(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 January 2025 (23.01.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/019251 A2

(51) International Patent Classification:
*G06Q 10/10* (2023.01) *G06N 20/00* (2019.01)

(21) International Application Number:
PCT/US2024/037519

(22) International Filing Date:
11 July 2024 (11.07.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/526,939 14 July 2023 (14.07.2023) US

(71) Applicant: **WALMART APOLLO, LLC** [US/US]; 702 SW 8th Street, Bentonville, Arkansas 72716 (US).

(72) Inventors: **DAS, Amlan Jyoti**; Purbachal Sports Complex, Kora Purbachal, Kora Chandigarh, Madhyamgram, North 24 Parganas, WB 700130 (IN). **ABRAHAM, Abin**; G6, Mana Candela, Kodathi Village, Sarjapur Road, Bangalore, KA 560035 (IN). **BALASUBRAMANIAN, Ravishankar**; A1102, Adarsh Rhythm Apartments, 71 Panduranga Nagar, Bengaluru, KA 560076 (IN). **SAHU, Ishant**; House 48/23, Street 1, Maitri Kunj Private, Risali, Bhilai Nagar, CG 490006 (IN). **SAHU, Subrat Sekhar**; Flat #404, Garden Residency, Garden Layout, HSR Sector 2, Bengaluru, KA 56102 (IN).

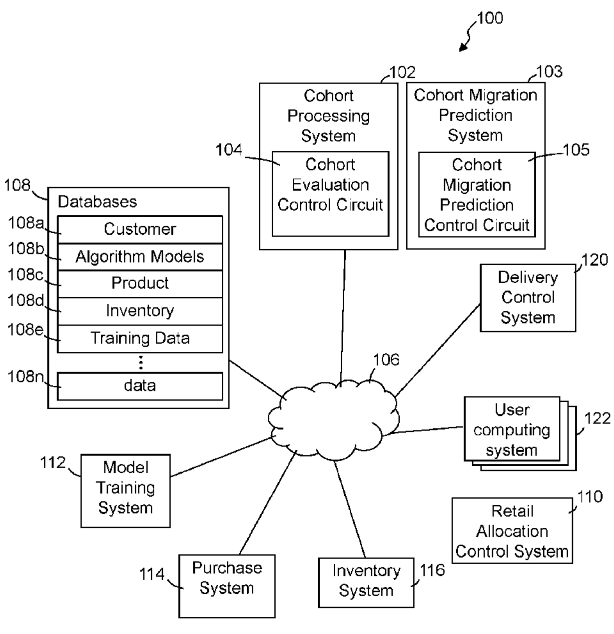
(74) Agent: **KRATZ, Rudy** et al.; Fitch, Even, Tabin & Flannery LLP, 120 South LaSalle Street, Suite 2100, Chicago, Illinois 60603 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,

(54) **Title:** SYSTEMS AND METHODS OF RETAIL CONTROL BASED ON CUSTOMER COHORT MIGRATIONS

*FIG. 1*




**(57) Abstract:** Some embodiments provide systems to predict reallocation of customer categorization, comprising: a customer database; and a cohort migration prediction system comprising a migration prediction control circuit configured to: identify migration of a first set of customers between multiple cohorts; apply a causation model to at least a first subset of the customer information and identify, based on the application of the causation model, at least a causation factor predicted to at least partially cause a change in behavior resulting in the migration of the first set of customers previously associated with the first cohort, of the first cohort type, to a second cohort of the first cohort type; and apply a prediction model to a subset of the customer information relative to the first cohort, and predict future migration between the first cohort and at least the second cohorts of the first cohort type.

WO 2025/019251 A2

*[Continued on next page]*

SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

# SYSTEMS AND METHODS OF RETAIL CONTROL BASED ON CUSTOMER COHORT MIGRATIONS

## CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 63/526,939 filed July 14, 2023, which is incorporated herein by reference in its entirety.

## TECHNICAL FIELD

[0002] This invention relates generally to machine learning in controlling retail inventory and distribution.

## BACKGROUND

[0003] Customer satisfaction is critical to retail entities. It is often difficult to meet customer expectations do to supply issues, stocking issues, and other such issues. Further, customers and their expectations and wants often change over time.

## BRIEF DESCRIPTION OF DRAWINGS

[0004] Disclosed herein are embodiments of systems, apparatuses and methods pertaining to the application of machine learning and predefined sets of rules for customer information evaluation to improve the control and operation of commercial entities. This description includes drawings, wherein:

[0005] FIG. 1 illustrates a simplified block diagram of an exemplary retail and/or retail facility control system, in accordance with some embodiments.

[0006] FIG. 2 illustrates a simplified flow diagram of an exemplary process of controlling retail allocation, in accordance with some embodiments.

[0007] FIG. 3 illustrates a simplified flow diagram of an exemplary process of evaluating customer information and identifying cohorts of customers based on the user defined hierarchical parameters, in accordance with some embodiments.

[0008] FIGS. 4A-4B illustrate a simplified flow diagram of an exemplary threshold algorithm process of implementing an exemplary configurable threshold algorithm model, in accordance with some embodiments.

[0009] FIG. 5 illustrates a simplified block diagram of an exemplary graphical representation of a multi-dimensional cohort distribution of multiple cohorts of a cohort type, in accordance with some embodiments.

[0010] FIGS. 6A-6B illustrate a simplified flow diagram of an exemplary process of implementing an exemplary hierarchical clustering or agglomerative scoring algorithm model, in accordance with some embodiments.

[0011] FIGS. 7A-7B illustrate a simplified flow diagram of an exemplary process of implementing an exemplary parametric hyperboloid algorithm model, in accordance with some embodiments.

[0012] FIG. 8 illustrates a simplified, two-dimensional representation of an exemplary hyperplane, with two hyperboloids dividing the two-dimensional hyperplane into three representative classes or cohorts, in accordance with some embodiments.

[0013] FIGS. 9A-9B illustrate a simplified flow diagram of an exemplary process of implementing an exemplary sequential repetition scoring algorithm model, in accordance with some embodiments.

[0014] FIG. 10 illustrates a simplified graphic representation of an exemplary multi-dimensional table, in accordance with some embodiments.

[0015] FIG. 11 illustrates a simplified flow diagram of an exemplary cohort migration identification process, in accordance with some embodiments.

[0016] FIG. 12 illustrates a simplified flow diagram of an exemplary cause detection process, in accordance with some embodiments.

[0017] FIG. 13 illustrates a simplified flow diagram of an exemplary migration prediction process in predicting customer migration between cohorts over one or more time windows, in accordance with some embodiments.

[0018] FIG. 14 illustrates a simplified flow diagram of an exemplary process of predicting customer categorization and controlling retail entities, in accordance with some embodiments.

[0019] FIG. 15 illustrates an exemplary system for use in implementing systems, devices, components, methods, techniques, apparatuses, servers, sources and providing evaluation of customer information to improve the control and operation of commercial entities, in accordance with some embodiments.

[0020] Elements in the figures are illustrated for simplicity and clarity and have not necessarily been drawn to scale. For example, the dimensions and/or relative positioning of some of the elements in the figures may be exaggerated relative to other elements to help to improve understanding of various embodiments of the present invention. Also, common but well-understood elements that are useful or necessary in a commercially feasible embodiment are often not depicted in order to facilitate a less obstructed view of these various embodiments of the present invention. Certain actions and/or steps may be described or depicted in a particular order of occurrence while those skilled in the art will understand that such specificity with respect to sequence is not actually required. The terms and expressions used herein have the ordinary technical meaning as is accorded to such terms and expressions by persons skilled in the technical field as set forth above except where different specific meanings have otherwise been set forth herein.

