computing system 500 includes a bus 505 or other communication component for communicating information and a processor 510 coupled to the bus 505 for processing information. The computing system 500 also includes main memory 515, such as a random access memory (RAM) or other dynamic storage device, coupled to the bus 505 for storing information, and instructions to be executed by the processor 510. Main memory 515 can also be used for storing position information, temporary variables, or other intermediate information during execution of instructions by the processor 510. The computing system 500 may further include a read only memory (ROM) 510 or other static storage device coupled to the bus 505 for storing static information and instructions for the processor 510. A storage device 525, such as a solid state device, magnetic disk or optical disk, is coupled to the bus 505 for persistently storing information and instructions.

The computing system 500 may be coupled via the bus 505 to a display 535, such as a liquid crystal display, or active matrix display, for displaying information to a user. An input device 530, such as a keyboard including alphanumeric and other keys, may be coupled to the bus 505 for communicating information, and command selections to the processor 510. In another implementation, the input device 530 has a touch screen display 535. The input device 530 can include a cursor control, such as a mouse, a trackball, or cursor direction keys, for communicating direction information and command selections to the processor 510 and for controlling cursor movement on the display 535.

In some implementations, the computing system 500 may include a communications adapter 540, such as a networking adapter. Communications adapter 540 may be coupled to bus 505 and may be configured to enable communications with a computing or communications network 545 and/or other computing systems. In various illustrative implementations, any type of networking configuration may be achieved using communications adapter 540, such as wired (e.g., via Ethernet), wireless (e.g., via WiFi, Bluetooth, etc.), pre-configured, ad-hoc, LAN, WAN, etc.

According to various implementations, the processes that effectuate illustrative implementations that are described herein can be achieved by the computing system 500 in response to the processor 510 executing an arrangement of instructions contained in main memory 515. Such instructions can be read into main memory 515 from another computer-readable medium, such as the storage device 525. Execution of the arrangement of instructions contained in main memory 515 causes the computing system 500 to perform the illustrative

processes described herein. One or more processors in a multi-processing arrangement may also be employed to execute the instructions contained in main memory 515. In alternative implementations, hard-wired circuitry may be used in place of or in combination with software instructions to implement illustrative implementations. Thus, implementations are not limited to any specific combination of hardware circuitry and software.

Although an example processing system has been described in FIG. 5, implementations of the subject matter and the functional operations described in this specification can be carried out using 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.

Implementations of the subject matter and the operations described in this specification can be carried out using digital electronic circuitry, or in computer software embodied on a tangible medium, firmware, or hardware, including the structures disclosed in this specification and their structural equivalents, or in combinations of one or more of them. Implementations 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 one or more 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 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-readable 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 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 components or media (e.g., multiple CDs, disks, or other storage devices). Accordingly, the computer storage medium is both tangible and non-transitory.

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" or "computing device" encompasses all kinds of apparatus, devices, and machines for processing data, including by way of example, a 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 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.

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 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 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 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 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 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 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, implementations of the subject matter described in this specification can be carried out using 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 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 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.

Implementations of the subject matter described in this specification can be carried out using a computing system that includes a back-end component, e.g., as a data 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 backend, middleware, or frontend components. The components of the system 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 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 implementations, 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 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.

In some illustrative implementations, the features disclosed herein may be implemented on a smart television module (or connected television module, hybrid television module, etc.), which may include a processing circuit configured to integrate internet connectivity with more traditional television programming sources (e.g., received via cable, satellite, over-the-air, or other signals). The smart television module may be physically incorporated into a television set or may include a separate device such as a set-top box, Blu-ray or other digital media player, game console, hotel television system, and other companion device. A smart television module may be configured to allow viewers to search and find videos, movies, photos and other content on the web, on a local cable television channel, on a satellite television channel, or stored on a local hard drive. A set-top box (STB) or set-top unit (STU) may include an information appliance device that may contain a tuner and connect to a television set and an external source of signal, turning the signal into content which is then displayed on the television screen or other display device. A smart television module may be configured to provide a home screen or top level screen including icons for a plurality of different applications, such as a web browser and a plurality of streaming media services, a connected cable or satellite media source, other web "channels", etc. The smart television module may further be configured to provide an electronic programming guide to the user. A companion application to the smart television module may be operable on a mobile computing device to provide additional information about available programs to a user, to allow the user to control the smart television module, etc. In alternate implementations, the features may be implemented on a laptop computer or other personal computer, a smartphone, other mobile phone, handheld computer, a tablet PC, or other computing device.

