Predicting Human Movement Behaviors Using Location Services Model

Background

[0001] People in today's society have many different modes of transportation available to them, and many factors that determine when, where, and how each individual will travel from one place to another. In many cases, a person's movements have a set of defined, rational criteria, but those criteria can vary from person to person and from situation to situation.

[0002] When human movements can be predicted accurately, transportation systems may be studied to answer many interesting questions. For example, how might a transportation network be affected by a temporary or permanent change in design, where would be a good location for a new store, or how would policy decisions for a workforce affect transportation network loading?

Summary

[0003] A location services model may include behavior models of individual people that may simulate movement in a network. The behavior models may be derived from many data sources, including cellular telephone data, and may predict a person's behavior and responses to various inputs, as well as people's affinities to hobbies, interests, brands, and other items. The behavior models may further include direct and indirect relationships between people. Analysis may be performed using measured data to tune and confirm a model, then simulations performed with perturbations, additions, deletions, or other changes made to the model to predict behavior changes.

[0004] This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used to limit the scope of the claimed subject matter.

Brief Description of the Drawings

[0005] In the drawings,

[0006] FIGURE 1 is a diagram illustration of an embodiment showing a method for generating a knowledge graph from various databases.

[0007] FIGURE 2 is a diagram illustration of an embodiment showing a network environment with devices that gather data and create a knowledge graph from those data.

[0008] FIGURE 3 is a diagram illustration of an embodiment showing various applications that may be delivered using graph data.

[0009] FIGURE 4 is a diagram illustration of an embodiment showing an example graph of nodes and edges representing ontological elements and their relationships.

[0010] FIGURE 5 is a timeline illustration of an embodiment showing a method for gathering data.

[0011] FIGURE 6 is a flowchart illustration of an embodiment showing a method for adding or augmenting a knowledge graph with objects collected from a database.

[0012] FIGURE 7 is a flowchart illustration of an embodiment showing a method for using knowledge analyzers to infer and strengthen relationships in a graph.

[0013] FIGURE 8 is a diagram illustration of an example embodiment showing several types of nodes and examples of those types.

[0014] FIGURE 9 is a diagram illustration of an example embodiment showing an activities grid table.

[0015] FIGURE 10A is a diagram illustration of an example embodiment showing relationships.

[0016] FIGURE 10B is a diagram illustration of a second example embodiment showing relationships.

[0017] FIGURE 11A is a diagram illustration of an example embodiment showing a graph with relationships.

[0018] FIGURE 11B is a diagram illustration of an example embodiment showing a computed graph with computed relationships.

[0019] FIGURE 12 is a diagram illustration of an example embodiment showing interactions when updating and using a model.

[0020] FIGURE 13 is a flowchart illustration of an example embodiment showing a method for updating a graph.

[0021] FIGURE 14 is a flowchart illustration of an example embodiment showing a method for updating a graph with comprehensive data.

[0022] FIGURE 15 is a flowchart illustration of an example embodiment showing a method for performing analysis on a network representation.

[0023] FIGURE 16 is a diagram illustration of an example embodiment showing an analysis of a transportation network.

Detailed Description

[0024] Predicting Human Movement Behaviors Using Location Services Model

[0025] Human behavior may be predicted by building a model of people's physical movements along with a model that represents their affinities to various ontological elements, and their relationships to other people. The model may include a graph of interrelated people, as well as their hobbies, interests, employment, and other elements.

[0026] The model may be analyzed by simulating people's activities as those activities interact with real or hypothetical networks. The real networks may be used to verify the model by comparing measured parameters to predictive parameters derived from simulation to calibrate the models. A calibrated model may then be used with a modified or hypothetical network to analyze the effects of changes to the real network.

[0027] In one use scenario, a model may simulate the commuting habits of a city. The model may include behavior models of individual people in the city, and the model may be verified by measuring footfall traffic at various subway stations. A hypothetical model may be analyzed that includes a new subway line, temporary construction along a corridor, or even the effects of a sporting event on the transportation system.

[0028] The model may be updated periodically. Data from user's cellular telephones, web or television browsing habits, subway usage, or other sources may be periodically added to the model. Such data may be used to add new people to the model, as well as to update models of existing people.

[0029] Knowledge Model for Personalization and Location Services

[0030] A knowledge model may be constructed by aggregating data from many different sources into a unified graph where nodes represent ontological elements of user's lives, and edges represent relationships between the ontological elements. The graph may be built using metadata gathered from many different communications networks, including mobile telephony, internet access, television and other media, and various geo-location information. The graph may be supplemented by other data sources, and the resultant graph may be mined to identify implied or indirect relationships that may be inferred from the graph. The graph may be analyzed to provide various services, including geo-location services, demographic services, recommendation services, and other services.

[0031] The knowledge model may have some elements and relationships that are transitory or time-based, while other elements may be longer lived. The time-based elements may fall into several categories, including ephemeral relationships that may occur for a period of time and then degrade or end all together, as well as recurring relationships that may depend on time of day, day of week, or have other seasonality.

[0032] Examples of time-based relationships may include the relationship between a person and their employer. Such a relationship may have a recurring element that reflects the time the person spends at their job throughout the day, and such a relationship may terminate abruptly when the person moves to a different job. A person may also develop a time-based relationship with a certain hobby, which a person may enjoy intensely for a period of time, then the intensity may slowly degrade as the person's interests change to a different hobby or pastime.

[0033] The knowledge model may have various modifiers on the observed or implied relationships. The modifiers may reflect the strength of a relationship, which may be increased by the number of repeated observations or by applying a strength relationship using a heuristic or other algorithm.

[0034] A use scenario for a knowledge model may include the following: a graph may be constructed that combines metadata from a user's interactions with a mobile telephone service with the user's interactions with the world wide web and the same user's media consumption habits. Metadata from the mobile telephone network may provide time of day information and person-to-person interactions. From these metadata, relationships between the user and other entities, such as friends and family, coworkers, and various businesses may be uncovered. The user's home and work schedules may also be uncovered from these metadata. From the user's media consumption habits, the user's interests may be uncovered.

[0035] In the use scenario, interesting and useful queries may be performed. For example, demographic information about certain locations may be uncovered. Based on the location of a potential restaurant near many officer workers, what are the demographics of people likely to have lunch in the area? What will the demographics be in the evening, when people go out to dine and socialize? Such information may help a restaurateur select appropriate décor, menu, and other items to attract the people who are in the area.

[0036] In another query, key influencers for a specific brand's products may be identified within a neighborhood. Those influencers may be awarded coupons, special promotions, samples of new products, or other items that may further cultivate the influencer's interest in a brand's products. The brand's highly targeted marketing strategy may focus on these people who may generate a much higher return for the advertising investment than traditional, broad-stroke advertising.

[0037] In yet another query, a recommendation engine may suggest things to do, restaurants, media, and other items to a user that may be of interest. The recommendation engine may tailor the recommendations based on a user's likes and dislikes. For example, a person who enjoys cooking and food, which may be identified from their media viewing habits, may be shown opportunities to attend a cooking class when the user is in a new location away from work and home. Such a location may indicate that the user is on holiday, and such an assumption may be buttressed by the user's contact with various travel planning businesses.

[0038] When personally identifiable information (PII) is collected and/or processed in the various databases and use scenarios (including those described below), users will be (a) requested to give consent and/or (b) given an opportunity to opt in or opt out of such data processing, whichever is applicable in compliance with applicable data protection laws. In many embodiments, PII will usually be properly obfuscated prior to being aggregated into a graph representation with other data sources. Various other mechanisms or measures may be used to ensure that PII is not collected, stored, transmitted, analyzed or otherwise processed in any manner not in compliance with applicable data protection laws.

[0039] Throughout this specification, like reference numbers signify the same elements throughout the description of the figures.

[0040] In the specification and claims, references to "a processor" include multiple processors. In some cases, a process that may be performed by "a processor" may be actually performed by multiple processors on the same device or on different devices. For the purposes of this specification and claims, any reference to "a processor" shall include multiple processors, which may be on the same device or different devices, unless expressly specified otherwise.

[0041] When elements are referred to as being "connected" or "coupled," the elements can be directly connected or coupled together or one or more intervening elements may also be present. In contrast, when elements are referred to as being "directly connected" or "directly coupled," there are no intervening elements present.

[0042] The subject matter may be embodied as devices, systems, methods, and/or computer program products. Accordingly, some or all of the subject matter may be embodied in hardware and/or in software (including firmware, resident software, micro-code, state machines, gate arrays, etc.) Furthermore, the subject matter may take the form of a computer program product on a computer-usable or computer-readable storage medium having computer-usable or computer-readable program code embodied in the medium for use by or in connection with an instruction execution system. In the context of this document, a computer-usable or computer-readable medium may be any medium that can contain, store, communicate, propagate, or transport the program for use by or in connection with the instruction execution system, apparatus, or device.

[0043] The computer-usable or computer-readable medium may be, for example but not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, device, or propagation medium. By way of example, and not limitation, computer readable media may comprise computer storage media and communication media.

[0044] Computer storage media includes volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can accessed by an instruction execution system. Note that the computer-usable or computer-readable medium could be paper or another suitable medium upon which the program is printed, as the program can be electronically captured, via, for instance, optical scanning of the paper or other medium, then compiled, interpreted, of otherwise processed in a suitable manner, if necessary, and then stored in a computer memory.

[0045] When the subject matter is embodied in the general context of computer-executable instructions, the embodiment may comprise program modules, executed by one or more systems, computers, or other devices. Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Typically, the functionality of the program modules may be combined or distributed as desired in various embodiments.

[0046] Figure 1 is an illustration showing creation of a market knowledge graph. The market knowledge graph may be created from multiple data sources, which may be processed through an information extraction process, then again through a knowledge formation process to create a graph 124. Once the graph 124 is created, various queries may be performed against it using a query engine 126.

[0047] The market knowledge graph 124 may assemble many data elements to create a graph of various relationships between ontological elements, which may be items

such as people, places, things, brands, locations, media, and other items. The relationships between the ontological elements may be strengthened, implied, or otherwise modified by knowledge formation processes, which may include heuristics or other analyses.

[0048] Various data collectors 102 may gather data from many different sources. The data sources may be various media or communication networks, include cellular telephone networks, internet service provider networks, digital television media networks, or many other data sources.

[0049] Examples of the data sources may include location data 104, which may be gathered through any device that a person may have that can detect location. For example, a cellular telephone or other mobile device may be able to detect location through an internal Global Positioning System (GPS) receiver. In many cases, a mobile device may be able to determine its location through triangulation of wireless access points, cellular telephony system, or other mechanism.

[0050] The location data 104 may be associated with a specific user. For a cellular telephone, the location of the device may assumed to be the same location as the subscriber to whom the telephone belongs. Many other mobile devices may also gather location data 104, such as tablet and laptop computers, personal fitness devices, or other wearable or portable computing devices.

[0051] Many other devices may have a linkage between a specific person and a location. For example, an automobile, bus, train, airplane, boat, or other transportation vehicle may be associated with people in the vehicle and the route and timing of the vehicle. In such cases, the location data 104 may be culled from a transportation system's computer, which may track which passengers are on which vehicle. Such databases may be provided by airlines, ferry operators, transit system operators, or other transportation system.

[0052] Call metadata 106 may reflect any communication metadata between two or more people. The call metadata 106 may be gathered from a mobile telephone operator, for example, where various metadata for each call may be collected. The metadata may identify the origin and receiving stations, the time and duration of the call,

and other information. The call metadata 106 may additionally include location data 104 in some cases.

[0053] The call metadata 106 may not include content or details of the communication between users, but may only include various metadata.

[0054] Weblogs 108 may include metadata for a user's browsing history or other communications over the world wide web or other data network. The weblogs 108 may include the time, duration, and other connection parameters for communication sessions for a computer. While the contents or payloads of a communication may not be gathered, the location, timing, and direction of the communication may be.

[0055] Digital television logs 110 may include a user's viewing history of a digital television system. In many cases, the digital television system may have a digital video recorder, which may capture broadcast video for later viewing. The digital television logs 110 may include the shows that were viewed on the system, as well as the shows that were captured and various parameters about the replay sessions for the recorded shows.

[0056] Other data sources 112 may include metadata gathered from game consoles, electronic reader platforms, or other systems. The other data sources 112 may also include information from electronic social networks or any other source of various ontological elements or relationships of the ontological elements that may be available.

[0057] An information extraction process 114 may process the various data sources to identify nodes and edges for the graph 124. The nodes may represent an ontological element and the edges may represent relationships between the ontological elements.

[0058] The ontological elements may be any element that may relate to human beings and their surroundings, such as individual people, their homes, places of work, businesses that supply goods and services, media outlets, media content, hobbies, interests, or any other element. The type of elements in a graph 124 may depend on the type of raw data that may be available as well as the types of queries that may be supported by a query engine 126.

[0059] Various items of interest may be extracted from the first layer of information extraction 114. Various destinations 116 may be identified from the various

location data 104. The destinations 116 may include places of work or recreation, homes, vacation locations, business locations including restaurants and shopping, and other destinations.

[0060] A social network 118 may be constructed from the various data sources that may link various people together by different affiliations. Such affiliations may include familial relationships, as well as coworker relationships, religious and other social group relationships, and other person to person relationships.

[0061] Interests 120 may include interests in various topics as well as intents or goals for a person. An interest may be a work related topic, such as computer programming or project management, as well as a hobby or leisure interest, such as travel, cooking, or gardening. The interests 120 may also include companies, brands, or their products that may be of interest to a person.

[0062] Knowledge formation 122 may be a secondary analysis of the information extracted from the raw data. The knowledge formation 122 may perform various tasks such as inferring relationships, prediction, and otherwise further enhancing the graph 124. The knowledge formation 122 may be performed using heuristics or other algorithms. In some cases, such algorithms may be developed for specific types of inferred or enhanced relationships.

[0063] A query engine 126 may perform queries against the graph 124. Examples of various queries may be discussed later in this specification.

[0064] Figure 2 is a diagram of an embodiment 200 showing components that may collect data to generate a graph showing ontological elements and relationships between those elements. The components are illustrated as being on different hardware platforms as merely one example topology.

[0065] The diagram of Figure 2 illustrates functional components of a system. In some cases, the component may be a hardware component, a software component, or a combination of hardware and software. Some of the components may be application level software, while other components may be execution environment level components. In some cases, the connection of one component to another may be a close connection where two or more components are operating on a single hardware platform. In other cases, the connections may be made over network connections spanning long distances.

Each embodiment may use different hardware, software, and interconnection architectures to achieve the functions described.

[0066] Embodiment 200 illustrates a device 202 that may have a hardware platform 204 and various software components. The device 202 as illustrated represents a conventional computing device, although other embodiments may have different configurations, architectures, or components.

[0067] In many embodiments, the device 202 may be a server computer. In some embodiments, the device 202 may still also be a desktop computer, laptop computer, netbook computer, tablet or slate computer, wireless handset, cellular telephone, game console or any other type of computing device. In some embodiments, the device 202 may be implemented on a cluster of computing devices, which may be a group of physical or virtual machines.

[0068] The hardware platform 204 may include a processor 208, random access memory 210, and nonvolatile storage 212. The hardware platform 204 may also include a user interface 214 and network interface 216.

[0069] The random access memory 210 may be storage that contains data objects and executable code that can be quickly accessed by the processors 208. In many embodiments, the random access memory 210 may have a high-speed bus connecting the memory 210 to the processors 208.

[0070] The nonvolatile storage 212 may be storage that persists after the device 202 is shut down. The nonvolatile storage 212 may be any type of storage device, including hard disk, solid state memory devices, magnetic tape, optical storage, or other type of storage. The nonvolatile storage 212 may be read only or read/write capable. In some embodiments, the nonvolatile storage 212 may be cloud based, network storage, or other storage that may be accessed over a network connection.

[0071] The user interface 214 may be any type of hardware capable of displaying output and receiving input from a user. In many cases, the output display may be a graphical display monitor, although output devices may include lights and other visual output, audio output, kinetic actuator output, as well as other output devices. Conventional input devices may include keyboards and pointing devices such as a mouse, stylus, trackball, or other pointing device. Other input devices may include various

sensors, including biometric input devices, audio and video input devices, and other sensors.