## DETAILED DESCRIPTION

[0021] The following description is not to be taken in a limiting sense, but is made merely for the purpose of describing the general principles of exemplary embodiments. Reference throughout this specification to “one embodiment,” “an embodiment,” “some embodiments”, “an implementation”, “some implementations”, “some applications”, or similar language means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, appearances of the phrases “in one embodiment,” “in an embodiment,” “in some embodiments”, “in some implementations”, and similar language throughout this specification may, but do not necessarily, all refer to the same embodiment.

[0022] The evaluation of information is critical to the operation of retail entities as well as other commercial entities. Some embodiments evaluate customer information to improve the operation of retail entities, including but not limited to retail stores, distribution facilities, warehouses, shipping entities, suppliers, manufactures and/or other entities associated with retail and customer purchases. Large amounts of customer information can be obtained over time for numerous different customers in various geographic areas. Such information can provide insight into the operation of a retail entity and be used to identify modifications that can improve the operation of such retail entities while further improving the experience for customers. Improvements can include the improved distribution of information to enhances the operation of a retail facility and/or enhance the purchase experience off customers. Some embodiments evaluate customer information to identify modifications to be implemented at one or more retail facilities and distribute information in controlling and implementing these modifications. Customer data is processed by one or more cohort engines, algorithms, machine learning models, artificial intelligence engines and other such methods to try and understand customer behavior and identify correlations to detect multiple different types of cohorts of customers based customer behavior and customer data. Some embodiments measure and predict movements of customers across cohorts over time.

[0023] Some embodiments provide systems to control product allocation and information distribution, and can include a customer database, an algorithm database, and cohort processing systems. The customer database can store customer information corresponding to numerous different customers wherein the customer information comprises customer identifying information, customer location, purchase history information, and shopping pattern information. The algorithm database can store at least in part a plurality of evaluation algorithm models, which can comprise machine learning algorithm models, rules based algorithm models, mathematical and/or statistical based algorithm models and/or other such models, that can be applied to evaluate customer data. The cohort processing system can comprise one or more cohort evaluation control circuits communicatively coupled over one or more distributed communication networks with the customer databases and the algorithm database, and the cohort evaluation control circuit can be accessible via the distributed communication network by a plurality of remote user computing systems. In some embodiments, the cohort evaluation control circuit can be configured to control user computing systems to respectively render a cohort parameter user interface that enables a user to specify one or more customer cohort type parameters of a plurality of customer cohort types, and receive based on user interaction a specification by the user of at least a customer cohort type of the plurality of cohort types. The cohort evaluation control circuit can identify multiple different first level customizable parameters each corresponding to the customer cohort type, control the user computing system to cause the user computing system to render a hierarchical parameter interface configured to enable the user to specify for each of multiple hierarchical parameter levels one or more respective values for each of one or more different levels of customizable parameters, and receive a respective value for at least one of the customizable parameters of at least two different levels of the levels of customizable parameters dependent on the first customer cohort type. The cohort evaluation control circuit, in some embodiments, can automatically set one or more of the customizable parameters for which settings are not received as a function of at least one of the one or more setting received. One or more algorithm models, of the plurality of algorithm models, can be applied by the cohort evaluation control circuit to the customer information stored in the customer database based on the parameter values for at least one of the customizable parameters of at least two different levels of the customizable parameters, and identify through the application of the one or more algorithm models multiple cohorts of customers based on multi-dimensional correlations consistent with the parameter values for the at least one of the customizable parameters of the at least two different levels of the levels of customizable parameters.

[0024] Some embodiments provide methods of controlling retail allocation, in part by controlling, by a cohort evaluation control circuit coupled over one or more distributed communication networks with a customer database and an algorithm database, a user computing system to render a cohort parameter user interface enabling the user to specify one or more customer cohort type parameters of a plurality of customer cohort types, and receive based on user interaction a specification by the user of at least a first customer cohort type of the plurality of cohort types. Multiple different first level customizable parameters can be identified that each correspond to the first customer cohort type, and the user computing system can be controlled to cause the user computing system to render a hierarchical parameter interface configured to enable the user to specify, for each of multiple hierarchical parameter levels, one or more respective values for each of one or more different levels of customizable parameters. A respective value can be received for at least one of the customizable parameters of at least two different levels of the levels of customizable parameters dependent on the first customer grouping scheme. Some embodiments automatically set one or more of the customizable parameters for which settings are not received as a function of at least one of the one or more setting received. One or more algorithm models, of the plurality of algorithm models, can be applied to the customer information stored in the customer database based on the parameter values for at least one of the customizable parameters of at least two different levels of the customizable parameters, and multiple cohorts of customers can be identified through the application of the one or more algorithm models based on multi-dimensional correlations consistent with the parameter values for the at least one of the customizable parameters of the at least two different levels of the levels of customizable parameters.

[0025] FIG. 1 illustrates a simplified block diagram of an exemplary retail and/or retail facility control system 100, in accordance with some embodiments. The retail control system in part can evaluate customer data in controlling product distribution, product allocation, data distribution and/or other retail control. In some embodiments, the retail control system 100 includes one or more cohort processing systems 102 that comprise one or more cohort evaluation control circuits 104. The one or more cohort evaluation control circuits 104 can be communicatively coupled with one or more distributed computer and/or communication networks 106. Further, in some implementations the cohort evaluation control circuits 104 can be geographically distributed in multiple different geographic locations while communicatively coupled with each other via the one or more communication networks 106, which can enable distributed computational processing and/or provide redundancy in the event of a failure, network problems, network congestion, geographically varying demands and/or other such issues. In some embodiments the retail control system 100 includes one or more cohort migration prediction systems 103 configured to predict customer migration between cohorts. One or more of the cohort migration prediction systems 103 may be implemented as part of the cohort evaluation control circuit 104, or separate from the cohort evaluation control circuit. Further, the cohort migration prediction systems may include one or more cohort migration prediction control circuits 105. The retail control system 100 further includes one or more databases 108, which can include one or more customer databases 108a, one or more algorithm databases 108b, one or more product databases 108c, one or more inventory databases 108d, one or more machine learning model training data databases 108e, and/or other databases 108n. The databases comprise one or more memory storage systems that are communicatively coupled with one or more of the distributed communication networks 106, and in some instances are geographically distributed in multiple different geographic locations. The cohort evaluation control circuits 104 can be communicatively coupled with one or more of the databases 108.

[0026] The one or more customer databases 108a can store customer information corresponding to numerous different customers. The customer information can include, but is not limited to, customer identifying information (e.g., name, address, customer identifier number, phone number, etc.), customer location, purchase history information, shopping pattern information, product preferences, pricing preferences, other such customer information, and typically a combination of two or more of such information. The algorithm databases 108b can store a plurality of evaluation algorithm models that can be applied to customer information, product information, sales information, inventory information and/or other relevant information. The evaluation algorithm models can include, for example, machine learning algorithm models, rules based algorithm models, mathematical and/or statistical based algorithm models, and/or other such models, which can be applied to evaluate customer data, product data, inventory data, and/or other relevant data. This customer data can correspond to substantially any number of customers (e.g. tens of thousands, millions, hundreds of millions, etc.), can include numerous different types of customer information that may be relevant to evaluation, and is typically continuously updated over time as further information is accumulated and/or associated with different customers. Because of the updates to the customer information, the retail facility control system 100 and/or evaluation of the customer information can be repeated over time to update the evaluation, the identification of groupings and/or detect changes over time of groupings and/or customers' migration between cohorts over time.