ABSTRACT

Systems, methods, and computer-readable storage media that may be used to conduct geo-user advertising experiments. One method includes determining a set of content items and filtering an activity log based on one or more characteristics of the set of content items to determine a subset of users. The method further includes retrieving geo-based data from the subset of users over a first timeframe. The method further includes designating one or more geographic regions associated with the subset of users as a treatment group and one or more geographic regions associated with the subset of users as a control group. The method further includes modifying a spend level associated with the treatment group and retrieving geo-based data over a second timeframe. The method further includes calculating an impact of the modification of the spend level and generating graphical interface data for presentation to a device of a content provider.

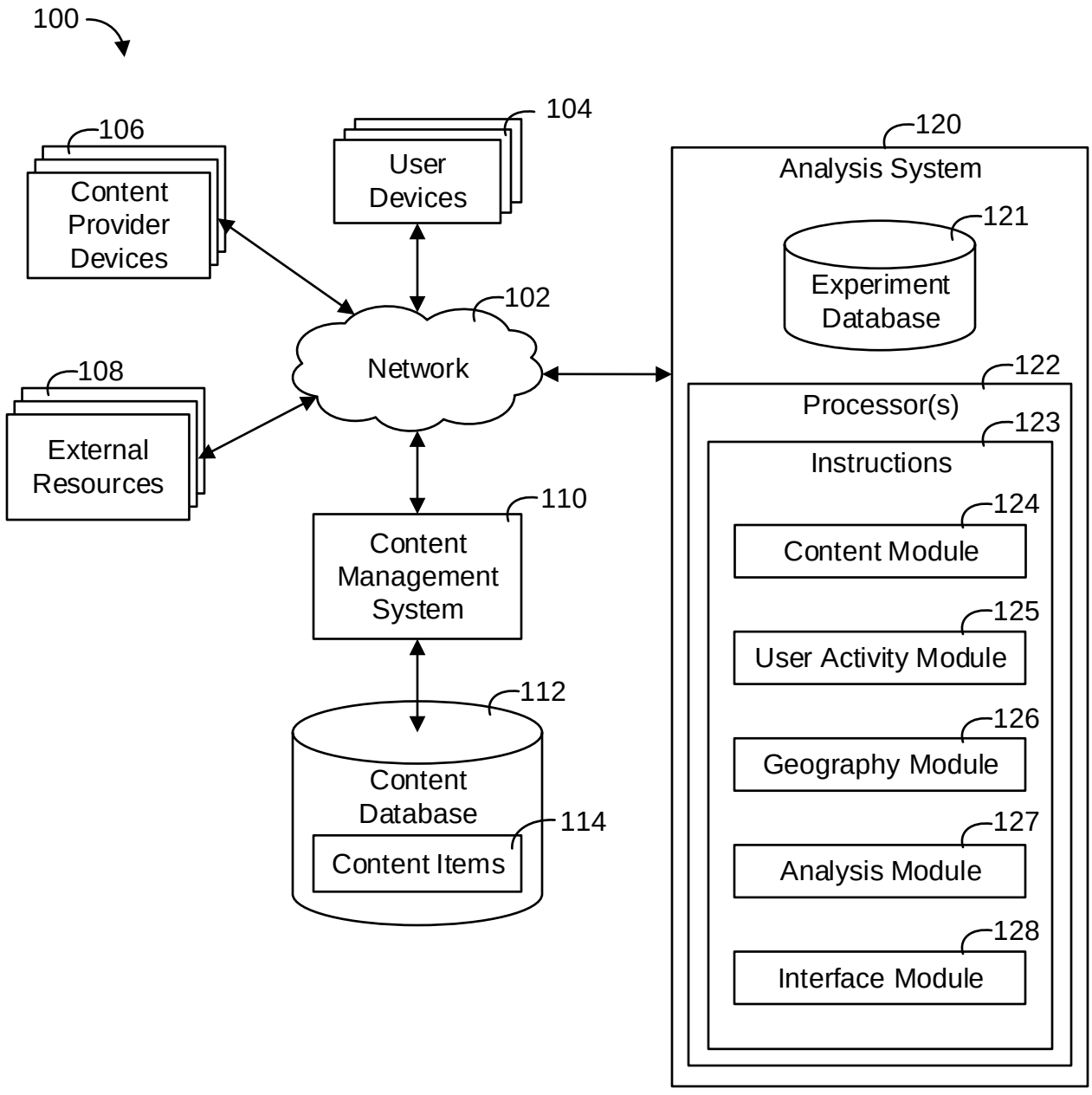
100
104
106
User Devices
Content Provider Devices
120
Analysis System
121
Experiment Database
102
Network
108
External Resources
122
Processor(s)
123
Instructions
110
Content Management System
124
Content Module
125
User Activity Module
112
126
Geography Module
Content Database
114
127
Content Items
Analysis Module
128
Interface Module