[0072] The network interface 216 may be any type of connection to another computer. In many embodiments, the network interface 216 may be a wired Ethernet connection. Other embodiments may include wired or wireless connections over various communication protocols.

[0073] The software components 206 may include an operating system 218 on which various software components and services may operate.

[0074] A graph builder application 220 may consist of database-specific extractors 222 and knowledge analyzers 224. The graph builder application 220 may create, update, and maintain a graph 226 that may be queried by a query engine 228.

[0075] The graph builder application 220 may update the graph 228 as new data are received. In many deployments, a graph builder application 220 may be continually updating and building the graph 226 in response to receiving data in real time or near-real time.

[0076] The database-specific extractors 220 may cull data from various databases and update or build the graph 226. For example, a database extractor may be created for a cellular telephone network metadata database and a second database extractor may be created for analyzing weblogs. Both extractors may add nodes and edges to the graph 226.

[0077] The knowledge analyzers 224 may be algorithms, routines, or other mechanisms to apply different rules to the data. The rules may infer relationships, as well as strengthen or weaken relationships based on the presence or absence of different elements defined in the algorithm. A simple example of a knowledge analyzer may add a coworker relationship between two people who work at the same place of business. Such a relationship may not be identified from the raw data, but the knowledge analyzer may create the relationship based on a heuristic or other algorithm.

[0078] The knowledge analyzers 224 may represent second level analyses that derive insights from the data. The knowledge analyzers 224 may be algorithms designed by data scientists with the help of domain knowledge experts, and may automatically generate insights from combinations of data sources. The insights may identify

relationships and relationship strengths that may not be present in a single data source. For example, a user that may watch a cooking show on television may be assumed to have a slight affinity for cooking as a hobby. When that user also patronizes high end restaurants and visits websites for haute cuisine cooking schools, that user may have a deep interest in cooking at a professional level. The combination of data from different sources may strengthen a relationship from a mere passing interest to a serious hobby or professional interest.

[0079] The query engine 228 may receive various requests and return data derived from the graph 226. The query engine 228 may have a visual user interface where a user may be able to visualize and interact with the graph. The query engine 228 may have an application programming interface (API) that may respond to programmatic requests from applications. The applications may use the query engine 228 to provide data for subsequent analyses or other purposes.

[0080] A network 230 may connect the device 202 to other systems. In some cases, the network 230 may be the Internet, a local area network, a wide area network, a wired network, a wireless network, some other network, or a combination of different networks.

[0081] Some of the data in the graph 226 may be derived from the operations of a communications network 232. The communications network 232 may be, for example, a wireless network where users communicate with mobile telephones. The communications network 232 may be a wired telephony network, a pager network, a private exchange network, a wireless radio network, or any other type of network.

[0082] A monitoring and control system 234 may manage the communications network 232. The monitoring and control system 234 may perform routing, network management, authentication, accounting, and administration functions, or other functions. Operational metadata 236 may be collected from the monitoring and control system 234. A data collector 238 may gather the metadata and transmit the metadata to a corresponding database specific extractor 222 to add these data to the graph 226.

[0083] Similarly, an internet service provider 240 may have a monitoring and control system 242. The internet service provider 240 may provide internet connections to various end users, and the monitoring and control system 242 may gather operational

metadata 244 during various administrative operations. The operational metadata 244 may include connections between users and various websites or services. A data collector 246 may and transmit the operational metadata 244 to a corresponding database specific extractor 222 to add these data to the graph 226.

[0084] Many other data sources 248 may be used to build or supplement the graph 226. The other data sources may be any data source which may relate to objects stored in the graph 226. For example, a data source may include a description or analysis of a television show that may define the topics of the show. The topics may be used to establish a relationship between a user's viewing of the show and the user's interests in the various topics. Such relationships may be inferred using the combination of data sources.

[0085] A client device 250 may be any device that may perform requests to the query engine 228 as part of an application. The client device 250 may have a hardware platform 252 which may be similar to that described for the hardware platform 204. Various applications 254 may perform queries against the graph 226 by sending a request to the query engine 228 and receiving the results of the query for further display or processing.

[0086] Figure 3 is a diagram illustration of an example embodiment 300 showing applications that may be delivered from a graph derived from multiple data sources. The graph 302 may represent the graphs 124 or 226 from embodiments 100 or 200, respectively. Embodiment 300 may illustrate two levels of analysis or applications that may be drawn from the graph 302.

[0087] A query engine 304 may perform various queries against the graph 302. From these queries, several analyses may be performed, such as user and location profiling 306, market segmentation 308, mobility pattern analysis 310, social mining 312, and others.

[0088] User and location profiling 306 may develop profiles of various users or locations. For example, a user profile may identify a user's likes, dislikes, and other characteristics. The profile may include any available data, such as a person's job, commute route and time, their hobbies, interests, likely purchasing habits, estimated purchasing power, and other information.

[0089] A location profile may include demographic information about the people who visit a location, along with their interests, occupations, and other factors.

[0090] Time-based factors may be present in some graphs. A time-based factor may identify certain parameters that vary with the time of day, day of week, season, holidays, or other interval. A query for a time-based factor may include a date and time parameter and a query engine may return a predicted or estimated value for a requested parameter that meets the date and time requested.

[0091] An example of a time-based factor may be the location of employees at a company. During a typical work day, the employees may be located at a place of business, but on a non-work day, the employees may be located at home or at a place of recreation.

[0092] Predictive models may be constructed for user and location profiles, along with many other data. The predictive models may be constructed by analyzing the time and endurance of various events, then determining a statistical model of those events. A predictive model may estimate the time-based factor using the statistical model for a given time or range of times.

[0093] Market segmentation 308 may be an analysis that identifies the components of a given market. An example may be to define the types of people who enjoy cooking as a hobby. The data may include a breakdown of the ages, occupations, and other hobbies enjoyed by this segment of population. Some factors may be time-based factors, such as the time of day that people engage media that are cooking related, such as cooking based websites or television shows.

[0094] Mobility pattern analysis 310 may be an analysis that identifies the flow patterns of people or other objects through various locations. Mobility pattern analysis may include traffic delays during rush hour as well as the flow and demographics of people through a shopping center or along a commercial road.

[0095] Social mining 312 may identify connections between individuals or groups of people. Such analysis may yield a rich demographic profile of groups of people, with their interests, relationships, affinity to various brands or products, and many other data points.

[0096] The various analyses such as user and location profiling 306, market segmentation 308, mobility pattern analysis 310, and social mining 312 may be combined into many different applications. Examples of such applications include various recommender systems 314, personalization services 316, market research 318, targeted advertising 320, and geo-intelligence services 322.

[0097] Recommender systems 314 may provide any type of recommendation for a person with a given profile. In the commercial world, there are many choices for different products. A recommender system 314 may narrow the choices for a person given their individual profile. For example, a person who enjoys many outdoor activities may be shown products that were selected by other people who also enjoy outdoor activities. A different person who enjoys indoor activities may be shown different products.

[0098] Recommender systems 314 may benefit from the diversity of data sources that generate the graph 302. For example, a metadata analysis of cellular telephone metadata may identify a user living in a certain neighborhood and commuting to work at a high technology company. Such analysis may begin to build a demographic profile of the user. Coupling that data with weblog analysis may uncover that the user may enjoy endurance athletics along with outdoor sports. A recommendation system may identify a demographic profile for a new person, which may be similar to the demographic profile of the analyzed user. The recommendation system may return outdoor sports information that may be interesting to the user because of the relationship uncovered in the metadata analysis.

[0099] A personalization service 316 may be any type of service that may customize or personalize an experience for a user. For example, a user who may be placing an online order from a restaurant may be presented with choices that are popular with other users who share the same interests and demographic profiles.

[00100] Market research services 318 may attempt to classify, categorize, and understand the interactions of customers and products. A demographic study may classify people who may be interested in certain products or classes of products, and may show detailed analyses that may include any type of relationship or ontological element

that may be related to the product or people. A product study may show which types of people may be interested in a certain brand or type of product.

[00101] Targeted advertising services 320 may be advertising mechanisms that can target people with very specific demographics and interests. The demographics, interests, and other factors for a user may be gathered from the user's profile as defined in the graph 302. When an advertising opportunity is identified, such as an online advertisement, customized outdoor advertisement, or other opportunity, elements of a user's profile may be sent to an advertisement arbitrage system that may match the user's profile with an advertiser that may be trying to reach that type of user. The arbitrage system may return an advertisement that may be targeted for that user, making the interaction more useful for the user, as well as potentially profitable for the advertiser.

[00102] A geo-intelligence service 322 may provide details about locations. Such information may assist a business person in evaluating current or potential business opportunities in a geographical area. For example, a shopkeeper or restaurateur may wish to know the demographic of potential customers that pass by a place of business. Based on the geo-intelligence results, a shopkeeper may stock certain items tailored to the demographic. Similarly, a restaurateur may optimize a menu or offer specials that cater to the tastes of potential patrons.

[00103] The geo-intelligence example may also be one example of time-based data, where the shopkeeper or restaurateur may wish to understand traffic patterns and demographics based on time of day, day of week, or other time-dependent parameters.

[00104] Figure 4 is a diagram illustration of an embodiment 400 showing an example graph of relationships between ontological elements. Embodiment 400 is a very simplified example of a graph that may be constructed from various databases and the relationships that may be inferred from the data.

[00105] The graph may illustrate three people, James 402, David 404, Sally 406, and Nancy 408. James 402 and David 404 may have a friend relationship 422 because James and David talk on the phone or send text messages. A similar friend relationship 418 may be uncovered between James 402 and Nancy 408. These relationships may be derived from cellular telephone metadata.

[00106] Location analysis of James 402 and Sally 406 may indicate that both people are at a work location 410 at the same time. Both persons may attend their job at relatively the same time and on a daily basis, indicating that they both have work relationships 412 and 414 to the work location 410. An inferred coworker or colleague relationship 416 may be discovered by knowledge analysis, where a knowledge analysis routine may establish implied coworker relationships between people who work at the same location.

[00107] Sally 406 and David 404 may share a home location 424. The home location 424 may be identified by matching a geo-location with an address lookup database that may identify the geo-location coordinates as a residential area. Sally's location at the home location 424 may coincide with their time away from work location 410.

[00108] A cooking channel 428 may be viewed from a digital television device at the home 424. In some cases, the precise person who may be using a device may be unknown. However, David 404 may have visited a cooking website 426, which may infer that David 404 has an interest in cooking. This inference may be augmented by both the visit to the cooking website 426 as well as viewing the cooking channel 428, even though it may be unknown who in the home location 424 was actually viewing the cooking channel 428.

[00109] Figure 5 is a timeline illustration of an embodiment 500 showing a method for collecting data. Embodiment 500 may illustrate the operations of a client device 502 which may act as a data collector, and a server device 504, which may receive and process the data.

[00110] Other embodiments may use different sequencing, additional or fewer steps, and different nomenclature or terminology to accomplish similar functions. In some embodiments, various operations or set of operations may be performed in parallel with other operations, either in a synchronous or asynchronous manner. The steps selected here were chosen to illustrate some principals of operations in a simplified form.

[00111] Embodiment 500 may illustrate a simple example of interactions that may be performed between a device that collects data and a server device that may receive and process the data.

[00112] A client device 502 may gather data on the device in block 506, anonymize data in block 508, and transmit the data to a server in block 510.

[00113] The server device 504 may receive the data in block 512 and store the data for later processing in block 514.

[00114] The client device may perform various pre-processing steps in block 508, including anonymizing the data. Some data that may be collected may include PII, and such PII may be scrubbed or anonymised from the data prior to transmitting and storing the data. Several mechanisms may be used to anonymize the data, including randomization, masking or other mechanisms. In some embodiments, anonymizing of the data may occur at a server device prior to processing.

[00115] Figure 6 is a flowchart illustration of an embodiment 600 showing a method for adding or augmenting a graph with objects in a database. Embodiment 600 may illustrate a first level of data extraction that may be performed when receiving data from a database.

[00116] Other embodiments may use different sequencing, additional or fewer steps, and different nomenclature or terminology to accomplish similar functions. In some embodiments, various operations or set of operations may be performed in parallel with other operations, either in a synchronous or asynchronous manner. The steps selected here were chosen to illustrate some principals of operations in a simplified form.

[00117] Embodiment 600 may illustrate a process that may be performed to initially populate or augment a graph. The process may be performed on a repeating basis to add new nodes or edges to a graph, where the nodes may represent ontological elements, such as a person, location, interest, brand, product, or other element, and the edges may represent relationships between those elements. In many embodiments, a relationship may include a strength relationship, where the relationship may be strengthened or weakened based on additional data points.

[00118] The graph building or augmentation may begin in block 602.

[00119] Each database for which new data are available may be processed in block 604. For each database in block 604, a connection may be established to the database in block 606.

[00120] Each object in the database may be evaluated in block 608. For each object in block 608, a corresponding instance of an ontological element may be identified in the graph in block 610. If the object is not found in the graph in block 612, a new node may be created in the graph in block 614. The new node may represent a new instance of an ontological element in the graph.

[00121] The ontological element of a graph may be identified by matching a schema defining the database being analyzed with a schema of ontological elements of the graph. In some cases, an object in a database may not have a corresponding element in the graph and the object may be discarded.

[00122] The existing or potential relationships between the node and other nodes may be identified in block 616. For each of the relationships in block 618, if the relationship does not exist in block 620, a new relationship may be created in block 622. When the relationship does exist in block 620, the existing relationship may be strengthened in block 622.

[00123] The relationships identified in block 616 may be common relationships that may be derived directly from the data being analyzed. For example, if a person is physically located at a restaurant, nodes may be generated representing the person and the restaurant and a relationship may be established between the two nodes.

[00124] After processing each object in each database in block 608 and each database in block 604, the process may proceed to knowledge analysis in block 626.

[00125] Figure 7 is a flowchart illustration of an embodiment 700 showing a method for applying knowledge analyzers to graph data. Embodiment 700 may illustrate a second level of analysis where inferences may be made on the data combined from multiple databases.

[00126] Other embodiments may use different sequencing, additional or fewer steps, and different nomenclature or terminology to accomplish similar functions. In some embodiments, various operations or set of operations may be performed in parallel with other operations, either in a synchronous or asynchronous manner. The steps selected here were chosen to illustrate some principals of operations in a simplified form.

[00127] Embodiment 700 may apply algorithms or heuristics to identify implied relationships between elements in the graph. Embodiment 700 may illustrate a system

where multiple knowledge analyzers may be applied to a graph. Such an embodiment may be expandable to accept new knowledge analyzers that may be developed by data scientists or domain experts to address different insights that may be gleaned from the data.

[00128] Embodiment 700 is merely one example of an architecture that may perform knowledge analysis on a graph. In the example of embodiment 700, the knowledge analyzers may operate by finding a first element or target, then searching for a second element that may be near the target. When the second element is present, a relationship may be inferred.

[00129] In a simple example of such a knowledge analyzer, a person who attends a work location on a regular basis may have an inferred work relationship between the company at the work location and the person. The first target element may be a place of work, and the second element may be a person who visits the location for most of the normal business hours.

[00130] Other knowledge analyzers may identify two, three, four, or more elements that may be present for an inferred relationship to occur. In the example of embodiment 700, a knowledge analyzer is illustrated that infers a relationship between two elements.

[00131] The knowledge analysis may begin in block 702. Each knowledge analyzer may be processed in block 704. For each knowledge analyzer in block 704, a target element type may be identified in block 706.

[00132] A search may be made in block 708 for the target element, and each target element may be analyzed in block 710. For each target element in block 710, a second element type may be identified in block 712 and searched in block 714.

[00133] Each of the second elements found in block 714 may be analyzed in block 716. For each of the second elements in block 716, if the relationship between the target element and the second element is not found in block 718, the relationship may be created in block 720. When the relationship is found in block 718, the relationship may be strengthened in block 722.

[00134] After processing all of the knowledge analyzers in block 704, the graph may be ready for queries in block 724.

[00135] Figure 8 is a diagram illustration of an embodiment 800 showing examples of nodes that may be present in a graph. Embodiment 800 is merely a simplified example of the different types of nodes and examples of these nodes.