[0027] Some embodiments include one or more machine learning model training systems 112 communicatively coupled over the distributed network 106 with the algorithm databases 108b and at least the model training data database. The model training system can utilize the machine learning training data to train one or more evaluation algorithm models. Further, the model training system, in some implementations, continues to acquire data and/or feedback over time to be incorporated into the model training data and used to repeatedly re-train one or more models over time to improve the effectiveness and/or accuracy of the machine learning algorithm models. Such information and/or feedback can include subsequent customer purchase information, subsequent changes in customer behaviors information, control commands executed to modify the operation of the retail system, changes in inventory, changes in sales, and/or other such information. Still further, some embodiments utilize artificially generated training and/or re-training data. Such data can be generated to simulate one or more conditions (e.g., error conditions, boundary and/or threshold conditions, desired changes in user behavior, undesirable user behaviors, other such conditions, or a combination of two or more of such information). The training data can be dependent on the type of machine learning model or models employed. The machine learning models and/or modeling applications further include the trained, deep learning models that process the data. The learning models can be substantially any relevant modeling, whether custom developed or acquired by a third party. For example, in some embodiments, the trained learning models may include decision trees, XGBOOST, GRIDSEARCHCV, unsupervised learning, regression, clustering, TENSORFLOWLITE model, MOBILENETV2 model, ML KIT for FIREBASE, and substantially any other relevant modeling and supporting applications (e.g., CORE ML, VISION FRAMEWORK, CAFFE, KERAS, XGBOOST, TENSORFLOW, etc.) to implement the modeling. Additionally, or alternatively, the machine learning models can comprise a neural network machine learning model, a convolutional neural network, Bayesian network learning, dynamically learned behavior based on, for example, decision tree learning, association rule learning, inductive logic learning, support vector learning, cluster analysis learning, Bayesian network learning, and/or similarity and metric learning, and/or other such modeling.

[0028] Some embodiments include one or more retail allocation control systems 110 that control the allocation and/or distribution of products through the retail entity, between retail facilities and/or within a retail facility, distribution of product information, allocation and/or distribution of marketing information, allocation and/or distribution of pricing and/or discounts, allocation of resources (e.g., delivery personnel, delivery vehicles, robotic systems, store associates, and/or other such retail resources), and other such allocations. The retail allocation control system 110 can, in some implementations, utilize the customer cohort associations, predicted migrations between cohort associations, causes of migration and/or other such factors as described above and further below in managing and controlling the allocation of resources, information, products and other such retail assets and inventory.

[0029] One or more product purchase systems 114 are included in the retail control system 100 to manage and complete sales of one or more products to different customers. The purchase system 114 can include associate point of sale (POS) systems at retail facilities, customer self-checkout POS systems, online POS systems, distributed purchase systems and/or other relevant systems that enable the identification of products to be purchased, and obtaining payment for those products. The retail control system 100, in some embodiments, further includes one or more inventory systems 116 that track inventory information for the numerous products available for purchase through one or more retail locations and/or entities. One or more delivery control systems 120 can be included in the retail control system 100, in some embodiments, that manage, control and implement delivery of products to customers.

[0030] The retail control system 100 further includes and/or is accessible by a plurality of different user computing systems 122. In some embodiments, for example, the cohort evaluation control circuit 104 is accessible via one or more of the distributed communication networks 106 by the plurality of remote user computing systems 122. These user computing systems 122 can include computers, laptops, smartphones, tablets, wearable computing systems and/or other such computing systems that enable users, through one or more of the distributed communication networks 106, to interact with the system, provide input to the system, obtain information from the system and implement controls over the system based on information obtained through the retail control system 100. Further, the user computing systems can be positioned locally at a retail facility or remote from one or more facilities. In some implementations, the user computing system can be implemented through multiple processors and/or microprocesses, and/or multiple computing systems that are geographically distributed at different locations and in communication via one or more of the communication networks.

[0031] In some embodiments, the cohort evaluation control circuit 104 and/or a customer data evaluation system processes extensive amounts of customer data captured corresponding to numerous different customers in numerous different geographic areas, and can identify relevant correlations and/or dissimilarities between different customer behaviors and/or patterns that can be used to control numerous different aspects of retail operation. This correlation data can be used in substantially all different levels and/or stages of retail include retail levels corresponding to the operation of brick and mortar retail facilities, as well as retail levels prior to and/or peripheral to a brick and mortal facility. Accordingly, such data correlations and/or dissimilarities can be used in part by a retail entity, product supplier, product manufacturer, product distribution entity, marketing entity, and other such entities that are related to retail to control different aspects of retail sales, product distribution, product allocation, product placement, product delivery, marketing, sales, other such retail aspects, and combinations of two or more of such aspects. Similarly, such controls can include the control of one or more automated systems in the movement and/or placement of products within a retail facility. For example, some retail facilities utilize unmanned vehicles and/or robots that can move products through a retail facility. The movement of such robots can be controlled based on one or more aspects of cohort identification in modifying product distribution and/or placement of products within a retail facility. Such modifications can be based on predicted demands as a function of determined cohorts, and certain identified products may be positioned in a location that provides more rapid retrieval, positioning of products with less predicted demand in locations that can reduce congestion, other such product distribution and typically a combination of two or more of such product allocation and/or distribution. Examples of such automated product storage and retrieval systems are described in U.S. Patent Nos.: 9,139,363; 10,508,010, 10,815,057; 10,919,701; 10,984,375; 11,142,398; 11,142,402; 11,203,486; 11,267,651; 11,315,072; 11,332,311; 11,623,826; U.S. Application Publication Nos.: US2018/0341908; US2021/0261335; US2021/0300664; and US2021/0323769; U.S. Application Nos.: 18/114,150; and International Application Publication

Nos.: WO2021243059; and WO2022133353; each of which is incorporated by reference herein in its entirety.

[0032] The one or more of cohort evaluation control circuits 104 are configured to control respective user computing systems 122 to render one or more cohort parameter user interfaces that can enable a user of that user computing system to specify one or more customer cohort type parameters of a plurality of customer cohort types. In some embodiments, the user computing systems implement a locally stored executable application that communicates with the cohort evaluation control circuit 104 and executes commands and/or implements instructions in response to commands. Based on interaction by the user with the cohort parameter user interface, the cohort evaluation control circuit 104 can receive a specification by the user of at least a first customer cohort type of the plurality of cohort types, and can identify multiple different first level customizable parameters each corresponding to the first customer cohort type. The user computing system can be controlled to cause the user computing system to render, through a display of the user computing system, a hierarchical parameter interface configured to enable the user to specify for each of multiple hierarchical parameter levels one or more respective values for each of one or more different levels of customizable parameters. The cohort evaluation control circuit 104 can receive a respective value (e.g., an entered numerical value, an entered word, a toggling of an option, an activation of an option, a selection, etc.) for at least one of the customizable parameters of at least two different levels of the levels of customizable parameters dependent on the first customer cohort type. Some embodiments automatically set one or more of the customizable parameters for which settings are not received as a function of at least one of the one or more setting received.

[0033] The cohort evaluation control circuits 104 can apply one or more algorithm models, of the plurality of algorithm models, to some or all of the customer information stored in the customer database, based on the parameter values set for at least one of the customizable parameters of at least two different levels of the customizable parameters. Multiple cohorts of customers are identified through the application of the one or more algorithm models based on multi-dimensional correlations consistent with the set customized parameter values for one or more of the customizable parameters of the at least two different levels of the levels of customizable parameters. Again, the customers can be associated with the one or more cohorts based on the determined correlations of customer behaviors. Types of customer cohort types can include, but are not limited to, one or more customer loyal cohort types, one or more brand affinity cohort types, one or more product category affinity cohort types, one or more purchase price affinity cohort types, one or more a geographic cohort types identifying correlation based on geography, marketing cohort corresponding to one or more marketing materials presented to customers, house hold size cohort, one or more holiday cohort types, other such cohort types, or combinations of two or more of such cohort types.