FIG. 1

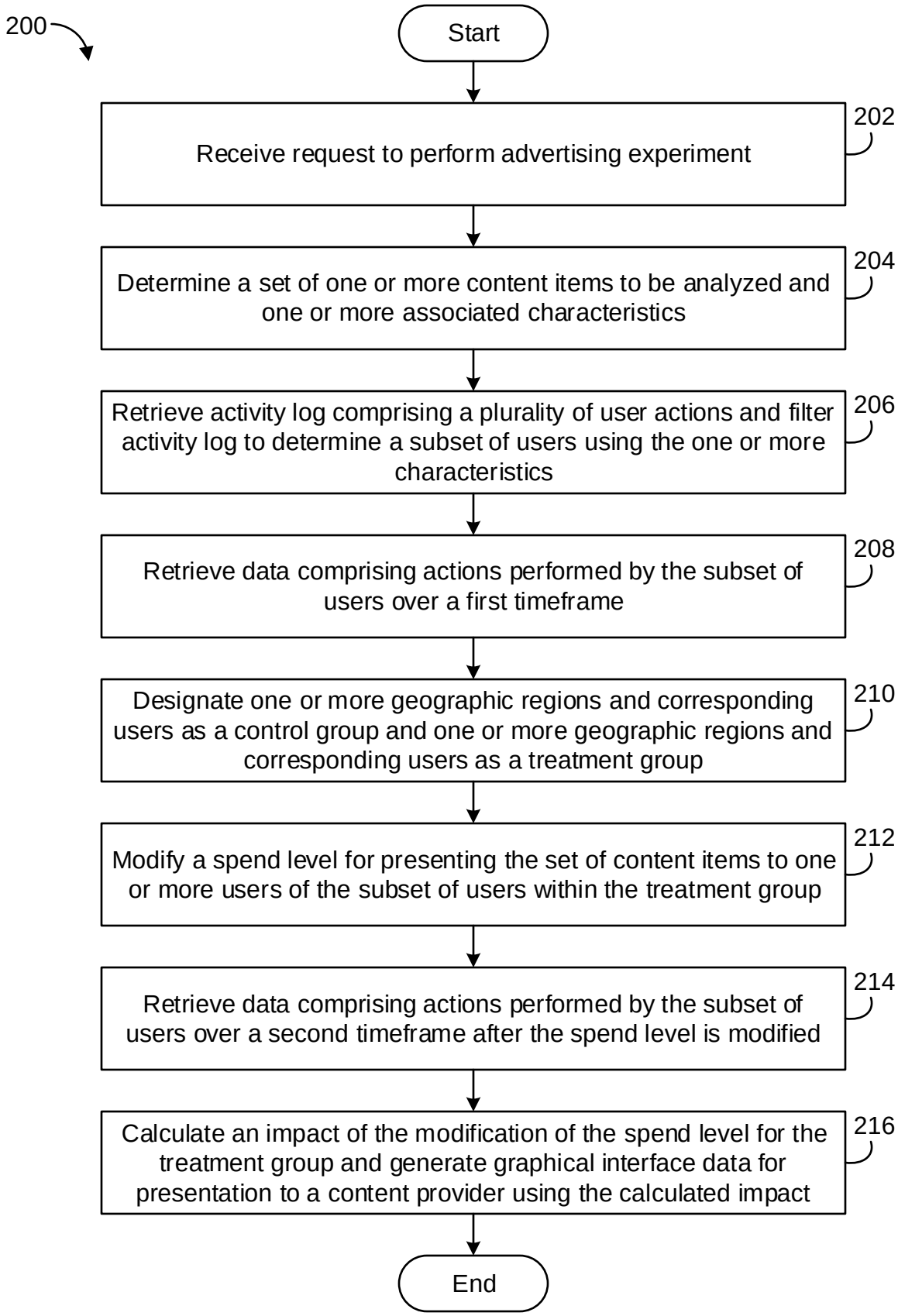
200
Start
Receive request to perform advertising experiment
202
Determine a set of one or more content items to be analyzed and one or more associated characteristics
204
Retrieve activity log comprising a plurality of user actions and filter activity log to determine a subset of users using the one or more characteristics
206
Retrieve data comprising actions performed by the subset of users over a first timeframe
208
Designate one or more geographic regions and corresponding users as a control group and one or more geographic regions and corresponding users as a treatment group
210
Modify a spend level for presenting the set of content items to one or more users of the subset of users within the treatment group
212
Retrieve data comprising actions performed by the subset of users over a second timeframe after the spend level is modified
214
Calculate an impact of the modification of the spend level for the treatment group and generate graphical interface data for presentation to a content provider using the calculated impact
216
End