[00136] Location nodes 802 may have various properties, such as latitude and longitude pairs, names, polygons describing a boundary, addresses, and other parameters. Examples of location nodes 802 may be street segments 804, country 806, airport 808, park 810, shopping mall 812, subway station 814, and other locations.

[00137] User nodes 816 may have properties such as age and gender. Example nodes may be James 818, Amy 820, and Keen 822. In many embodiments, the user nodes may be anonymized or obfuscated so that PII may not be present in the graph. The identifiers of James 818, Amy 820, and Keen 822 may be placeholders or labels that may be associated with an anonymized or obfuscated identity and may not relate directly to any specific person.

[00138] A category node 824 may be any category of interest. Typically, a category may be a parameter for which searches may be performed or for which relationships may be inferred. Categories derived from various data sources, such as a television or movie database that may classify television shows or movies into categories.

[00139] Examples of different categories may be cultural 826, shopping 828, food 830, travel 832, industrial 834, healthcare 836, leisure 838, sport 840, and others.

[00140] Product nodes 842 may have properties such as name, price, language, medium, and other properties. A product may be a hard good, service, software, media content, or any other type of service. Examples may include a car 844, taxi 846, shoes 848, massage 850, social network 852, office supplies 854, and other products.

[00141] Brand nodes 856 may represent a company, product line, or other aggregation of products. Examples may be Adidas 858, BMW 860, Singapore Airlines 862, Facebook 864, and others.

[00142] Figure 9 is a diagram illustration of an example embodiment 900 showing an activities grid table. An activities grid table may illustrate a time-based interaction between to nodes in a graph. In the example of embodiment 900, the interaction may be between a user and a location, but this type of time-based interaction may be used for any type of relationship.

[00143] Relationships between any two nodes in a graph may be characterized by different time-based interactions. Some time-based relationships may be a periodic relationship, such as the times that an employee attends their place of work. Such relationships may have a distinct periodicity and may be summarized into a periodic time-based model.

[00144] Some relationships may persist over a long period of time and may have distinct beginning and ending events. In an employment relationship, for example, an employee's hiring and firing dates may mark calendar events that begin and end that relationship. After a separation event, employment relationships may not have much residual interaction.

[00145] Some relationships may persist over time and may degrade smoothly or change abruptly in response to some event. For example, a person may enjoy following a sports team. The person's enthusiasm for the team may swell in the initial stages and may persist at a near steady state for some period of time. An event may occur, such as mediocre performance of the team, and the person's enthusiasm may slowly decline. In some cases, the person's enthusiasm may decline rapidly if, for example, the team were to unceremoniously trade a star player. A person's relationships with brands, products, and other categories may have a similar ebb and flow.

[00146] In the activities grid table 900, the day of week 904 is displayed in the vertical axis and hour of day 906 in the horizontal axis. A grid of each hour of the week is formed, and each hour may have a value that may represent the strength of an interaction. In the example of embodiment 900, the interaction may simulate a worker's time spent at work. During normal business hours, the worker may be present at a place of employment. Some days, the worker may come in late and stay late, but in general the worker may be present from 8am to 5pm on Monday through Friday.

[00147] In the example of a worker and their place of employment, the relationship made be defined after observing the worker's behavior over many weeks, months, or years. The strength of relationship for a particular hour may be increased as the number of observations increases for that hour.

[00148] A relationship may be strengthened or weakened by the intensity of interactions. Some interactions may indicate a stronger relationship than others. For

example, a user may watch merely one portion of a television show, which may indicate a lackluster and weak relationship. A second user may watch the entirety of the show, visit the show's website, and engage on social media relating to the show. The second user's interaction level is very strong, while the first one is weak. In such cases, a strength of a relationship may be determined by specific actions, each of which may have a different strength and may be aggregated by summing or other heuristic to determine a strength value.

[00149] Figures 10A and 10B are diagram illustrations of example relationships 1000 and 1002 that may show relationships, some of which may vary over time.

[00150] The example relationships 1000 may illustrate relationships between a physical mobile user 1004 with various objects. From the objects, a relationship may be inferred to different categories. The examples of embodiments 1000 and 1002 use the notion of a physical mobile user. This term is meant to mean a physical person that may be detected using data gathered from mobile phone metadata. This is merely one example of the type of data that may be analyzed.

[00151] A physical mobile user 1004 may have various relationships, such as with a physical object 1006, a virtual game or application 1008, a physical service 1010, and a website 1012. The relationship with a physical object 1006, which may be a piece of sports equipment, may be defined by an interest relationship 1014. The interest relationship 1014 may be identified when the user scans a UPC barcode of an object when the user visits a sporting goods store, for example. Such a relationship may be a single event at a point in time. When the user plays a virtual game or application 1008, the usage relationship 1016 may have a time-based relationship that may include the time of day the user plays the game, the number and duration of playing sessions, and in some cases the intensity or interaction the user has while playing the game.

[00152] The user may request a physical service 1010, such as requesting a taxi, by placing a telephone call 1018. The time-based factors for such a relationship may be the time of day, day of week, season, proximity to a personal or calendar holiday, or some other factor. Similarly, a user may visit a website 1012, which may afford many different metrics for capturing engagement. Such metrics may include the time of day

and other related factors, as well as the pages viewed, requests submitted, or other metrics.

[00153] Relationships to categories may be established from each of the various objects, games, services, websites, or other items with which a user interacts. From the knowledge that the physical object 1006 relates to sports 1022, an inferred relationship between the physical mobile user 1004 is made to sports 1022. Similarly, the relationship between the application or game 1008 to music 1024 may infer the user's taste and affinity for music. The physical service 1010 has a relationship 1034 to transportation, so the user's preferences for transportation may be inferred. The user's level of interest in a website 1012 may infer a relationship 1036 to travel 1028, which may be the content of the website.

[00154] The example relationships 1002 may illustrate inferences that may be drawn from evaluating time dependent factors of a relationships. Three different mobile users 1038, 1040, and 1042 and how those users interact with a location of a building 1044.

[00155] User 1038 may be assumed to live 1046 in the building 1044 when the user's interaction pattern matches a conventional person's living habits. For example, the user may spend a large majority of their evenings and weekends at the building.

[00156] User 1040 may visit 1048 the building 1044 when the user's interactions are sporadic and do not have a consistent pattern of either a place of employment or a home.

[00157] User 1042 may have an inferred work 1050 relationship with the building 1044 when the user's visitation habits have the consistency and regularity of working hours.

[00158] Figures 11A and 11B are diagram illustrations of example graphs 1100 and 1102. Graph 1100 may illustrate raw data while graph 1102 may illustrate computed relationships that may be derived from the graph 1100.

[00159] In the full graph 1100, a user 1104 may be a user who views a cooking show 1106. A viewing relationship 1108 may be measured or determined from the user's viewing habits 1108. The user 1104 may visit a recipes website 1110, which may be measured from internet metadata as a visit relationship 1109. From the cooking show

1106, there may be a relationship 1112 to cooking, as well as a second relationship 1114 from the recipes website 1110 to cooking 1114.

[00160] The user 1104 may play the game Halo 1118, and a playing relationship 1120 may be detected. The game Halo 1118 may be related to the general category of gaming 1122 through a gaming relationship 1124.

[00161] In performing a query against the graph 1100, relationships between the user 1104 and the categories cooking 1114 and gaming 1122 may be given as computed relationships 1126 and 1128.

[00162] The computed relationship 1126 between the user 1104 and cooking 1114 may be computed by determining the viewing relationship 1108 of the television program and the television program's relationship 1112 to cooking. A second path of the user's visitation relationship 1109 to the recipes website 1110 and the websites' relationship 1116 to cooking 1114. The two paths may be aggregated by summing or applying another heuristic or algorithm.

[00163] The computed relationship 1128 between the user 1104 and gaming 1122 may be computing by determining the playing relationship 1120 to Halo 1118 and Halo's relationship 1124 to gaming 1122.

[00164] The computed graph 1102 may be determined in response to a query that may request the strength of relationships between various users and categories, such as cooking 1114 and gaming 1122.

[00165] Figure 12 is a diagram illustration of an embodiment 1200 showing various interactions with updating and using a behavioral model. Embodiment 1200 is merely one simplified example of various components that interact with a graph model 1202 of people to create, update, and use the model in different scenarios.

[00166] Embodiment 1200 is merely one example of the interactions that may occur with a graph model of users 1202 and various other components. The components may include perform maintenance, calibration, and analysis using the model. The maintenance operations may add new information and cull old or outdated information from the model. The calibration operations may include performing an analysis of real world conditions and comparing an estimated result from the model with actual results

measured from the real world. The analysis operations may analyze the model against changes from the real world to estimate effects that the changes may have.

[00167] The maintenance operations may be performed by a model manager 1206, which may use various data sources 1204 to update the graph model of users 1202. The data sources 1204 may be any type of data that may be related to the model, including information gleaned from cellular telephone movement and usage, interactions of people with the Internet or devices connected to the Internet, geophysical information, semantic or ontological data, or any other data source.

[00168] In many maintenance operations, a data item from a data source may be searched in the model. The new data item may be a node on the model. When the data item is found in the model, the model may be updated to reflect the model's confidence in the data item. The confidence may be reflected in a numerical confidence value that may be increased each time an update is performed and that data item is verified. If the data item is not found in the model, the model may be updated to add the data item in the model.

[00169] Some models may include a historical parameter that may include information on how data items in a model may change over time. In some cases, the changes may ebb and flow over time, such as a user's affinity for a certain hobby, brand, sports team, or other interest. In other cases, the changes may be distinct and complete, such as a user's place of employment when the user changes from one job to another. With each maintenance operation of the model, the changes to the model may be made to the model to store historical changes to various nodes and edges.

[00170] Updates to the model may be applied differently using heuristics or other algorithms that have a historical context. For example, an update to a user's employment location may be a binary change that removes an old employment location and inserts a new employment location, yet a user's interest in a sports team may be increased slightly when the user is identified as attending a sporting event.

[00171] In many instances, the data item from a data source may identify relationships or interactions between two nodes in the model. In such instances, the relationships between nodes may be searched to determine where the data item may have

been used to establish a relationship between nodes in the model. In such instances, a given data item may be applied to many relationships in a model.

[00172] The model manager 1206 may perform updates to the graph model of users 1202 on a periodic basis using data sets that may be partial or incomplete. For example, a data source 1204 may include cellular telephone usage data gathered from a subset of cellular telephone access points. In such an example, the model manager 1206 may be able to add new nodes and edges to the model.

[00173] The model manager 1206 may perform updates to the model using comprehensive or relatively complete sets of data. In such cases, the model manager 1206 may have data that may be complete enough to identify nodes or edges that may be incorrect, outdated, or no longer relevant. In such cases, the model manager 1206 may rebuild sections of the model using the new, comprehensive set of data and may prune or remove portions of the model. In some embodiments, outdated elements of the model may be denigrated but not removed from the model. Such embodiments may keep the older elements in the model to enable historical analysis of the elements.

[00174] An analysis engine 1208 may analyze the graph model of users 1202 against a network and may estimate parameters that may be found in the network. The network may be any set of interactions that may be mapped against the model. Examples of such networks may be a transportation network, a social network, a utility network, and other network.

[00175] An analysis engine 1208 may start with a network, then may determine how each node and edge of the model may interact with the network. In an example of a transportation network, the analysis engine 1208 may attempt to map how each person's transportation habits may apply to the network. For example, a user's employment start time and end time may determine when a user would commute between home and work. The user's commuting habits may be affected by the user's spouse's employment start time, the user's children's school start times, the user's distaste for crowded subways, and many other factors. The analysis engine 1208 may combine these factors with similar factors from many other users to determine a commuting pattern for a transportation network.

[00176] The analysis of a transportation network may be used to predict different measurable factors. For example, an analysis of a transportation network may be used to predict the footfall of traffic in a subway station during the day, determine the utilization of busses at off-peak hours, analyze the effect of a road closure, or other analyses.

[00177] A utility network may be analyzed using a similar mechanism. A utility network may be a water system, electrical distribution system, sewer or waste disposal system, communications network, or any other utility. An analysis engine 1208 may be used to model the usage of the network, determine peak usage, determine factors that affect usage, and other analyses.

[00178] The analysis engine 1208 may be used to calibrate results by introducing a real network 1210 and various measured parameters 1212. The real network 1210 may be a model of a network from which the graph model of users 1202 may have been derived. An analysis of the real network 1210 may be performed and compared against the measured parameters 1212. The comparisons may be used to calibrate results such that the results from hypothetical networks 1216 may have a high confidence value.

[00179] A hypothetical network 1216 may be any network for which predicted results 1218 may be desired. In many cases, a hypothetical network 1216 may reflect a change from a real network 1210. For example, an analysis of a transportation network may be performed with a hypothetical network that has a simulated closure of a single road. In such a case, the hypothetical network may have merely one factor or element that varies from the real network, and the analysis may identify perturbations or changes in the predicted results 1218 from a set of measured results 1212.

[00180] In another example, a hypothetical network 1216 may be a substantially new or different network from the real world. For example, an analysis may be performed using a network that includes a new sports team in a city. The social network of the city may be affected, as people who have an interest in the sport may be enthused by the team; the transportation network may be affected by the addition of a sports stadium; the distribution network of goods by be affected as fans purchase sports related items; the economic network may be affected by people's attendance at sports venues; the communication network may be affected by viewing and discussing sports matches; and many other effects may be identified and estimated.

[00181] Figure 13 is a flowchart illustration of an embodiment 1300 showing a method for updating a behavior model with new data. A behavior model may reflect user's behavior that may be captured in various data sets and the model may be updated by incorporating new data or by analyzing new data as the data may change over time.

[00182] Other embodiments may use different sequencing, additional or fewer steps, and different nomenclature or terminology to accomplish similar functions. In some embodiments, various operations or set of operations may be performed in parallel with other operations, either in a synchronous or asynchronous manner. The steps selected here were chosen to illustrate some principals of operations in a simplified form.

[00183] Embodiment 1300 may illustrate a simplified method for updating a behavior model. The behavior model may capture relationships between other people as well as relationships between people and locations, brands, interests, hobbies, employers, and other elements. The behavior model may capture user's interactions with various networks, which may include transportation networks, utility networks, communication networks, social networks, and other networks.

[00184] Embodiment 1300 may illustrate a method for updating a behavior model where a new data set may not be complete or comprehensive. Because a data set may not be complete, elements may be added or modified during an update but not removed from the model.

[00185] An example of an incomplete data set may be the weekend receipts from a restaurant. Within the data set of the receipts, a determination may be made that the restaurant patrons expressed an affinity for the restaurant brand and type of food and were at the restaurant's location at a specific time. These data points may be added to a behavior model for each patron of the restaurant. However, previous patrons of the restaurant who did not attend the restaurant for the time period cannot be removed from the set of all previous patrons.

[00186] An example of a comprehensive data set may be the monthly renters of an apartment complex. Each of the people in the renters data set would be added to the behavior model as living in the apartment complex, and any people who were previously in the set of renters may be removed from the set.

[00187] Recent data that include people behaviors are received in block 1302. Each person identified in the new data may be processed in block 1304.

[00188] For each person in block 1304, any personal data parameters may be determined from the new data in block 1306. The personal data parameters may be any types of nodes in a behavior graph that may be derived from the data set. In the example of restaurant receipts given above, the parameters may be a location of the person at a date and time, the affinity for a restaurant brand and food type, and other similar parameters.

[00189] The graph representing the behavior model may be searched in block 1308 to retrieve elements relating to the identified person. If the person is not in the graph in block 1310, the person may be added in block 1312.

[00190] For each element or data parameter relating to the person derived from the data set in block 1314, if the element does not exist in block 1316, the element may be added in block 1318. If the element does exist in block 1316, the confidence level of the element is added in block 1320.

[00191] Figure 14 is a flowchart illustration of an embodiment 1400 showing a method for updating a portion of a behavior model when using a comprehensive data set. As described above regarding embodiment 1300, a comprehensive data set may contain data that may allow for removal or pruning of a behavior model.

[00192] Other embodiments may use different sequencing, additional or fewer steps, and different nomenclature or terminology to accomplish similar functions. In some embodiments, various operations or set of operations may be performed in parallel with other operations, either in a synchronous or asynchronous manner. The steps selected here were chosen to illustrate some principals of operations in a simplified form.