[0034] The identification of the cohorts of multiple customers can subsequently be utilized in controlling the operation of one or more retail entities. For example, marketing materials may be distributed that is relevant to a particular cohort of customers, placement of one or more products within a retail facility may be changed in response to customers and/or a threshold number of customers being associated with one or more cohorts, inventory system 116 can be controlled to adjust levels of one or more products based on the association of customers with one or more cohorts, pricing of one or more products may be adjusted in response to customers being associated with one or more cohorts, discounts of one or more products may be presented to customers associated with a particular cohort, other such retail controls, or a combination of two or more of such retail controls. For example, the cohort evaluation control circuit can be configured to issue instructions that when implemented control the presentation or display of a particular content to a first set of customers that have been identified in a first cohort of the multiple cohorts, and issue instructions controlling the presentation or display of a second content that is different than the first content to a second set of customers identified in a second cohort of the multiple cohorts while preventing the second content from being presented to the first set of customers based on the first set of customers being associated with the first cohort, and preventing the first content from being presented to the second set of customers based on the second set of customers being associated with the second cohort. Further, the ability to allow a user to customize the parameters enables the user to define parameter values to control evaluation of data to obtain customer information particular to that user's particular considerations and/or interest. This can greatly increase the efficiency of control of the retail entity and/or facilities in relation to identified correlations between customers.

[0035] The cohort evaluation control circuit, in some embodiments, is further configured to automatically define multiple parameter values, of at least a subset of the customizable parameters of at least one of the levels of customizable parameters, that were not defined by the user. The automatically defined values can be based on the parameter values of at least one customizable parameter of the same level or a superior level of the multiple hierarchical parameter levels. Many of the customizable parameters at subordinate hierarchical levels are dependent upon and/or limited based on one or more of the values set for one or more parameters of one or more superior hierarchical levels. As such, the cohort evaluation control circuit 104 can be configured to limit available first subordinate customizable parameters at a first subordinate level based on a first parameter value of a first customizable parameter of a superior level. This limiting can include preventing one or more customizable parameters from being available, limiting respective values and/or ranges of values that can be set for one or more customizable parameters, restricting and/or preventing some algorithm models, restricting some potential cohort types, and/or other such limits.

[0036] Further, some embodiments subdivide the customer data based on one or more user specified sub-durations of time, and evaluate the subdivided customer data. A first customizable parameter can, for example, comprise a time duration that is user definable and that restricts an evaluation by applying a set of algorithm models to customer data that occurred within the time duration. The cohort evaluation control circuit 104 can be configured to divide this time duration into multiple sub-durations that further restrict the application of one or more of the set of algorithm models to subsets of the customer data corresponding to the respective sub-durations. In some embodiments, a second customizable parameter can comprise a cohort quantity wherein the cohort evaluation control circuit 104, in identifying the multiple cohorts of customers, associates the relevant customers with a number of cohorts equal to the user specified cohort quantity.

[0037] The cohort evaluation control circuit 104, in some embodiments, can be configured to limit, at a subordinate level that is subordinate to a superior level, the plurality of evaluation algorithm models to a subset of algorithm models available to be applied, based on a first parameter value of a first customizable parameter of the superior level specified by the user, and enable a selection by the user of one or more of the limited subset of algorithm models to be applied in identifying multiple cohorts of customers. For example, a user may specify at a first hierarchy a cohort type parameter identifying a cohort type to which customers may be associated, and based on the selected cohort type the cohort evaluation control circuit 104 can limit the available algorithm models available to be applied to be applied in evaluating the customer data and associate customers with the plurality of cohorts of the selected cohort type.

[0038] In some embodiments the cohort evaluation control circuit can be configured to limit, at a second subordinate level, which is subordinate to the first subordinate level, the plurality of evaluation algorithm models to a subset of algorithm models available to be applied, based on the first parameter value of the first customizable parameter and the a second parameter value of one of the first subordinate customizable parameters, and enable a selection by the user of one or more of the limited subset of algorithm models to be applied in identifying multiple cohorts of customers. The cohort evaluation control circuit can, in some implementations, identify a first subset of at least one algorithm model, of the plurality of algorithm models, that each correspond to the selected first customer cohort type while excluding a second subset of at least one algorithm models of the plurality of algorithm models that do not correspond to the selected first customer cohort type. The cohort evaluation control circuit can further be configured to control the user computing system 122 to receive a selection of one or more retail sales channels of a plurality of retail sales channels (e.g., in-store, catalog, online, third-party service, etc.), and identify, based on the selected one or more retail sales channels, a limited set of granularity parameters limited by the selection of the first retail sales channel. The granularity parameter may include, for example but not limited to retail departments, product categories, promotions and/or promotional products, discounted products, geographic granularity parameters, other such granularity hierarchy parameters, or a combination of two or more of such granularity hierarchy parameters. A user computing system 122 can be controlled to receive a selection by the user of one or more granularity parameters of the limited set of granularity parameters. The multiple different first level customizable parameters can comprise the multiple different first level customizable parameters at a subordinate level to the selection of the one or more granularity parameters, and the control circuit in limiting the multiple different first levels of customizable parameters can limit the parameters based on the selection of the one or more granularity parameters.

[0039] FIG. 2 illustrates a simplified flow diagram of an exemplary process 200 of controlling retail allocation, in accordance with some embodiments. In step 202, the cohort evaluation control circuit 104, which is coupled over the one or more distributed communication networks 106 with the customer database 108a and the algorithm database 108b, can control a user computing system 122 to render a cohort parameter user interface enabling the user to specify one or more customer cohort type parameters of a plurality of customer cohort types and/or cohort type schemes. In step 204, a selection or specification of at least a first customer cohort type of the plurality of cohort types is received from the user based on user interaction through the graphical user interface presented through the display of the user computing system 122. The cohort types can include multiple different cohort types. For example, in relation to retail, some cohort types can include but are not limited to customer loyal cohort type, a brand or item affinity cohort type, a purchase price preference or affinity cohort type, and a geographic cohort type.

[0040] In step 206, multiple different first level customizable parameters are identified that each correspond to the first customer cohort type. Some embodiments limit available subordinate customizable parameters at a subordinate level based on a parameter value specified for a customizable parameter of a superior level. This limiting of the subordinate customizable parameters can include preventing a user from defining or selecting one or more of these limited parameters and/or excludes these limited parameters from consideration in the evaluation of customer data.

[0041] In step 208, the cohort evaluation control circuit 104 controls the user computing system 122 to cause the user computing system to render a hierarchical parameter interface configured to enable the user to specify, for each of multiple hierarchical parameter levels, one or more respective values for each of one or more different levels of customizable parameters. In step 210, a respective value is received for at least one of the customizable parameters of at least two different levels of the levels of customizable parameters dependent on the first customer grouping scheme.