FIG. 2

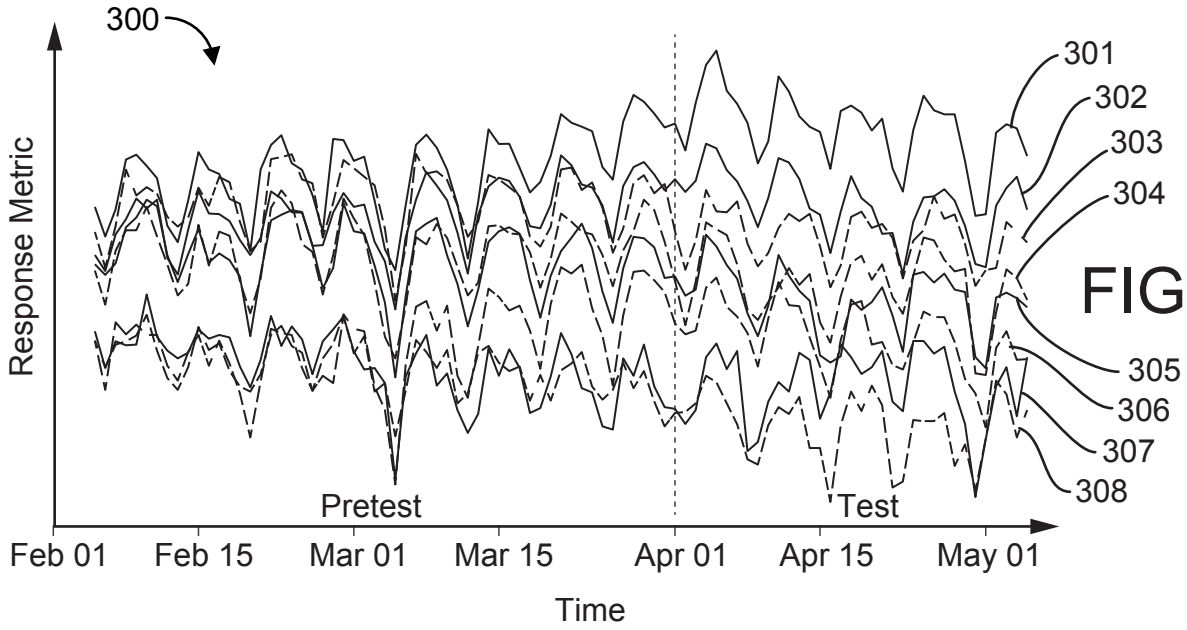

FIG. 3A

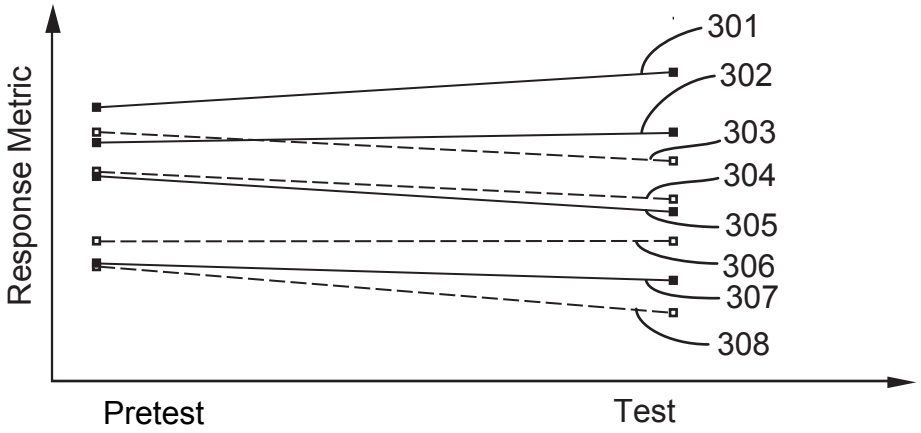

FIG. 3B

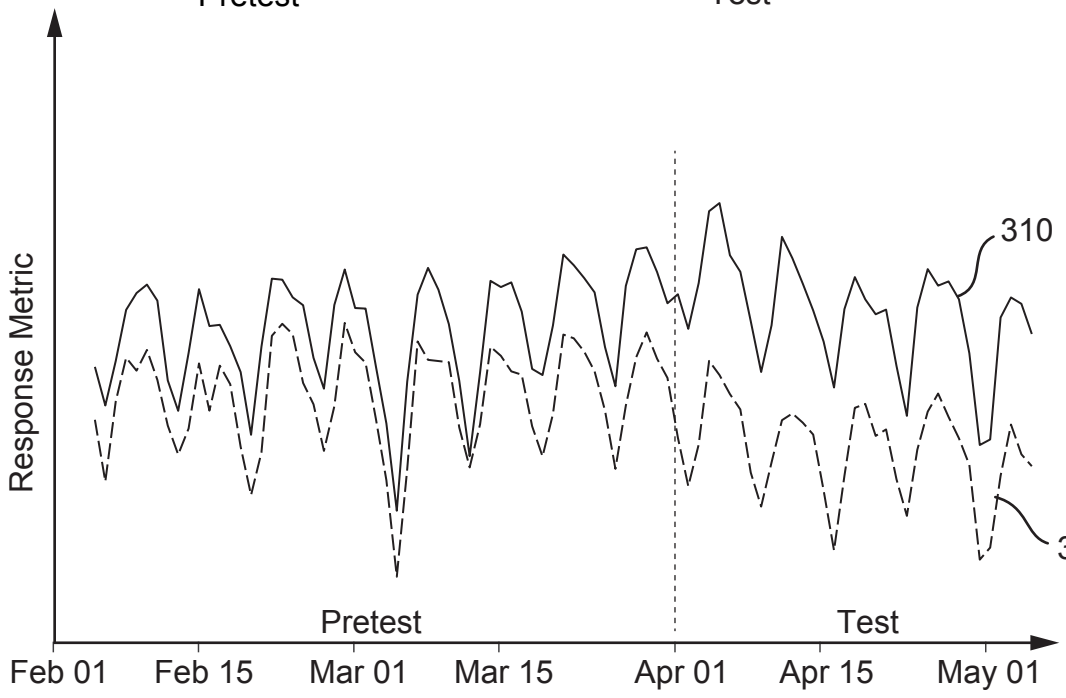

FIG. 3C

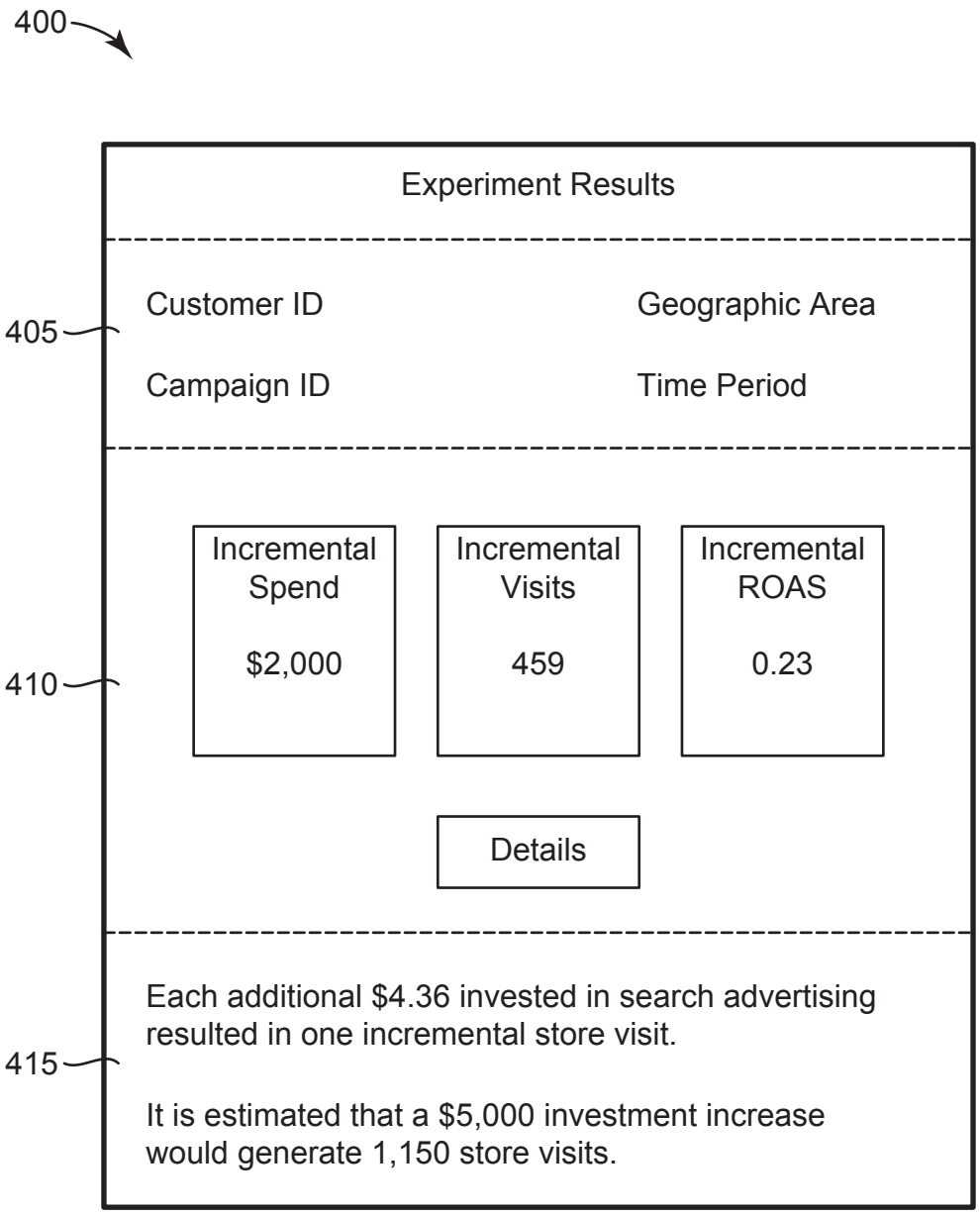
400
Experiment Results
Customer ID
Geographic Area
405
Campaign ID
Time Period
Incremental Spend
$2,000
Incremental Visits
459
Incremental ROAS
0.23
410
Details
Each additional $4.36 invested in search advertising resulted in one incremental store visit.
415
It is estimated that a $5,000 investment increase would generate 1,150 store visits.