[00193] Embodiment 1400 illustrates an update method whereby a decision may be made to remove or deprecate elements in a behavior graph. A comprehensive data set may be any set of data from which such decisions may be made.

[00194] Examples of comprehensive data sets may include a complete listing of employees for a company, a complete listing of renters in an apartment complex, or other such data sets. The term “comprehensive” may indicate that the data set may be contain

all of the members of at least one particular set, not that the data set contains all members of all sets.

[00195] The comprehensive data set may be received in block 1402. All elements related to the data set may be identified in block 1404. The elements in the data set may be any parameters that may be derived from the data set and applied to a behavior model. In an example of a data set that contains the renters of an apartment complex, the elements derived from the data set may include the renter's home location, and by extension the proximity of the renter to a transportation network, such as bus stops, highways, subway locations, for example. Additional elements to which a renter may have a relationship may be to the local restaurants and shopping areas near the apartment complex, their access to and use of networks associated with the apartment complex, such as communication networks, utility networks, social networks, and other elements.

[00196] Each element may be processed in block 1406. If the element is not in the model in block 1408, the element may be added to the model in block 1410. If the element is not in the model in block 1408, the element may be updated in block 1412. An update may be, for example, increasing the confidence indicator for the element, making changes to the element based on the new data set, or other changes.

[00197] A search may be performed in block 1414 to identify all elements in the model that are not found in the new comprehensive data set. For each identified element in block 1416, the element may be identified as removed in block 1418. In some embodiments, the element may be removed from the model, while in other embodiments, the element may be deprecated or otherwise identified as no longer being active or current.

[00198] Figure 15 is a flowchart illustration of an embodiment 1500 showing a method for analyzing a network using a behavior model. Embodiment 1500 is merely one example of a simplified analysis that may be performed using a behavior model against a network.

[00199] Other embodiments may use different sequencing, additional or fewer steps, and different nomenclature or terminology to accomplish similar functions. In some embodiments, various operations or set of operations may be performed in parallel

with other operations, either in a synchronous or asynchronous manner. The steps selected here were chosen to illustrate some principals of operations in a simplified form.

[00200] Embodiment 1500 may illustrate a method for analyzing a scenario using a behavior model. The behavior model may include relationships, affinities, and other interactions for people, and those behavioral attributes may be analyzed against a network to determine how people may behave within the network.

[00201] An example of such analyses may include people's behavior with respect to a transportation network within a city. Each person may have certain parameters that may affect how they move within the city, such as their employment or school, as well as their shopping habits, hobbies, recreational interests, entertainment interests, religious and cultural affiliations, and other behavior parameters.

[00202] Each person may also have affinities and preferences that may be affected by the status of the transportation network. For example, some people may prefer to avoid crowded subways and may be inclined to wait to use the subway when the subway is not crowded. Other people may have a strict deadline for arrival at their jobs and may change other habits to meet the deadline.

[00203] The mobility representation of the users may be received in block 1502. The network representation may be received in block 1504.

[00204] The network representation may include a definition of a network suitable for simulation, as well as the rules, algorithms, or other heuristics that may relate the various elements in the behavioral model to the network. For example, a transportation network may be represented by the physical location of various connections in the network, such as bus stops or subway stations, as well as the locations of employers, businesses, shops, restaurants, venues for entertainment, recreation, sports, or other interests, as well as social and religious organizations. The behavioral model for each person may affect the likelihood that a person may wish to visit any location in the transportation network.

[00205] A person's behavioral model may be used to determine the likelihood that a person may travel from a home location to any other location. For example, a person may commute from home to a job location according to a work schedule. The person may also visit a grocery store on a regular basis, as well as other stores,

restaurants, religious institutions, recreational areas, and other locations within the transportation network. Each person's interaction may be analyzed in block 1506, with the person's simulated usage of the network may be determined in block 1508.

[00206] In the example of a transportation network, a typical analysis may simulate the movements of people over a period of time, such as a typical workday, a week, month, or other time period.

[00207] After analyzing each person's interaction with the network in block 1506, each node of the network may be analyzed in block 1510 and the results of the individual people's interactions may be aggregated in block 1512.

[00208] In block 1514, the results may be iterated upon. The iteration may re-analyze the person's behavior based on the intermediate results from the previous iteration. Iterative analysis may be useful in situations where a user's behavior may be affected by the intermediate results. For example, a user may defer travel during busy periods to avoid crowded subways.

[00209] If an iteration is to be performed in block 1514, each person may be analyzed in block 1518 and the person's behavior may be updated based on the intermediate results set in block 1520. If there is a difference in block 1522, the person's behavior results may be updated in block 1524. Each node on the network may be analyzed in block 1526 and the results may be aggregated in block 1528.

[00210] In many cases, one person's behavior with a network may affect another person's behavior. For example, several people who are friends may join each other for a social event. When one person attends the event, the other friends are likely to also attend. Thus, a first pass at the analysis may determine that one of the friends attends an event, but a second iteration on the data set may determine that a second friend is more likely to attend the event because the friend may be present. In such a case, the second friend's activities on the transportation network may be updated to correspond with the second friend attendance at the event.

[00211] The network representation may be a simplified model of a network, and generally such models are designed to include results of interest by the analyst. For example, an analyst may wish to estimate the pedestrian traffic of a restaurant that may be located near a subway station. The analyst may construct a network representation of

a transportation network that includes the subway station of interest, along with pedestrian destinations near the subway station. The analysis may model the subway system as a whole, with more detail relating to the specific station and businesses in its vicinity.

[00212] Continuing with the example, an analysis may apply a heightened or more thorough analysis of commuters and people who may regularly travel through the subway station of interest, as well as people who may have a particular interest or affinity for the restaurant of interest and any nearby attractions. People who may not traverse the subway station or have little affinity for nearby attractions may be modeled using more simplistic models or not modeled at all in some cases.

[00213] Figure 16 is a diagram illustration of an example embodiment 1600 showing an analysis that may be performed using behavior models on a network. In the example of embodiment 1600, a transportation network may be analyzed using the known behavior models of three users, Alice 1602, Bob 1604, and Charlie 1606, as well as behavior models of new residents of an apartment complex.

[00214] In the study of example embodiment 1600, the footfall traffic at various subway stations are being analyzed, and the effects of a new apartment complex may be identified by performing the analysis with and without the apartment complex in the analysis and comparing the results. The results of the analysis may be an estimated footfall traffic 1638 and 1640 that may occur during a typical day.

[00215] In the example, Alice 1602, Bob 1604, and Charlie 1606 all are employed at the same office 1608. The office 1608 is near a subway hub 1620, and both Alice 1602 and Charlie 1606 regularly commute through the subway hub 1620 on their way to the office 1608.

[00216] As part of the analysis, behavior models 1610, 1612, and 1614 are available for Alice, Bob, and Charlie. The behavior models may represent the user's times that they typically spend at the office 1608. The behavior models may be used to determine when the various employees travel during the day to arrive and leave work.

[00217] Alice 1602 may have a home 1614 that is near a subway station 1616. Subway station 1616 is on the subway green line 1618, which connects to a subway hub 1620, which is near the office 1608.

[00218] Bob 1604 may have a home 1622 that is not near a subway station. Bob may commute by walking or by car. Since the analysis focuses on the subway ridership and Bob does not use the subway for commuting, Bob's influence to the analysis may be minimal.

[00219] Charlie 1606 may have a home 1624 that is near a subway station 1626 connected to the subway purple line 1628. The subway purple line 1628 may connect with the subway hub 1620.

[00220] In a first pass or calibration analysis, the activities of the various known users may be modeled and compared to the footfall traffic patterns 1638 and 1640 that may be measured from actual data. Once the model may be calibrated, a change to the model may be introduced and the effects measured.

[00221] A new apartment complex 1630 may be introduced near the subway station 1626. The new apartment complex 1630 may have many residents 1632, 1634, and 1636. Each of the residents may be modeled using a generic model that may represent similar people expected to live in the apartment complex.

[00222] An analysis may be performed by modeling the behaviors of the new residents 1632, 1634, and 1636 to determine how the footfall traffic 1638 and 1640 changes.

[00223] The foregoing description of the subject matter has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the subject matter to the precise form disclosed, and other modifications and variations may be possible in light of the above teachings. The embodiment was chosen and described in order to best explain the principals of the invention and its practical application to thereby enable others skilled in the art to best utilize the invention in various embodiments and various modifications as are suited to the particular use contemplated. It is intended that the appended claims be construed to include other alternative embodiments except insofar as limited by the prior art.

CLAIMS

What is claimed is:

1. A method performed on at least one computer processor, said method comprising:

receiving a mobility pattern model comprising:

a transportation network representation having a plurality of location nodes and transportation links between said location nodes;

a plurality of human behavior representations, each of said human behavior representations derived from a human interaction with transportation mechanisms;

modeling human transportation using said mobility pattern model using a representation of a first transportation network to determine a first traffic pattern on a first location node;

modifying said transportation network representation to represent a second transportation network and modeling said human transportation using said mobility pattern model to determine a second traffic pattern on a second location node.

2. The method of claim 1, said second transportation network having at least one new location node added to said first transportation network.

3. The method of claim 1, said first transportation network representing an actual transportation network and said second transportation network representing a hypothetical transportation network.

4. The method of claim 3 further comprising:

measuring an actual traffic pattern at a physical location represented by said first location node; and

determining that said actual traffic pattern differs from said first traffic pattern and updating at least one of said plurality of human behavior representations.

5. The method of claim 1 :

said first transportation network representing an actual transportation network at a first time with a first set of conditions;

said second transportation network representing said actual transportation network at a second time with a second set of conditions;

measuring a first actual traffic pattern observed at said first time and with said first set of conditions;

measuring a second actual traffic pattern observed at said second time and with said second set of conditions; and

modifying at least one human behavior representations based on said first actual traffic pattern and said second actual traffic pattern.

6. The method of claim 5, said first actual traffic pattern comprising transportation patterns measured using a plurality of devices carried by individual humans.

7. The method of claim 6, at least one of said plurality of devices being mobile telephones.

8. The method of claim 1, said human behavior representations being derived from cellular telephone location data.

9. The method of claim 8, said cellular telephone location data being anonymized data.

10. A system comprising:

at least one processor;

said processor configured to perform a method comprising:

receiving a mobility pattern model comprising:

a transportation network representation having a plurality of location nodes and transportation links between said location nodes;

a plurality of human behavior representations, each of said human behavior representations derived from a human interaction with transportation mechanisms;

modeling human transportation using said mobility pattern model using a representation of a first transportation network to determine a first traffic pattern on a first location node;

modifying said transportation network representation to represent a second transportation network and modeling said human transportation using said mobility pattern model to determine a second traffic pattern on a said second location node.

11. The system of claim 10, said second transportation network having at least one new location node added to said first transportation network.

12. The system of claim 10, said first transportation network representing an actual transportation network and said second transportation network representing a hypothetical transportation network.

13. The system of claim 12, said method further comprising:

measuring an actual traffic pattern at a physical location represented by said first location node; and

determining that said actual traffic pattern differs from said first traffic pattern and updating at least one of said plurality of human behavior representations.

14. The system of claim 10:

said first transportation network representing an actual transportation network at a first time with a first set of conditions;

said second transportation network representing said actual transportation network at a second time with a second set of conditions;

said method further comprising:

measuring a first actual traffic pattern observed at said first time and with said first set of conditions;

measuring a second actual traffic pattern observed at said second time and with said second set of conditions; and

modifying at least one human behavior representations based on said first actual traffic pattern and said second actual traffic pattern.

15. The system of claim 14, said first actual traffic pattern comprising transportation patterns measured using a plurality of devices carried by individual humans.

16. The system of claim 15, at least one of said plurality of devices being mobile telephones.

17. The system of claim 10, said human behavior representations being derived from cellular telephone location data.

18. The system of claim 17, said cellular telephone location data being anonymized data.

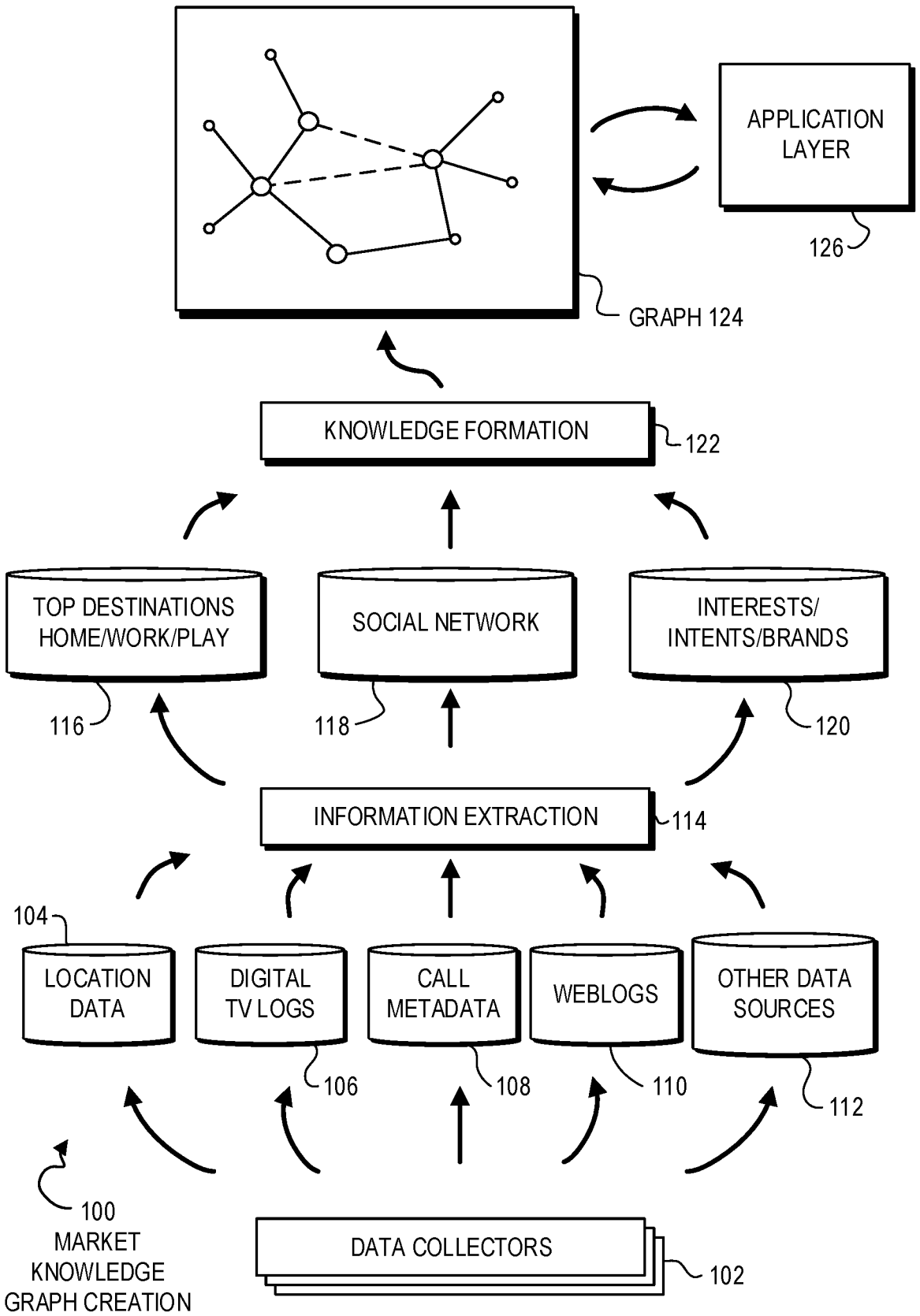

APPLICATION LAYER
126
GRAPH 124
KNOWLEDGE FORMATION
122
TOP DESTINATIONS HOME/WORK/PLAY
SOCIAL NETWORK
INTERESTS/ INTENTS/BRANDS
116
118
120
INFORMATION EXTRACTION
114
104
LOCATION DATA
DIGITAL TV LOGS
CALL METADATA
WEBLOGS
OTHER DATA SOURCES
106
108
110
112
100
MARKET KNOWLEDGE GRAPH CREATION
DATA COLLECTORS
102