[0042] In some instances, the cohort evaluation control circuit 104 can, in step 212, automatically define multiple parameter values of at least a subset of the customizable parameters of at least one of the levels of customizable parameters that were not defined by the user based on the parameter values of at least one customizable parameter of the same level or a superior level of the multiple hierarchical parameter levels. The parameters, in some embodiments, can include one or more algorithm models that are to be applied to some or all of the customer data. Different algorithm models can be relevant to different types of cohorts. In some embodiments, the cohort evaluation control circuit 104 can limit, at a subordinate level that is subordinate to a superior level, the plurality of potential evaluation algorithm models to a subset of algorithm models available to be applied based on one or more parameter values of corresponding customizable parameters of the superior level specified by the user. The user computing system can be controlled to enable a selection by the user of one or more of the limited subset of algorithm models to be applied in identifying multiple cohorts of customers. For example, in some embodiments, the cohort evaluation control circuit 104 can identify a first subset of at least one algorithm model, of the plurality of algorithm models, that each correspond to a superior level selected first customer cohort type, and further identify a second subset of at least one algorithm models of the plurality of algorithm models that do not correspond to the selected first customer cohort type. The user computing device can be controlled in presenting a user interface providing parameter options that include options corresponding to the first subset of one or more algorithm models that correspond to the selected customer cohort type, while excluding options corresponding to the second subset of at least one algorithm models that do not correspond to the selected first customer cohort type. As another example, the user computing system 122 can be controlled to receive a selection of a retail sales channel of a plurality of retail sales channels in a first hierarchical level. Based on the selected first retail sales channel, the cohort evaluation control circuit 104 can identify a limited set of granularity parameters of a subordinate or lower hierarchy second level limited by the selection of the first retail sales channel. The cohort evaluation control circuit 104 can control and/or enable the user computing system to receive a selection of one or more granularity parameters of the limited set of granularity parameters. In the identifying multiple different level customizable parameters, the cohort evaluation control circuit 104 can identify multiple different level customizable parameters at a subordinate third level that is subordinate to the second level and the selection of the one or more granularity parameters with the multiple different level of customizable parameters being limited based on the selection of the one or more granularity parameters.

[0043] In step 214, one or more algorithm models, of the plurality of algorithm models, are applied to at least some of the customer information stored in the customer database 108a based on the parameter values for at least one of the customizable parameters of at least two different levels of the customizable parameters. In step 216, multiple cohorts of customers are identified through the application of the one or more algorithm models based on multi-dimensional correlations consistent with the parameter values for the at least one of the customizable parameters of the at least two different levels of the levels of customizable parameters. For example, in some implementations, a first customizable parameter can comprise a time duration that is user definable. The time duration can be an evaluation time duration that restricts an evaluation to customer data that occurred within the time duration (e.g., purchases, adding to a virtual cart, viewing products and/or other such customer actions occurring during the specified time duration). The evaluation time duration limits the evaluation of customer data by applying a set of one or more algorithm models to customer data that occurred within the time duration. The time duration, in some embodiments, can further be divided into multiple sub-durations further restricting one or more of the set of algorithm models to subsets of the customer data corresponding to respective sub-durations. A second customizable parameter can comprise a cohort quantity wherein the cohort evaluation control circuit in identifying the multiple cohorts of customers comprises a number of cohorts equal to the cohort quantity and the limited customer data is used in associating customers with one of the number of cohorts. The cohort evaluation control circuit can identify the multiple cohorts of customers by identifying correlations and associating customers with one of the number of cohorts equal to the specified cohort quantity. The association of customers with a cohort of a cohort type can be repeated any number of times over a period of time, repeated any number of times for any number of different cohort types, and/or repeated any number of times for any number of customers. For example, over an evaluation period of time a customer's association with one of multiple cohorts of a cohort type can be repeated multiple times to identify multiple cohort associations over the evaluation period.

[0044] Some embodiments include step 220 where the cohort evaluation control circuit 104 can be configured to cause a user computing system 122 to render a multi-dimensional graphical representation graphically illustrating a representative distribution of the multiple cohorts of customers in relation to customer historic behaviors and/or illustrate a positioning within the illustrated distribution of one or more customers. In step 222, the cohort evaluation control circuit 104 can control content (e.g., product advertising, product information, discount information, sales information, promotional information, supply information, other such information, or a combination of two or more of such information) that is presented to a first set of customers identified as being associated with and/or designated to be in a first cohort of the multiple cohorts. This can include controlling presenting of different content to customers of different cohorts. For example, a first content can be identified as relevant to a first cohort and controlled to be presented to a first set of some or all of the customers identified as being associated with and/or designated in the first cohort of the multiple. Similarly, a second content, that is different than the first content, can be identified as relevant to a second cohort and controlled to be presented to a second set of some or all of the customers identified as being associated with and/or designated in a second cohort of the multiple cohorts while preventing the second content from being presented to the first set of customers based on the first set of customers being associated with the first cohort. In some instances, the first content can be prevented from being presented to the second set of customers based on the second set of customers being associated with the second cohort.

[0045] Some embodiments include step 224 the cohorts of customers are evaluated relative to product placement at one or more retail facilities, and product allocation is controlled in response to the evaluation of the cohorts relative to the product placement. This product allocation can include controlling the inventory system 116 to order additional quantities of one or more products and/or reduce ordered quantities of one or more products. In some embodiments, one or more machine learning models are applied to evaluate one or more of the identified cohorts and/or customer data of the customers associated with the identified cohorts to identify the modifications of quantities of one or more products to be carried at one or more retail facilities. For example, one or more machine learning models can identify a predicted future modification of purchasing based on correlations of customer information of customers associated with one or more of the cohorts. Similarly, the allocation of products and/or product placement within a retail facility can be controlled based on the identified cohorts. One or more steps of the process 200 and/or the process 200 can be repeated any number of times, and is typically repeated over time to determine actions to be taken in controlling one or more aspects of a retail entity and/or retail operations based on the cohort association of customers and/or changes in time of customers' cohort associations.

[0046] FIG. 3 illustrates a simplified flow diagram of an exemplary process 300 of evaluating customer information and identifying cohorts of customers based on the user defined hierarchical parameters, in accordance with some embodiments. As described above and further below, the cohort evaluation attempts to evaluate the large numbers of customers and the corresponding customer information in relation to shopping and/or purchase behaviors in an attempt to identify correlations and/or dissimilarities between such behaviors and group or associate different customers based on the threshold correlations between their behaviors. The cohort engine can help identify different behavior profiles and control retail operations based on these identified correlations. In step 301, the process is initiated. In some implementations, the process can be initiated by the customer through the selection of a cohort evaluation option provided through a graphical user interface displayed on a display of a user computing system 122. Additionally or alternatively, a user may have previously defined a schedule defining when cohort evaluations are to be implemented. For example, the user may specify a cohort evaluation to be implemented every Sunday evening. This can allow the system resources to be utilized when computational demands of the system are relatively low and further enable the cohort evaluation results to be available on Monday morning.

[0047] In step 302, the user specifies and/or selects at a hierarchical parameter level one or more commerce channel parameters. In some embodiments, the cohort evaluation control circuit 104 controls the user computing system 122 to display a graphical user interface that enables the user to select one or more of the commerce channel parameters. These commerce channel parameters can correspond to channels through which customers purchase products, shop for products, and/or otherwise acquire products through one or more retail entities. In some embodiments, the commerce channel parameters can include in-store 303, e-commerce 304 or other such remote shopping, an omni-channel 305 including both in-store and e-commerce, third party vendor channel, other such channels, or a combination of two or more of such channels.