FIG. 4

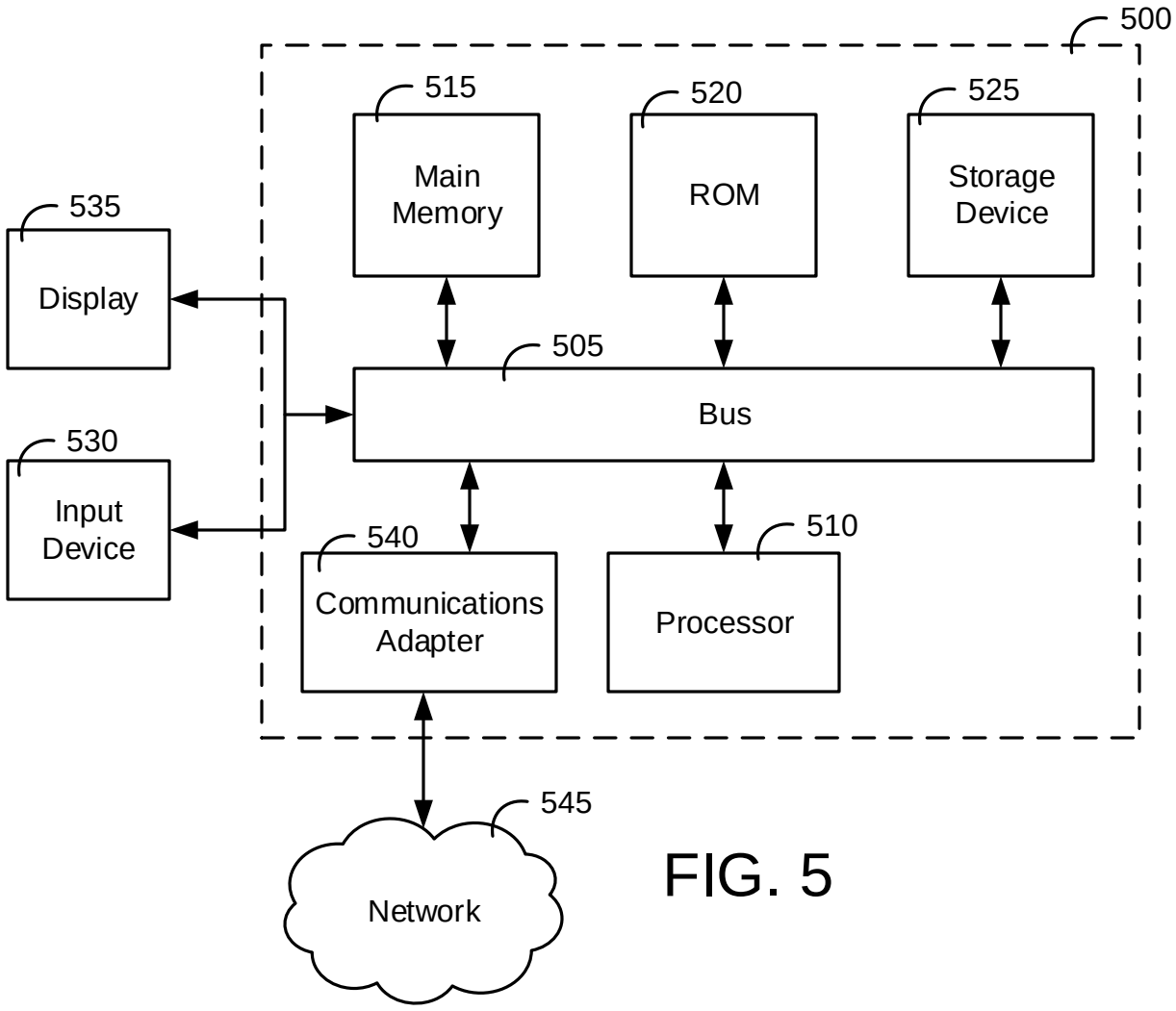
500
515
Main Memory
520
ROM
525
Storage Device
535
Display
505
Bus
530
Input Device
540
Communications Adapter
510
Processor
545
Network

FIG. 5