FIG. 1

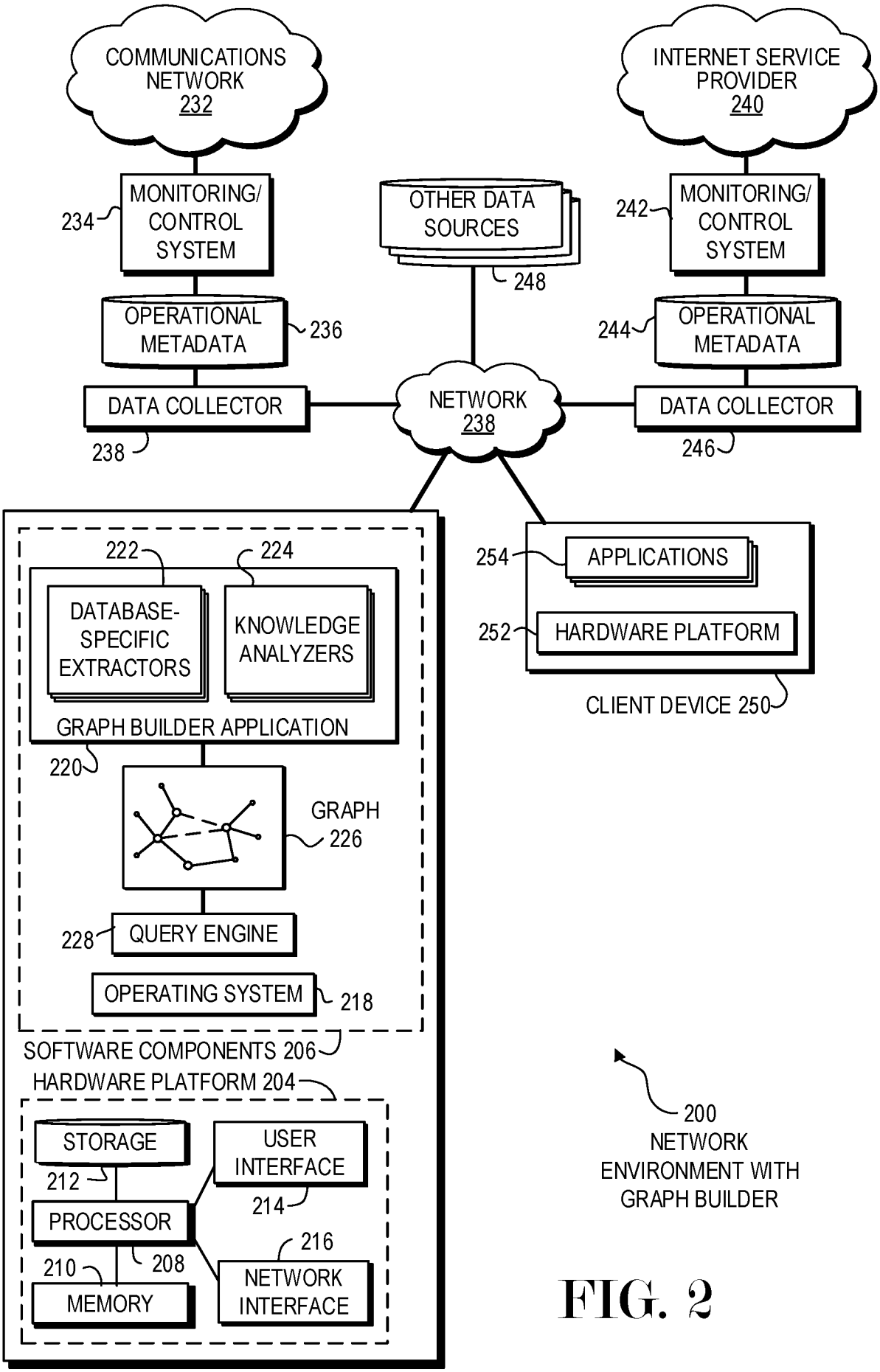

COMMUNICATIONS NETWORK 232
INTERNET SERVICE PROVIDER 240
234
MONITORING/ CONTROL SYSTEM
OTHER DATA SOURCES
248
242
MONITORING/ CONTROL SYSTEM
OPERATIONAL METADATA
236
244
OPERATIONAL METADATA
DATA COLLECTOR
238
NETWORK 238
DATA COLLECTOR
246
222
224
254
APPLICATIONS
DATABASE-SPECIFIC EXTRACTORS
KNOWLEDGE ANALYZERS
252
HARDWARE PLATFORM
GRAPH BUILDER APPLICATION
CLIENT DEVICE 250
220
GRAPH
226
228
QUERY ENGINE
OPERATING SYSTEM
218
SOFTWARE COMPONENTS 206
HARDWARE PLATFORM 204
STORAGE
USER INTERFACE
212
214
PROCESSOR
216
210
208
NETWORK INTERFACE
MEMORY
200
NETWORK ENVIRONMENT WITH GRAPH BUILDER

FIG. 2

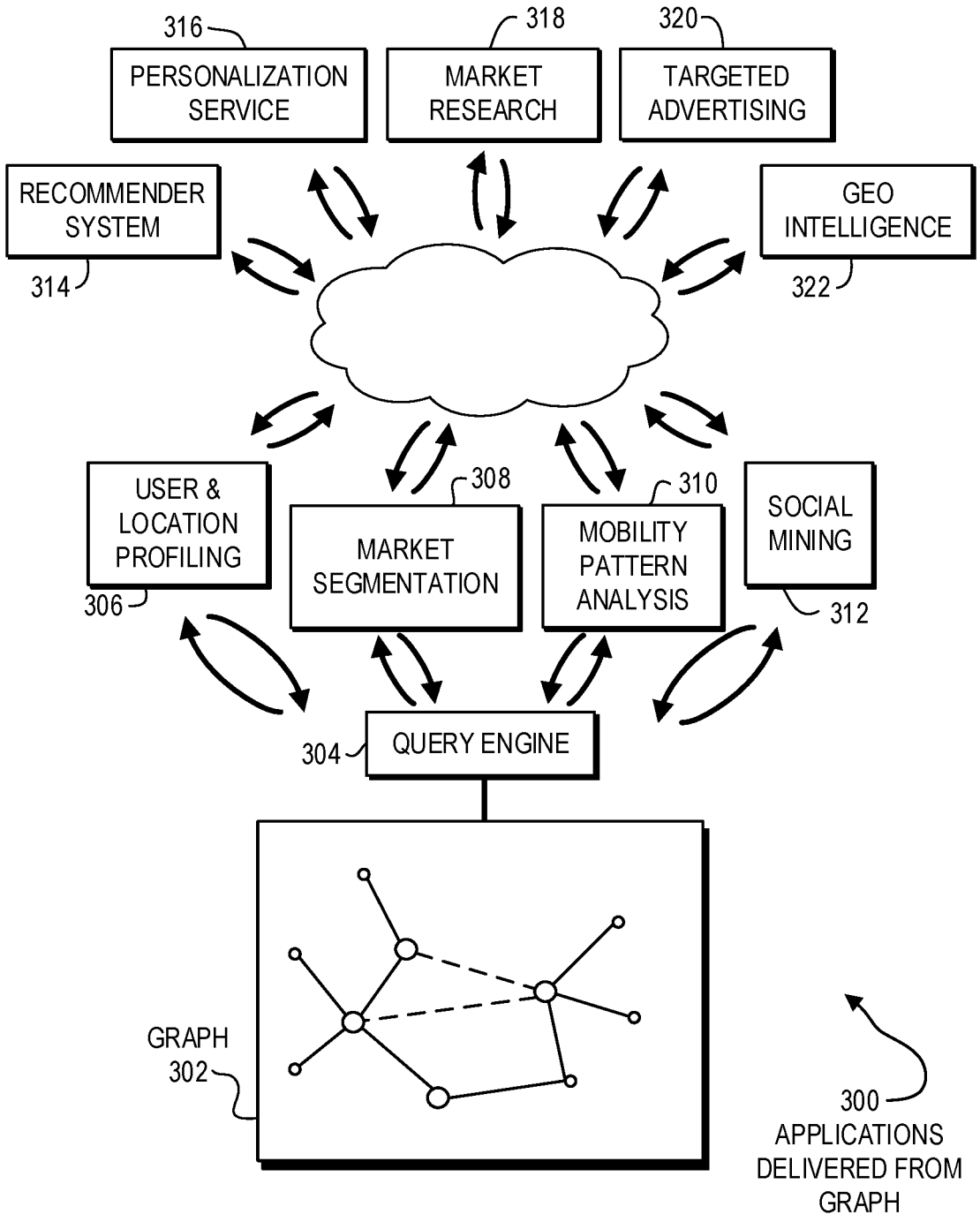

316
PERSONALIZATION SERVICE
318
MARKET RESEARCH
320
TARGETED ADVERTISING
RECOMMENDER SYSTEM
314
GEO INTELLIGENCE
322
USER & LOCATION PROFILING
306
308
MARKET SEGMENTATION
310
MOBILITY PATTERN ANALYSIS
SOCIAL MINING
312
304
QUERY ENGINE
GRAPH
302
300
APPLICATIONS DELIVERED FROM GRAPH

FIG. 3

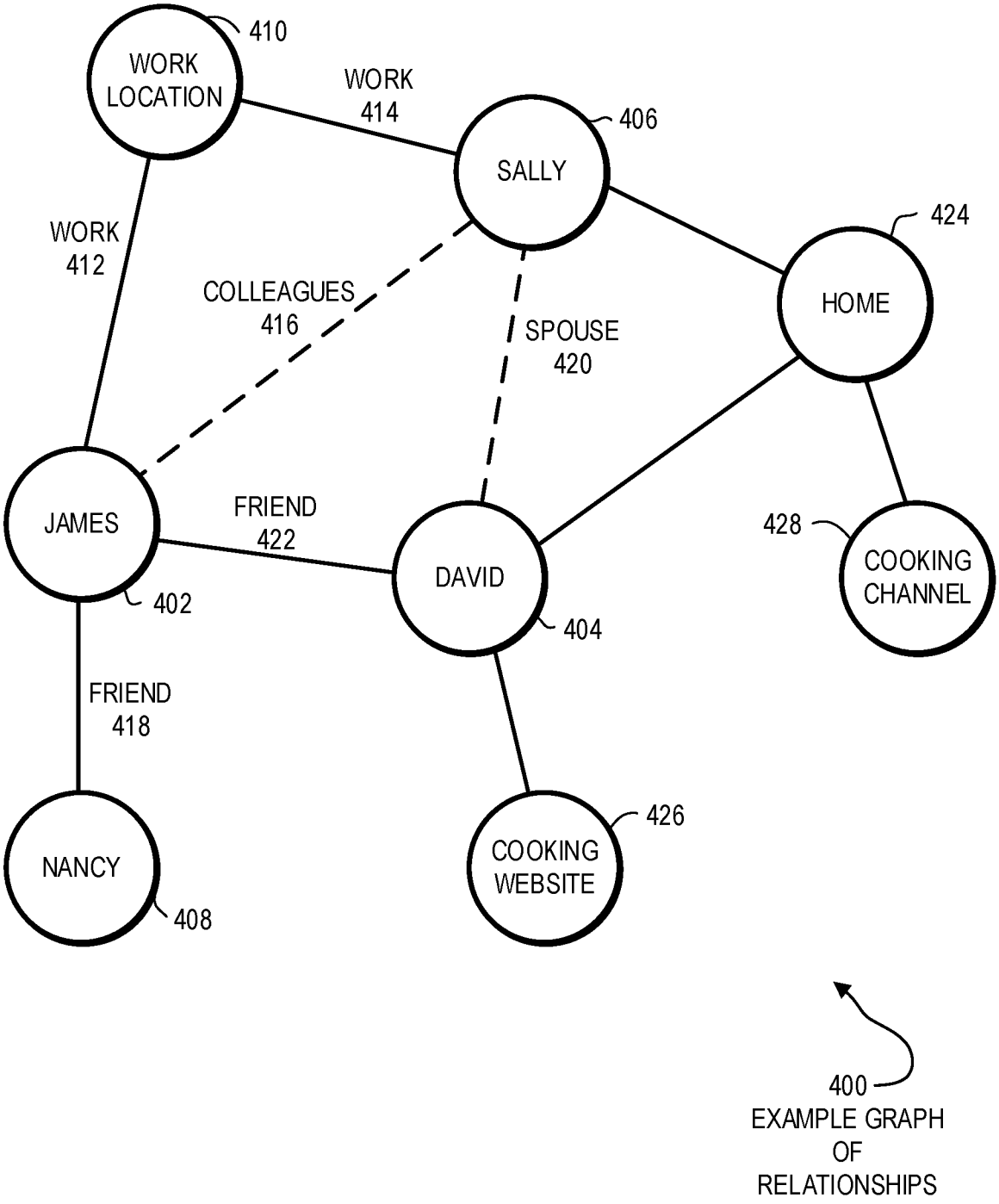

WORK LOCATION
410
WORK
414
SALLY
406
424
WORK
412
COLLEAGUES
416
SPOUSE
420
HOME
JAMES
FRIEND
422
DAVID
428
COOKING CHANNEL
402
404
FRIEND
418
NANCY
408
COOKING WEBSITE
426
400
EXAMPLE GRAPH OF RELATIONSHIPS

FIG. 4

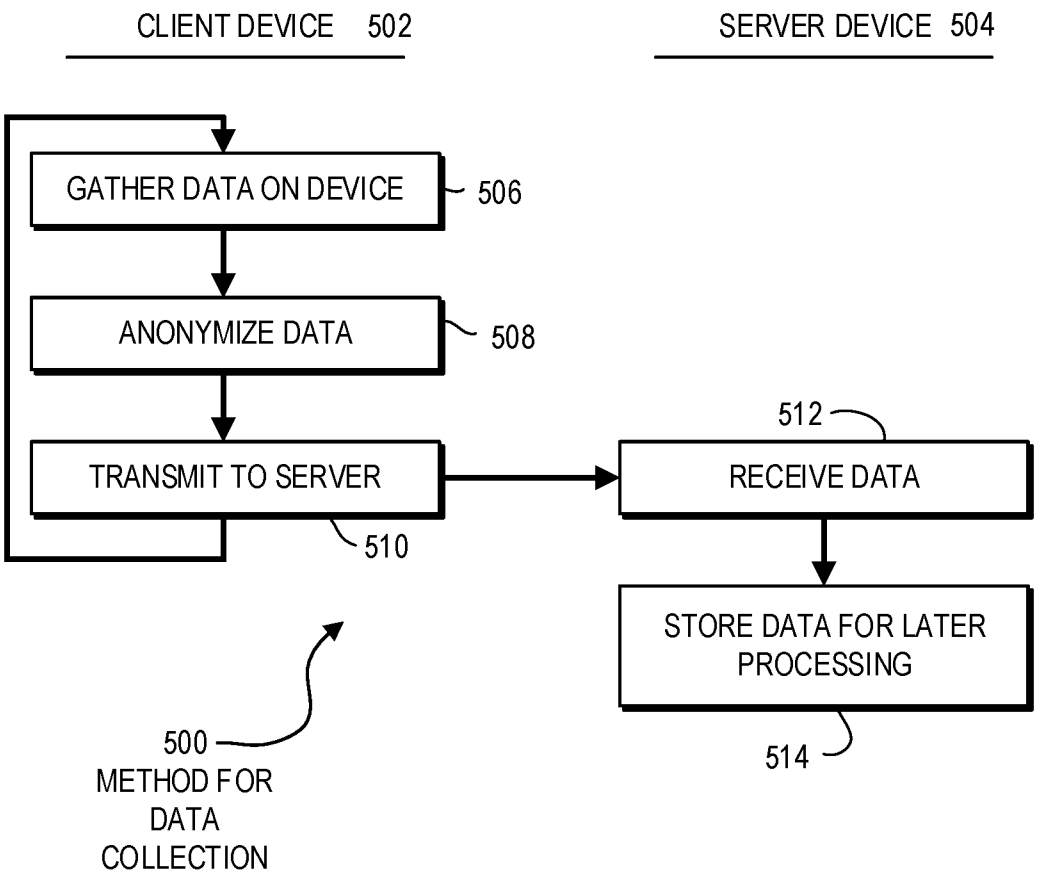

CLIENT DEVICE 502
SERVER DEVICE 504
GATHER DATA ON DEVICE
506
ANONYMIZE DATA
508
TRANSMIT TO SERVER
510
512
RECEIVE DATA
STORE DATA FOR LATER PROCESSING
514
500
METHOD FOR DATA COLLECTION