[0048] In some embodiments, the process 300 can include step 306 where one or more granularity hierarchy parameters can be selected. Such granularity hierarchy parameters can limit the customer data and/or product information that is evaluated. For example, the granularity hierarchy parameters can include retail departments (e.g., electronics, grocery, consumables, automotive, clothing, pharmacy, alcoholic beverages, and/or other such retail departments), product categories (e.g., dairy, meat, frozen foods, canned goods, soda, cereals, games and puzzles, sports equipment, children's clothes, shoes, candy, fruits, vegetables, and/or other such product categories), promotional products, discounted products (e.g., products on sale, products associated with a discount and/or rebate, and/or other such discounts), geographic granularity parameters (e.g., zip code(s), neighborhood(s), city(ies), county(ies), state(s), west coast, Asia, etc.), other such granularity hierarchy parameters, or a combination of two or more of such granularity hierarchy parameters. In some implementations, the identification or selection of the channel parameters can limit the available subordinate level granularity hierarchy parameters when the process 300 is established with the granularity hierarchy has a subordinate level to the channel parameter. For example, one or more granularity hierarchy parameters may not be relevant to an in-store channel while instead being relevant to one or more on-line channel(s). Similarly, some granularity hierarchy parameters may correspond exclusively with in-store purchases. Still further, some embodiments provide subordinate levels of granularity based on a superior level granularity parameter selection (e.g., selecting a "sports equipment" granularity parameter, subordinate granularity parameters may be provided and/or limited to categories of sports equipment (e.g., baseball, soccer, football, basketball, swimming, gymnastics, exercise equipment, etc.). Such granularity parameters can, in some embodiments, focus or narrow the evaluation to areas and/or products of interest.

[0049] In step 307, the cohort evaluation control circuit 104 controls the user computing system 122 to render one or more cohort parameter user interfaces presenting one or more groupings or cohort type parameters 308a-308n and/or corresponding options at a subordinate or lower hierarchical parameter level. Such cohort type parameters can focus the evaluation with respect to, for example, one or more particular behaviors and/or patterns. Some examples of cohort types can include, but are not limited to, loyalty and/or engagement cohort 308a, purchase price preference cohort 308b, brand affinity cohort 308c, multi-store shopping cohort, rate of return cohort, satisfaction cohort, product rating affinity cohort, purchase frequency cohort, other such cohorts or groupings, or a combination of two or more of such cohort types. Again, in some implementations the user may be able to specify or select two or more of the cohort type parameter options each corresponding to one of the multiple different cohort types 308. Further, some embodiments enable a user to define a customized cohort type 308d by identifying a series of parameters that limit the information that is evaluated, and/or for which algorithm models may be applied.

[0050] In step 310, the cohort evaluation control circuit 104 utilizes one or more of the user cohort type parameter setting values, the channel parameter setting values and/or the granularity parameter setting values and applies a matching algorithm to identify one or more of a plurality of evaluation algorithm models that may be applied to customer information in association with the parameter setting values to identify correlations between customers and/or shopping behaviors of customers. This identification may exclude one or more of the plurality of evaluation algorithm models as not being relevant to the intended cohort and/or parameter settings. It is noted that in many instances multiple different evaluation algorithms may be applied to process customer information consistent with the selected one or more cohort types. This can, in some instances, include a matching between parameter setting values and input data structures to identify one or more algorithm models that may be available to process the customer data. In one non-limiting example, a user may have selected a loyalty cohort, and the cohort evaluation control circuit 104 may identify based the loyalty cohort selection and/or one or more other parameters that a configurable threshold algorithm model, an agglomerative scoring algorithm model and a parametric hyperboloid algorithm model can be applied. As another non-limiting example, a user may intend to obtain cohort information for a price preference cohort 108b, and based on the price preference cohort and the specified parameter values identify that the parametric hyperboloid algorithm 312c and/or the sequential repetition scoring algorithm model are relevant to determine cohorts. As described above and further below, a user may in some implementations apply implement multiple cohort types. In such instances, some implementations implement different evaluation algorithm models for the different cohort types, while other implementations apply one or more overlapping evaluation algorithm models that can be applied to each of the multiple selected cohort types. Accordingly, the system 100 not only enables customization in some implementations through the customization of one or more cohorts to be identified but in some embodiments also a customization regarding how the cohort data is to be built and the customer data is to be evaluated. Again, when multiple different cohort build methods or evaluation algorithm models are identified some embodiments can be configured to allow the user to customize the process and select one or more of the multiple build methods to be applied.

[0051] In step 311, one or more of the potential algorithm models can be identified that can be selected by the user or the cohort evaluation control circuit 104 to be applied. In some embodiments, the cohort evaluation control circuit 104 enables further hieratical customization by allowing the user to specify and/or select one or more of the identified relevant algorithm models to be applied. The user may be presented, for example, with options of one or more of the relevant algorithm models and can allow the user to select one or more of the algorithm models to be applied. In other implementations, the cohort evaluation control circuit 104 can apply a default one of the relevant algorithm models. In step 312a-312n, one or more of the algorithm models to be applied are identified (e.g., based on a detected selection by the user, the cohort evaluation control circuit 104, default, etc.). While FIG. 3 shows each of step 312a-312n, it will be appreciated that not all of the algorithm models have to be implemented, and instead those that are selected and/or otherwise identified to be activated are identified. The cohort evaluation control circuit 104 can be configured in some instances to apply each of the identified algorithm models and perform statistical evaluations to obtain relevant results, while in other instances one or more of the identified algorithm models may be selected based on expected accuracy of results, which can often be dependent on the setting values of the one or more superior level parameters. In steps 314a-314n, the cohort evaluation control circuit 104 can cause one or more of the relevant algorithm models to be applied to at least the customer data. Some example algorithm models that may be applied can include but are not limited to a configurable threshold model, an agglomerative scoring model, a parametric hyperboloid model, a sequential repetition scoring model, and/or other such algorithm models. Such models can be machine learning models trained based on historic customer information relevant to particular information associated with customer shopping behaviors and variation of behaviors over time. Such data may be actual data from actual customer interactions and/or artificially generated data. Similarly, these models can be retrained over time based on subsequent customer data obtained and/or on feedback obtained in relation to the application of the algorithm models and/or subsequent customer actions following actions taken following cohort evaluations.

[0052] Typically, each of the algorithm models include multiple further hierarchal customizable parameters, and settings values for one or more of these customizable parameters can be defined, selected and/or otherwise specified by the user through one or more graphical user interfaces displayed through the user computing system 122. As described above and further below, the applicability of many of these customizable parameters are interdependent on and/or limited by settings values specified for superior level parameters specific to the algorithm model being applied and/or one or more superior levels of the process 300.

[0053] FIGS. 4A-4B illustrate a simplified flow diagram of an exemplary threshold algorithm process 401 of implementing an exemplary configurable threshold algorithm model 314a in determining a cohort association, in accordance with some embodiments. In this example, the configurable threshold algorithm model is being applied in determining loyalty cohorts 308a. It will be appreciated that the configurable threshold algorithm model 314a can be applied relative to different cohorts and different parameter settings. In some embodiments, the threshold algorithm process 401 presents one or more hierarchical parameter interfaces that enable the user to specify for each of multiple hierarchical parameter levels one or more respective values or settings of one or more hierarchal customizable parameters. In step 402, the user computing system 122 is controlled to enable a user, in this example, to customize at a subordinate hierarchical parameter level an evaluation duration parameter (n) that defines a historic duration of time over which customer data is to be evaluated. The evaluation duration can be substantially any relevant duration that the user intends to consider. In some embodiments, the duration can be specified in weeks, days, months, and/or other durations. As one example, a user may set the evaluation duration parameter (n) as 52 weeks value, which limits the customer data considered to be limited to customer data acquired over the immediately preceding 52 weeks. Other durations may be specified. In some implementations, the duration may be limited based on one or more factors, such as but not limited to a cohort type being evaluated, one or more of the superior parameters specified in one or more superior levels of hierarchy, based on available data, other such factors, or a two or more of such factors. Similarly, some embodiments present one or more preset values from which the user can select (e.g., 2 years, 52 weeks, 6 months, three months, and/or other such preset values). Such preset values may be dependent on one or more factors. In some embodiments, the evaluation duration parameter can automatically be set, such as to a default setting value and the user can elect to continue with the automatic value or customize the duration (n). This default setting value may be varied as a function of the cohort type being evaluated, one or more of the superior parameters specified in one or more superior levels of hierarchy, based on available data, other such factors, or a two or more of such factors.