FIG. 5

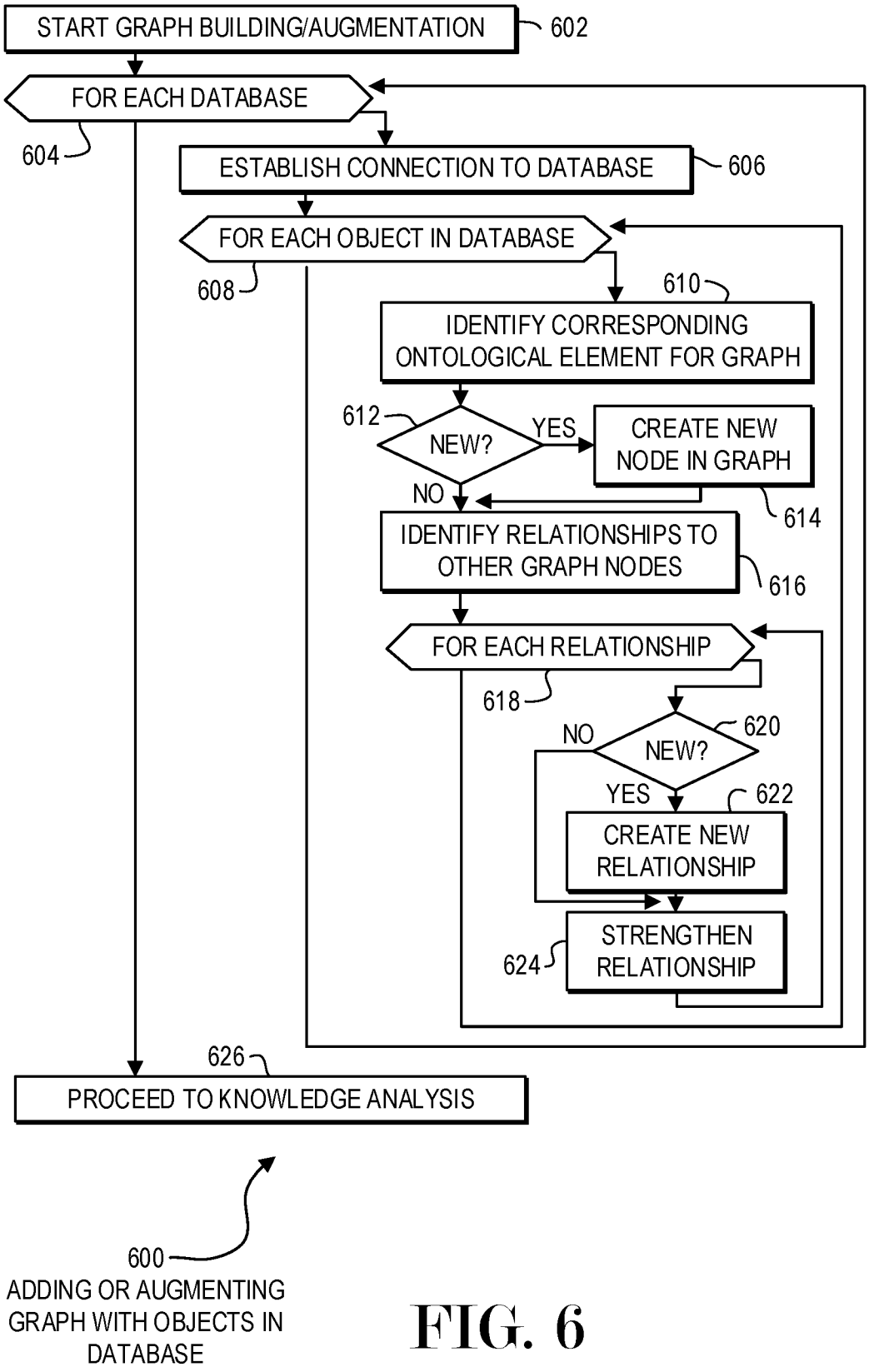

START GRAPH BUILDING/AUGMENTATION
602
FOR EACH DATABASE
604
ESTABLISH CONNECTION TO DATABASE
606
FOR EACH OBJECT IN DATABASE
608
610
IDENTIFY CORRESPONDING ONTOLOGICAL ELEMENT FOR GRAPH
612
NEW?
YES
CREATE NEW NODE IN GRAPH
614
NO
IDENTIFY RELATIONSHIPS TO OTHER GRAPH NODES
616
FOR EACH RELATIONSHIP
618
NO
NEW?
620
YES
622
CREATE NEW RELATIONSHIP
STRENGTHEN RELATIONSHIP
624
626
PROCEED TO KNOWLEDGE ANALYSIS
600
ADDING OR AUGMENTING GRAPH WITH OBJECTS IN DATABASE

FIG. 6

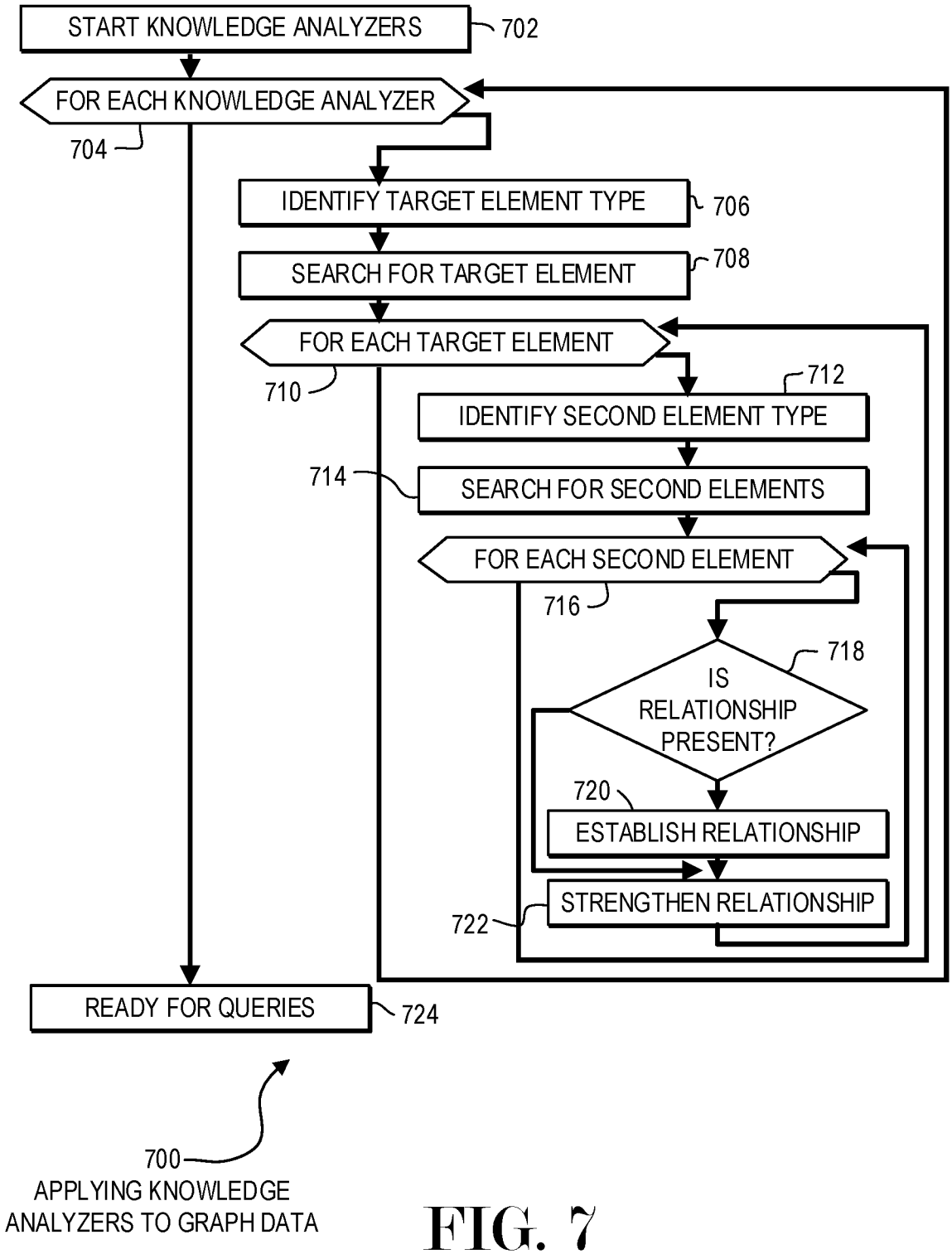

START KNOWLEDGE ANALYZERS
702
FOR EACH KNOWLEDGE ANALYZER
704
IDENTIFY TARGET ELEMENT TYPE
706
SEARCH FOR TARGET ELEMENT
708
FOR EACH TARGET ELEMENT
710
IDENTIFY SECOND ELEMENT TYPE
712
SEARCH FOR SECOND ELEMENTS
714
FOR EACH SECOND ELEMENT
716
IS RELATIONSHIP PRESENT?
718
ESTABLISH RELATIONSHIP
720
STRENGTHEN RELATIONSHIP
722
READY FOR QUERIES
724
700
APPLYING KNOWLEDGE ANALYZERS TO GRAPH DATA

FIG. 7

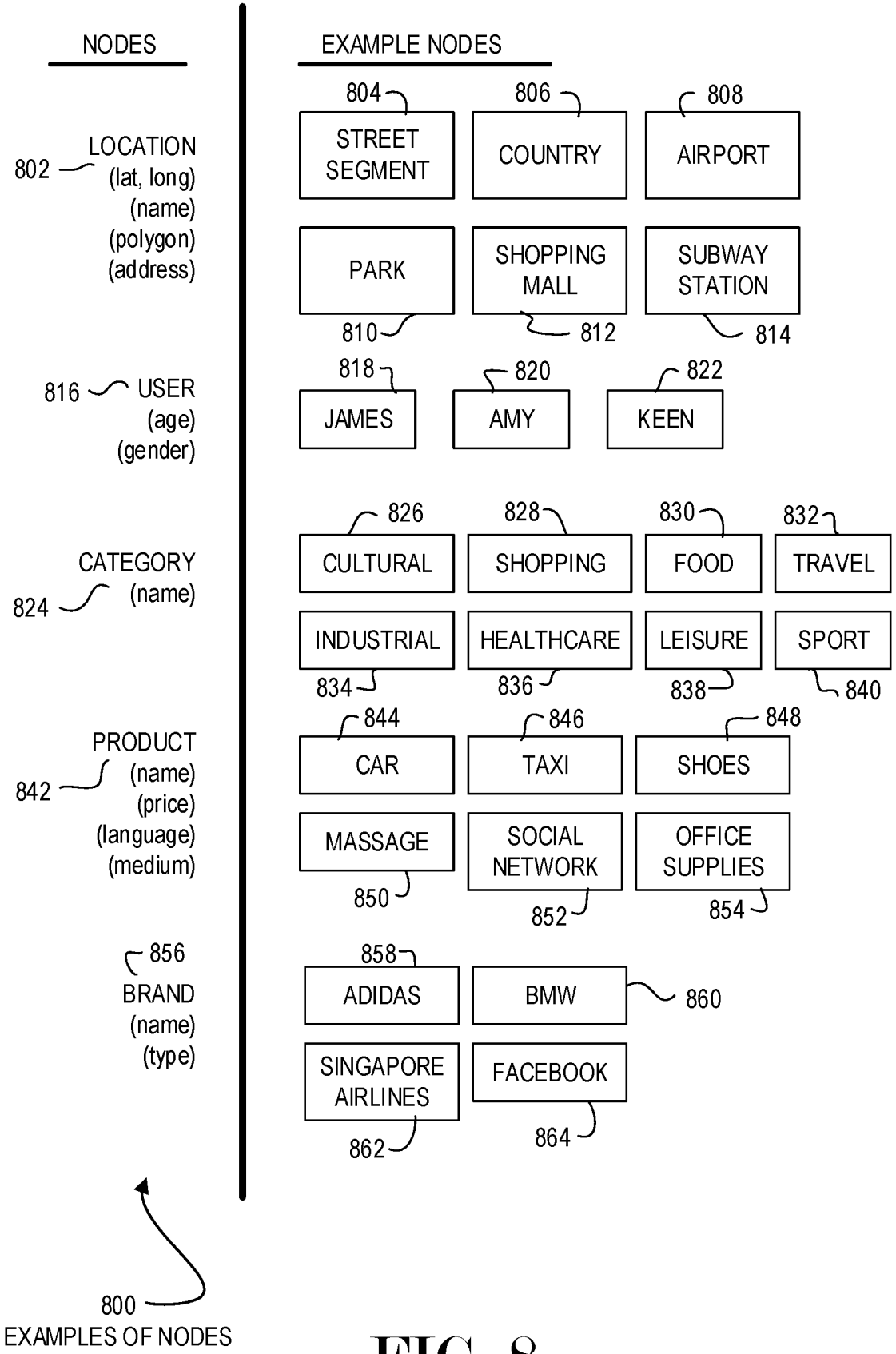

NODES
EXAMPLE NODES
LOCATION
(lat, long)
(name)
(polygon)
(address)
802
USER
(age)
(gender)
816
CATEGORY
(name)
824
PRODUCT
(name)
(price)
(language)
(medium)
842
BRAND
(name)
(type)
856
804
806
808
STREET SEGMENT
COUNTRY
AIRPORT
PARK
SHOPPING MALL
SUBWAY STATION
810
812
814
818
820
822
JAMES
AMY
KEEN
826
828
830
832
CULTURAL
SHOPPING
FOOD
TRAVEL
INDUSTRIAL
HEALTHCARE
LEISURE
SPORT
834
836
838
840
844
846
848
CAR
TAXI
SHOES
MASSAGE
SOCIAL NETWORK
OFFICE SUPPLIES
850
852
854
858
ADIDAS
BMW
860
SINGAPORE AIRLINES
FACEBOOK
862
864
800
EXAMPLES OF NODES

FIG. 8

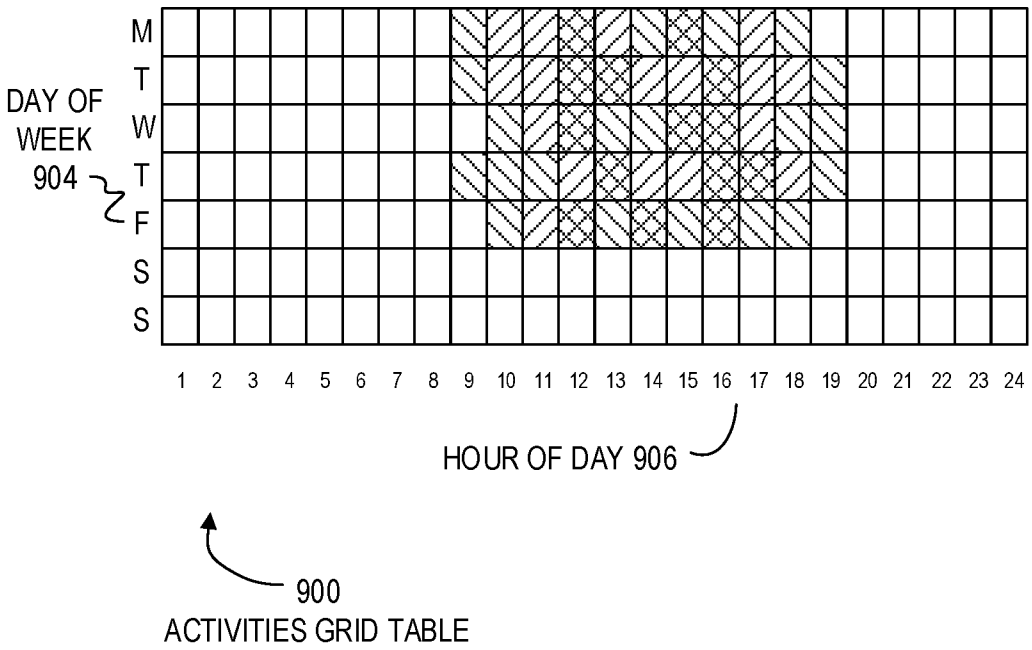

DAY OF WEEK 904
M
T
W
T
F
S
S
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
HOUR OF DAY 906
900
ACTIVITIES GRID TABLE

FIG. 9

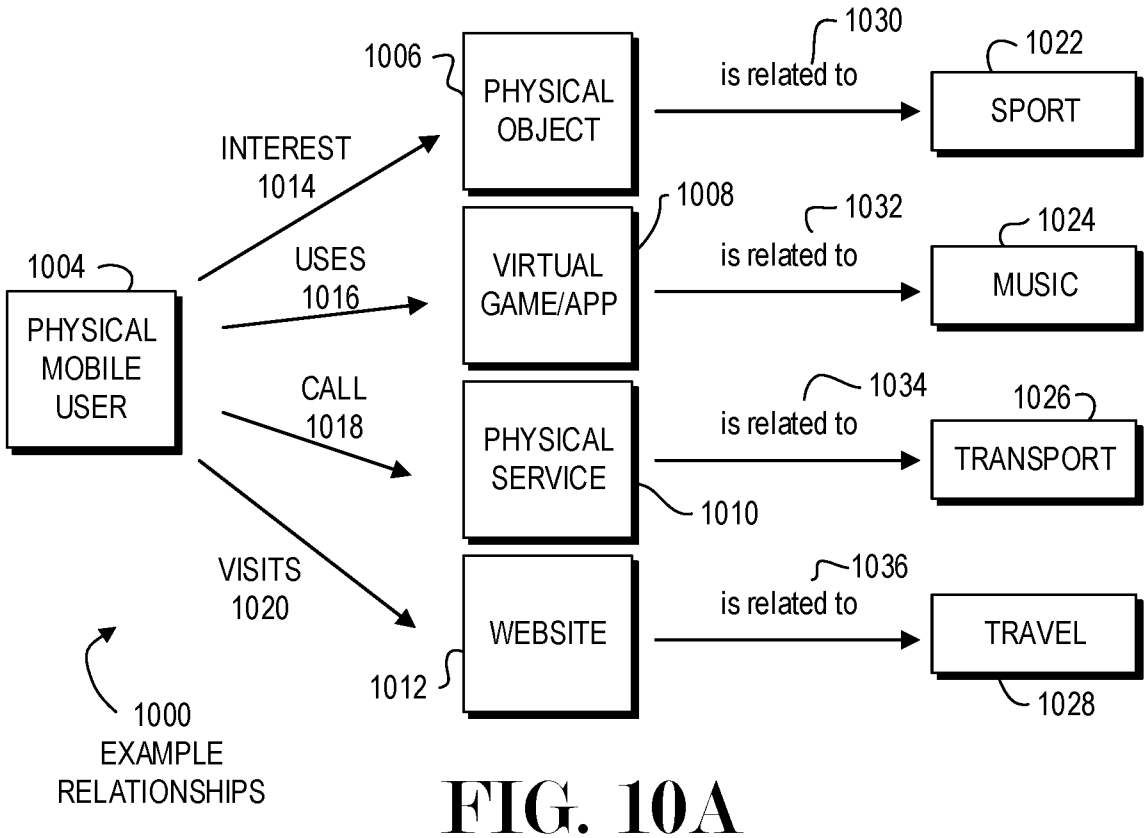


FIG. 10A

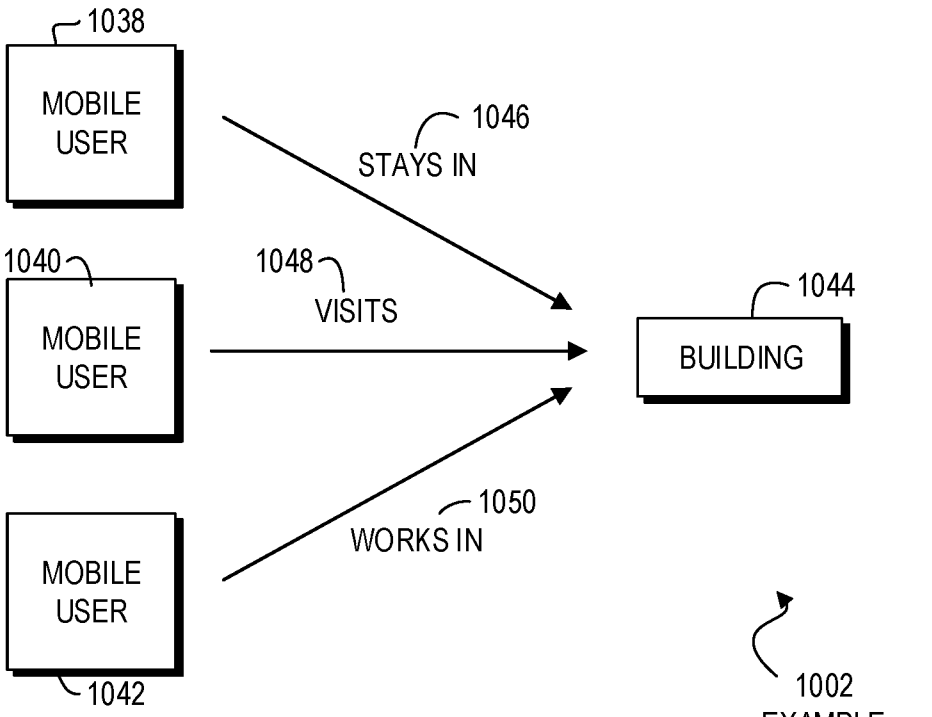


FIG. 10B

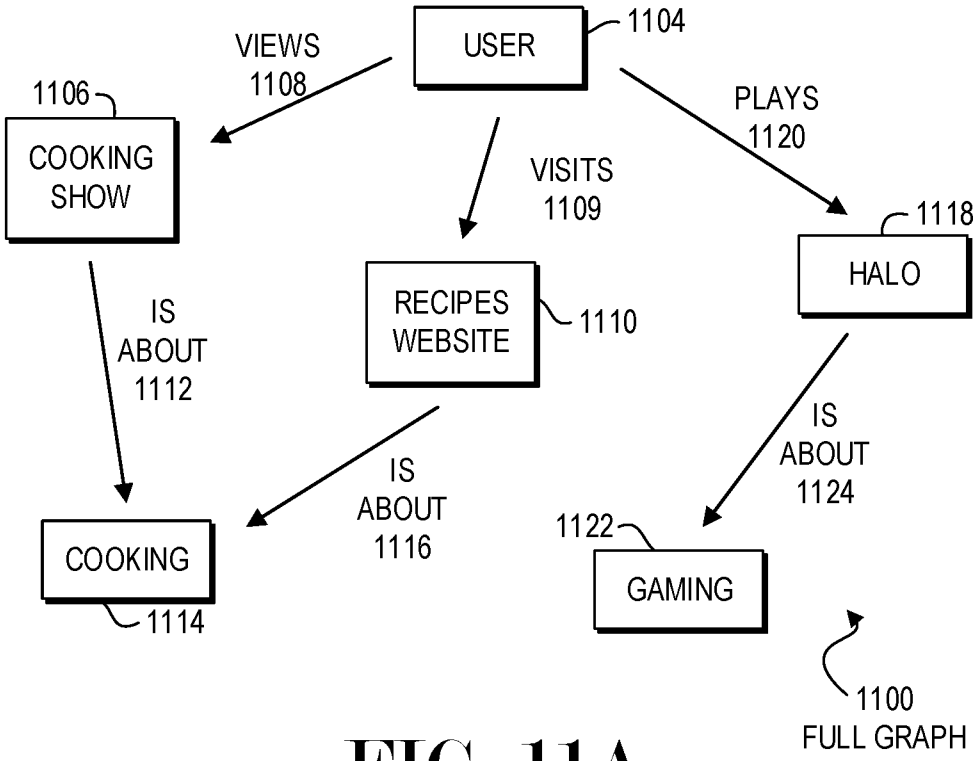


FIG. 11A

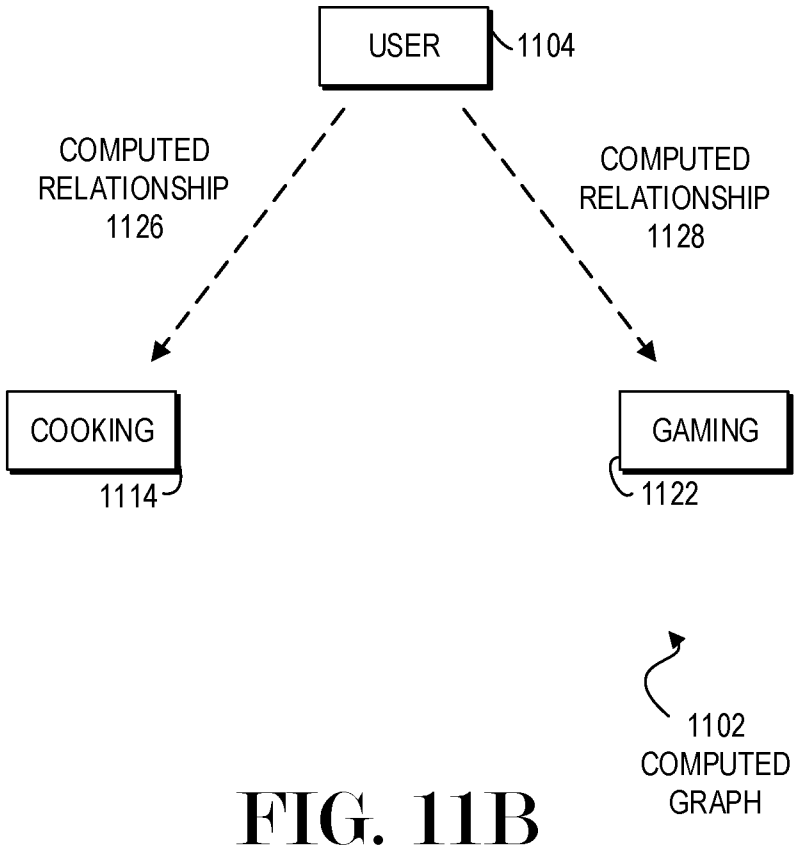


FIG. 11B

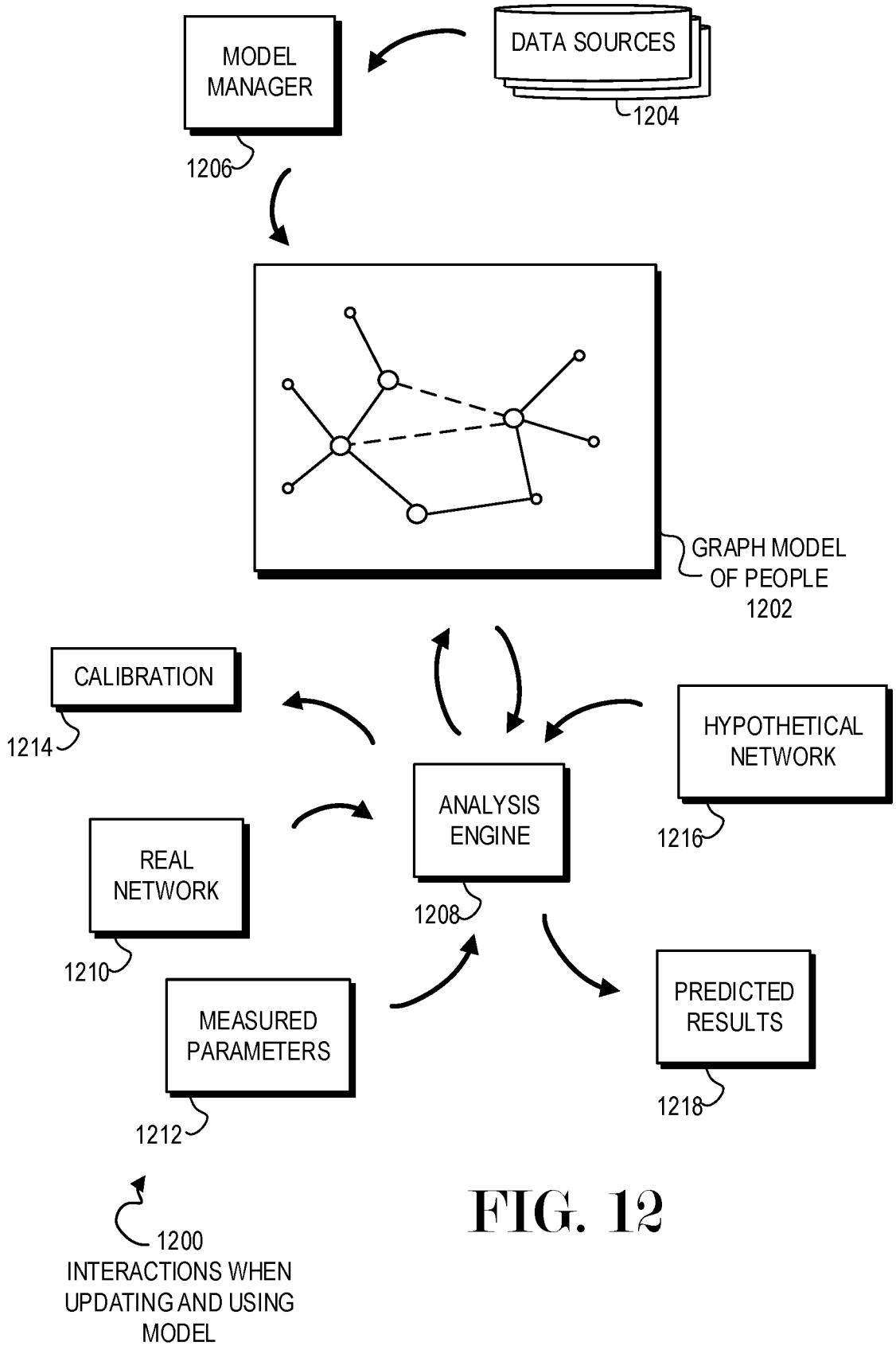

MODEL MANAGER
1206
DATA SOURCES
1204
GRAPH MODEL OF PEOPLE 1202
CALIBRATION
1214
HYPOTHETICAL NETWORK
1216
ANALYSIS ENGINE
1208
REAL NETWORK
1210
MEASURED PARAMETERS
1212
PREDICTED RESULTS
1218
1200
INTERACTIONS WHEN UPDATING AND USING MODEL