[0054] Some embodiments include step 403 where it is determined whether one or more dependent customizable parameters are to be user defined or default values utilized. Typically, the configurable threshold algorithm model 314a utilizes a set of multiple different parameters provided at one or more hierarchical parameter levels. The process advances to step 404 when default parameter values are to be set and applied. These default values may be predefined and/or determined over time as having provided more consistent and/or usable results. Further, one or more of the defaults may be dependent on values specified in one or more superior level parameters. When one or more parameters are to be customized the process proceeds to step 405 enabling the user to customize, specify and/or select one or more values for one or more parameters. The selection or specification of a value for some of the customizable parameters may cause one or more dependent subordinate parameters to be automatically set, while other subordinate parameters may additionally be set or default values utilized. For example, a user may not specify a value, based on one or more factors and/or superior level parameter values may limit the evaluation to a default duration, a user may select to use a default, and/or other such considerations.

[0055] Some embodiments include step 406 that applies one or more parameter constraints algorithms that specifies boundary limits or constraints on the values and/or ranges within which customizable values are to comply. For example, if the evaluation duration parameter is set to 52 weeks, an evaluation duration constraint algorithm can limit sub-duration thresholds or parameters to be within the specified 52 week value, and in some instances multiple sub-durations are equally spaced across the evaluation duration parameter value (e.g., if there are four sub-durations each sub-durations would be 13 weeks). As another example, with the evaluation duration parameter set to 52 weeks, an evaluation duration constraint algorithm can limit sub-durations to be within the specified 52 week value, and in some instances may further restrict a specified value for a sub-duration to comply with distribution restrictions within the specified evaluation duration (e.g., a first sub-duration parameter ($n_1$) may be restricted to being equal to or greater than n/2). A second sub-duration parameter ($n_2$) may be restricted based on the evaluation duration parameter value and the first sub-duration parameter value (e.g., $n_2 < n_1$), while a third sub-duration parameter ($n_3$) may be similarly restricted based on the second sub-duration parameter value (e.g., $(n_1/2) < n_3 < (n_2+1)$).

[0056] Other parameters that may be customized can include one or more thresholds, such as one or more percentage thresholds (q1, q2) applied to distinguish between cohorts or rankings within a cohort; quantity thresholds (e.g., number of shopping transactions thresholds (t1, t2, t3); product price thresholds; total transaction costs; etc.); and/or other such customizable parameters. One or more parameter constraints algorithms may apply two one or more of these setting limits and/or boundaries for values that can be specified. As a non-limiting example, a customizable percentage threshold q1 may define a monetary threshold corresponding to a quantity of money spent (e.g., first spending threshold (q1) can correspond to a 66% threshold where customers spending over the evaluation duration (n) is in the top 66% of customers identified within the determined cohort of customers, while a second spending threshold (q2) can correspond to a 33% threshold where customers can be identified having spending above or below the bottom 33% of customers identified within the determined cohort of customers). Still other parameters may correspond to different numbers or thresholds of purchase transactions (t1, t2, t3). Again, one or more parameter constraints algorithms may restrict and/or limit settings based on a dependency of settings of other parameters. As a non-limiting example, a user may specify a lower spending threshold (q1), and a spending constraints algorithms may restrict a second spending threshold (q2) to being limited between (q1+1) and 100% (e.g., with q1 = 40%, and 41% < q2 < 100%). As other non-limiting example, duration constraints algorithms may restrict a third sub-duration ($n_3$) based on settings values for a first sub-duration ($n_1$) and a second sub-duration ($n_2$) (e.g., with $n_1 = 26$ weeks and $n_2 = 8$ weeks, a third sub-duration ($n_3$) may be limited by 9 to 13 weeks according to $(n_2 + 1) < n_3 < (n_1/2)$); and a second transaction threshold (t2) may be constrained based on settings of a first transaction threshold (t1) and a third transaction threshold (t3) (e.g., with t1 = 15 transactions and t3 = 25 transactions, a second transaction threshold t2 may be limited to between 16 and 24 transactions (e.g., (t1 +1) < t2 < (t3

– 1)). In some embodiments, the constraints algorithms comprise predefined equations, while other algorithms are machine learning algorithms that can adjust over time based on feedback.

[0057] Some embodiments include step 407 where it is determined whether a default cohort sizing is to be applied in step 408 or the user intends to apply further customization to customize in step 410 the number or size cohorts and/or subdivisions. This enables a user to control the precision of grouping of customers based on the correlations of shopping and/or behaviors. Some embodiments define a default number of cohort divisions (P1-P4), but further allow a number of subdivisions. For example, when performing a loyalty cohort type evaluation, a default subdivision may specify a total of seven subdivisions with a first division P1 including two subdivisions, a second division including two subdivisions, a third division including a single division, and a fourth division including two subdivisions.

[0058] The cohort evaluation control circuit 104 utilizes the specified parameters 411 (whether user customized and/or default) to process customer data according to the configurable threshold algorithm model 314a. The customer data can be stored in one or more customer databases 108a and/or other storage, and can include information such as but not limited to dates of purchases and/or shopping (e.g., including a time (R) of and/or since a most recent purchase in accordance with the specified channel or channels), number of distinct weeks (F1) that a customer made a purchase, number of distinct days (F2) that a customer made a purchase, an average weekly monetary spending (M), a number of weeks (D) since a first purchase, other such customer data, and typically a combination of two or more of such data). In some embodiments, this includes step 412 where the cohort evaluation control circuit 104 evaluates customer data for each customer that is within a dataset in accordance with one or more superior parameters (e.g., within a commerce channel) to determine whether a most recent purchase time (R) (e.g., a last purchase at a particular retail entity) by a customer occurred within a first duration threshold ($n_1$) period of time. In a loyalty cohort example, this may filter customers to identify those customers that are generally associated with one or more lower level loyalty cohorts or customers that may be considered as churning customers that are shopping less or not at all with the particular retail entity.

[0059] In some implementations, when the cohort level (e.g., P4) only includes a single cohort the process can identify a customer having a most recent purchase more than the first threshold duration ($n_1$) as that single cohort (e.g., a disengaged cohort). For example, an optional step (not illustrated) may determine whether the cohort level P4 includes a single cohort (e.g., P4 = 1).

Some embodiments include a subprocess that comprises step 413 where the recent purchase time (R) corresponding to the customer being evaluated is within the evaluation duration value (n). When the recent purchase time is not within the evaluation duration (n), then the particular customer can, in step 414, be associated with a first level cohort (e.g., with loyalty cohorts, a lowest level might be considered the individual as no longer a customer or a “gone away” customer). When the recent purchase time is within the evaluation duration (n), it can be determined in step 415 whether a cohort group (P4) includes two cohorts) or more than two cohorts. When there are two cohorts and the customer’s recent purchase time is within the evaluation duration (n) then the particular customer can be associated, in step 416, with a second level cohort (e.g., with loyalty cohorts, a second level might correspond to customer that is “disengaged”). In some embodiments, when there are more than two cohorts, the particular customer can be associated in step 417, in order based on R, with one of the multiple cohorts 418 associated with a lower level (P4) (e.g., “disengaged 1” through “disengaged P4 - 1”).