FIG. 12

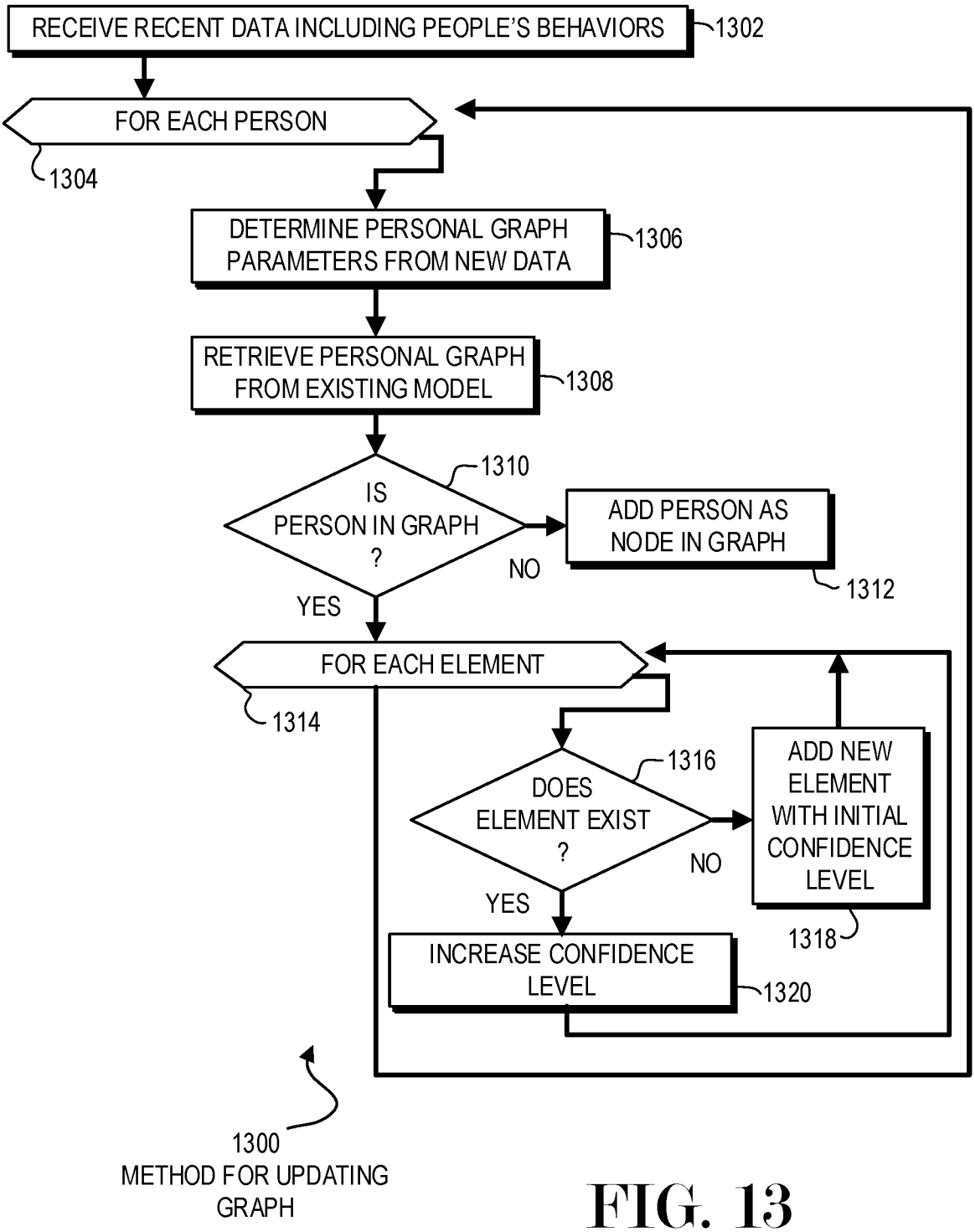
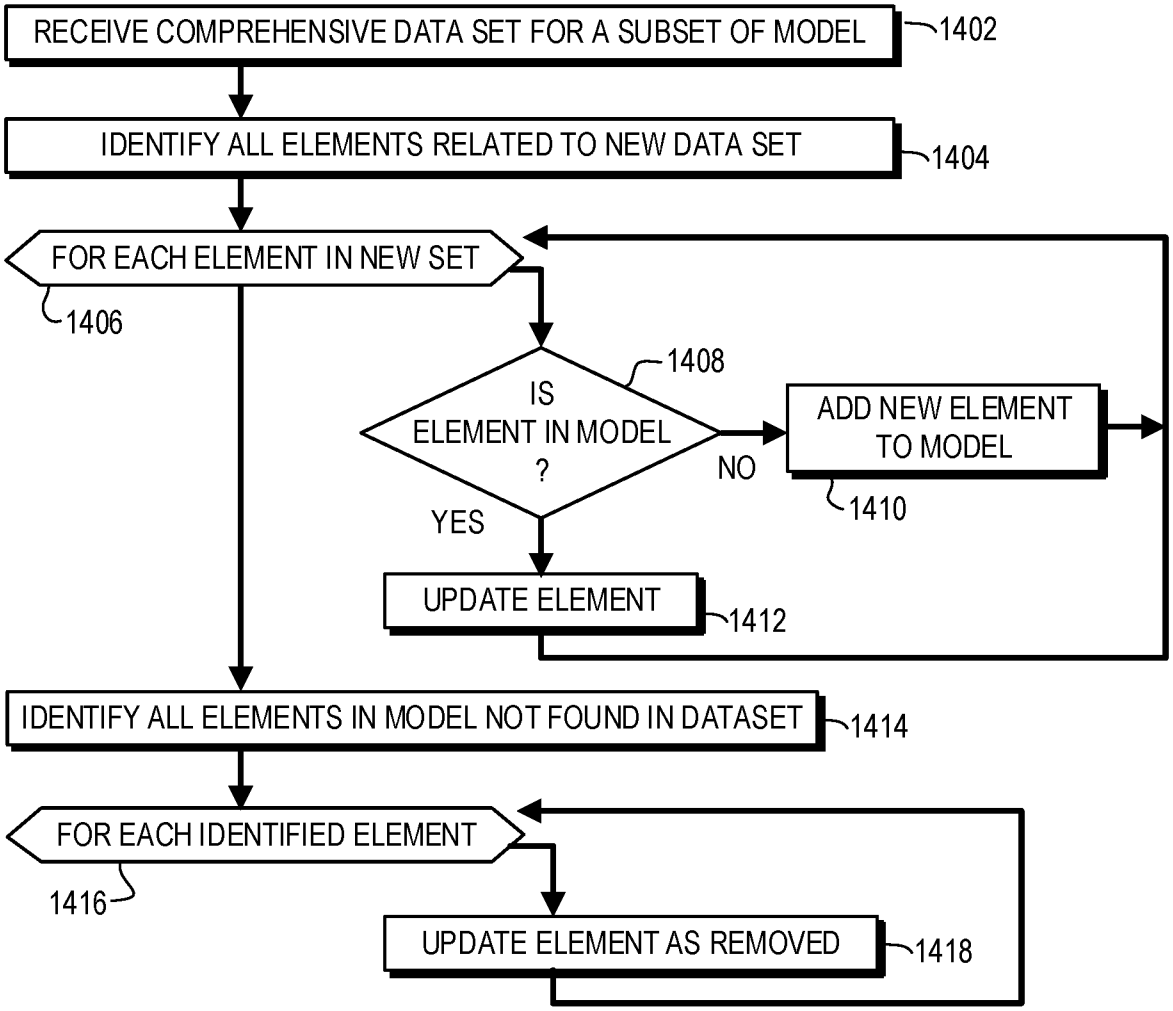

RECEIVE RECENT DATA INCLUDING PEOPLE'S BEHAVIORS
1302
FOR EACH PERSON
1304
DETERMINE PERSONAL GRAPH PARAMETERS FROM NEW DATA
1306
RETRIEVE PERSONAL GRAPH FROM EXISTING MODEL
1308
1310
IS PERSON IN GRAPH ?
NO
ADD PERSON AS NODE IN GRAPH
1312
YES
FOR EACH ELEMENT
1314
1316
DOES ELEMENT EXIST ?
NO
ADD NEW ELEMENT WITH INITIAL CONFIDENCE LEVEL
1318
YES
INCREASE CONFIDENCE LEVEL
1320
1300
METHOD FOR UPDATING GRAPH

FIG. 13



1400
METHOD FOR UPDATING USING COMPREHENSIVE DATA SET

FIG. 14

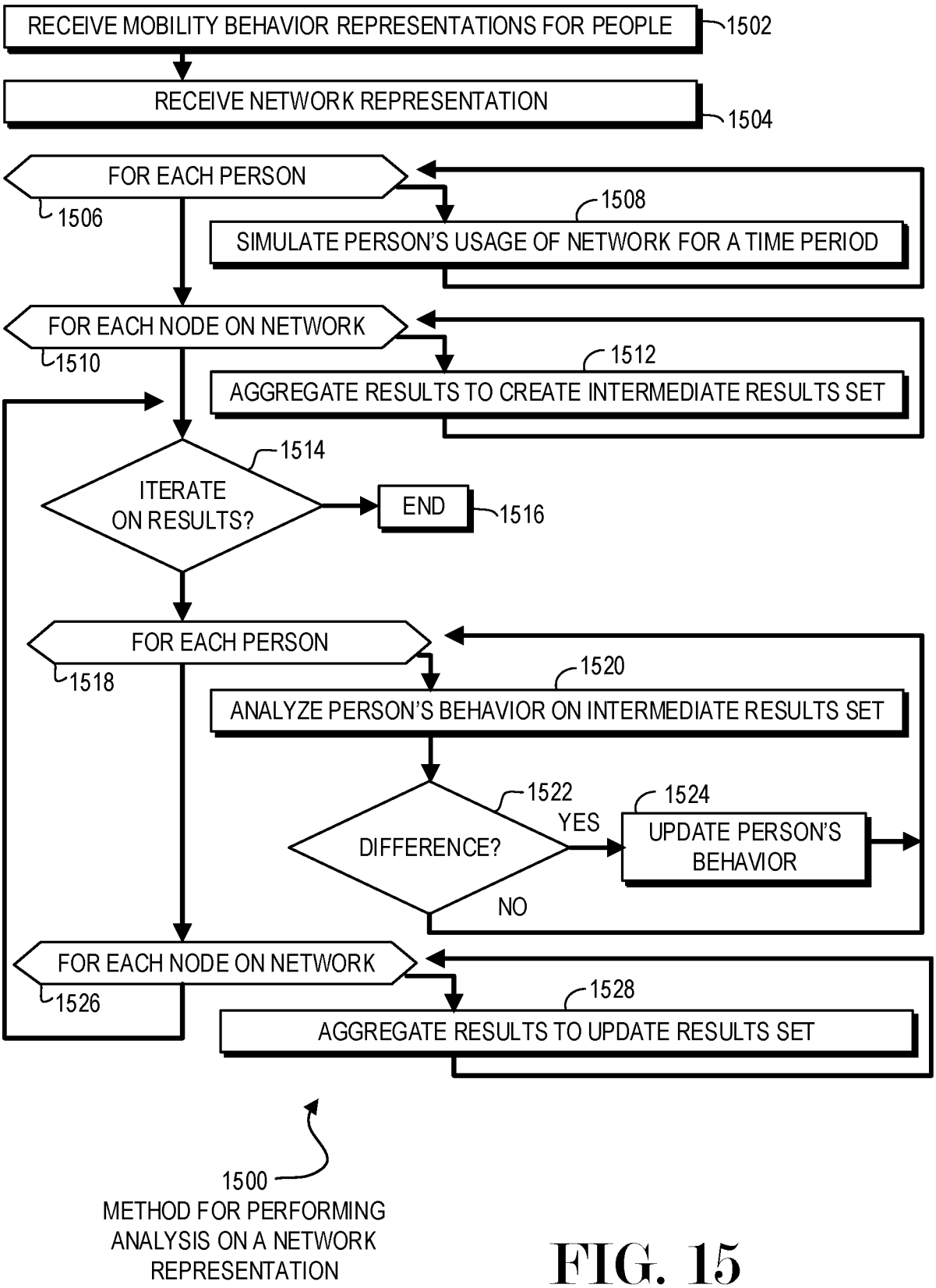

RECEIVE MOBILITY BEHAVIOR REPRESENTATIONS FOR PEOPLE
1502
RECEIVE NETWORK REPRESENTATION
1504
FOR EACH PERSON
1506
1508
SIMULATE PERSON'S USAGE OF NETWORK FOR A TIME PERIOD
FOR EACH NODE ON NETWORK
1510
1512
AGGREGATE RESULTS TO CREATE INTERMEDIATE RESULTS SET
1514
ITERATE ON RESULTS?
END
1516
FOR EACH PERSON
1518
1520
ANALYZE PERSON'S BEHAVIOR ON INTERMEDIATE RESULTS SET
1522
DIFFERENCE?
YES
1524
UPDATE PERSON'S BEHAVIOR
NO
FOR EACH NODE ON NETWORK
1526
1528
AGGREGATE RESULTS TO UPDATE RESULTS SET
1500
METHOD FOR PERFORMING ANALYSIS ON A NETWORK REPRESENTATION

FIG. 15

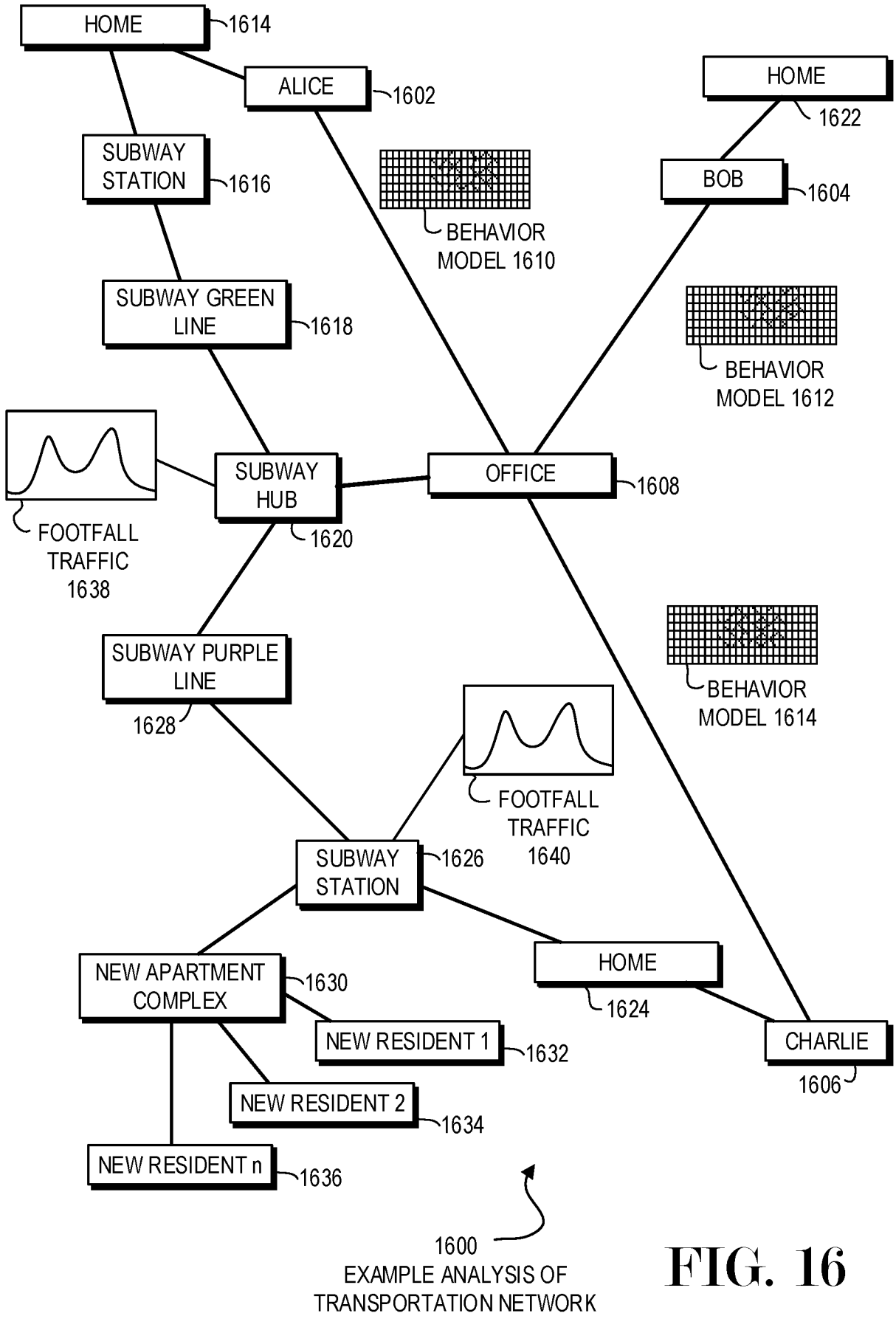

HOME
1614
ALICE
1602
HOME
1622
SUBWAY STATION
1616
BEHAVIOR MODEL 1610
BOB
1604
SUBWAY GREEN LINE
1618
BEHAVIOR MODEL 1612
SUBWAY HUB
1620
OFFICE
1608
FOOTFALL TRAFFIC 1638
SUBWAY PURPLE LINE
1628
BEHAVIOR MODEL 1614
FOOTFALL TRAFFIC 1640
SUBWAY STATION
1626
HOME
1624
NEW APARTMENT COMPLEX
1630
NEW RESIDENT 1
1632
CHARLIE
1606
NEW RESIDENT 2
1634
NEW RESIDENT n
1636
1600
EXAMPLE ANALYSIS OF TRANSPORTATION NETWORK

FIG. 16