[0060] Some embodiments include step 420, where customer data for customers that have a most recent purchase that is less than the first duration threshold ($n_1$) can be further processed to determine whether the particular customer has a first purchase date (D) occurring after (more recent than) a first date threshold (d) in determining whether this customer is considered a newer customer. When the customer has a first purchase date (D) after the first date threshold (d) some embodiments implement a subprocess that includes step 421 to determine whether a second cohort level includes one than one cohort (e.g., P3 = 1). When the second cohort level includes a single cohort, those customers having a first purchase date (D) that is prior to the first date threshold (d) can be associated in step 422 with that second level cohort (P3) (e.g., in a loyalty example, the customer can be associated with a “new” customer cohort). Alternatively, when the second cohort level includes more than a single cohort, the process can advance through one or more step (e.g., steps 423-425) to determine whether the first purchase date (D) correspond with and/or occurred after one or more additional date thresholds corresponding to the number of cohorts with the second level. For example, when the second cohort level includes four cohorts, step 423 can determine whether the first purchase date (D) occurred after the date threshold (d) minus the date threshold divided by one less than the number of cohorts (e.g., $D < (d – (d/(P3-1)))$ and assign a customer to a first new customer cohort 426. Similarly, for example, a step 424 can evaluate the first purchase date (D) relative to $(2*d)/(P3-1)$ in determining whether to associate a customer to a second new customer cohort 427, a step 425 can evaluate the first purchase date (D) to (d-1) in determining whether to associate a customer to a third new customer cohort 428, and a customer can otherwise be associated with a fourth new customer cohort 430 when the second level is specified to include four cohorts (P3=4 cohorts).

[0061] In some embodiments, the threshold algorithm process 401 can advance to step 431, when it is determined in step 420 that a particular customer has a first purchase date (D) occurring more recent than a first date threshold (d), where a duration (e.g., number of days or weeks (F1)) in which purchases are made by the customer are evaluated relative to a fourth duration or time quantity threshold ($n_4$) (e.g., determine whether the number of weeks (F1) during the evacuation duration (n) that a customer made a purchase is greater than the customizable fourth duration threshold ($n_4$), such as 36 weeks ($n_4$) during last 52 weeks (n)). For example, this fourth duration threshold ($n_4$) may indicate a high level of loyalty (P1) corresponding to one or more high loyalty cohorts (e.g., a “platinum” cohort and a “gold” cohort). Some embodiments may further consider spending, quantities of products purchased, rate of product purchases, and/or other behavior parameters in determining which of one or more cohorts a customer is to be associated in the current evaluation. For example, the process 401 may include a subprocess that comprises step 432 to determine whether a customer’s average weekly monetary spending (M) has a predefined relationship with a first quantity or spending threshold (q1). Some embodiments may further identify in step 433 whether a current cohort level (e.g., P1, such as a highest level loyalty) includes more than two cohorts. When the number of cohorts within a cohort level P1 is equal to two, the process can advance to step 434 to associate the current customer being evaluated with a first cohort (e.g., a highest cohort, which for example may be labeled “Platinum” or “most loyal”).

[0062] When it is determined in step 432 that a customer’s average weekly spending (M) does not have a predefined relationship with the first quantity or spending threshold (q1), the process can advance to step 435 to determine whether the customer’s average weekly spending (M) has a predefined relationship with a second quantity or spending threshold (q2). In some embodiments, when the average spending (M) does not have the predefined relationship with the second threshold (q2) the process advances to one or more further evaluation steps (e.g., steps 441, 448). When the average spending (M) does have the predefined relationship with the second threshold (q2), some embodiments continue to step i136 to determine whether a current cohort level (e.g., P1) includes more than two cohorts. When the number of cohorts within a cohort level P1 is equal to two, the process can advance to step 437 to associate the current customer being evaluated with a second cohort (e.g., a second highest cohort, which for example may be labeled "Gold" or "loyal").

[0063] In some embodiments, when it is determined in steps 433 or 436 that the number of cohorts within a cohort level P1 is not equal to two, the process can continue to apply one or more ordering steps in associating a customer with one of one or more cohorts specified for the cohort level (P1). For example, the process 401 can in some embodiments include step 438 to order or rank a customer based on the number of weeks the customer made purchases (F1), and/or step 440 there a customer is ordered or ranked based on average weekly monetary spending (M), and can associate the customer with one of the multiple cohorts 447 associated with the cohort level (P1).

[0064] Some embodiments can include one or more additional subprocesses based on the number of potential cohort levels and/or cohorts defined. For example, in some embodiments the exemplary process 401 can include a subprocess that includes step 441 where a parameter such as a number of distinct days a customer made a purchase over one or more periods of time (F2) can be evaluated relative to one or more quantity thresholds (t1), cost thresholds and/or other such thresholds. For example, it can be determined in step 441 whether a number of days when a purchase is made (F2) by the customer over a predefined duration has a predefined relationship with a first threshold number of days (t1). When the relationship exists, some embodiments advance to step 442 to determine whether the customer's average weekly spending (M) has a predefined relationship with a second quantity or spending threshold (q2). Similarly, some embodiments may additionally or alternatively include step 448, where it can be determined whether a number of days or weeks (F1) in which a customer made purchases is within one or more threshold boundaries (e.g., $(n_4 - n_3) < F1 < (n_4 - n_2)$). When the number of weeks (F1) is within the threshold boundaries, the process can advance to step 442.

[0065] In some embodiments, when it is determined in step 442 that the customer's average weekly spending (M) has a predefined relationship with a second quantity or spending threshold (q2), the process can advance to step 443 where the customer can be associated with a particular cohort (for example, when evaluating loyalty cohorts, an "bronze" cohort). Some embodiments advance to step 444 when one or more of a number of days when a purchase is made (F2) by the customer does not have a predefined relationship with a first threshold number of days (t1), when a number of weeks/days (F1) in which a customer made purchases does not have a predefined relationship with one or more threshold boundaries, and/or the customer's average weekly spending (M) does not have a predefined relationship with the second quantity or spending threshold (q2), where it can be determined whether the number of days when a purchase is made (F2) by the customer has a predefined relationship with one or more quantity threshold boundaries of time (e.g., t2 < F2 < t3), such as number of days.

[0066] In step 445 is can be determined whether the cohort level (P2) has two cohorts (e.g., “bronze” and “silver”). When the cohort level includes two cohorts, the customer being evaluated can be associated in step 446 with a second cohort (e.g., “silver” loyalty cohort). Alternatively, when the cohort level includes more than two cohorts some embodiments associate the customer in step 447 with one of the three or more cohorts associated with the cohort level (P2).

[0067] As such, some embodiments perform a multi-dimensional evaluation of customer data to identify groupings or cohorts of behaviors and/or customers based on the detected correlations and/or differences between customer behaviors. FIG. 5 illustrates a simplified block diagram of an exemplary graphical representation of a multi-dimensional cohort distribution 500 of multiple cohorts 502-507 of a cohort type (e.g., loyalty cohort type), in accordance with some embodiments. Further, the cohort distribution is based on multiple user configurable parameters. In this example, the customer data is evaluated as a function of recency (e.g., most recent ate of transaction) and/or frequency (e.g., total number of visits in a time period) of purchases (illustrated in this example along the x-axis) relative to a monetary factor (illustrated in this example along the y-axis), which may corresponding to total spent in a time period and/or average order value per visit. Example cohorts of the loyalty cohort type can include but are not limited to: Platinum cohort (e.g., regularly visiting customers doing highest annual spending); Gold cohort (e.g., customers that are less regular but do relatively high spending); Silver cohort (e.g., not very regular customers that purchased a threshold monetary value over a threshold time); Bronze cohort (e.g., customers that shop irregularly and in relatively small quantities); or New cohort (e.g., customers that first shopped within a threshold time); Lapsing cohort (e.g., customers that have not made a purchase and/or not shopped in a threshold time (e.g., >30 weeks)); and Inactive cohort (e.g., customers that have not shopped in more than a threshold period and/or purchased a threshold monetary quantity in a threshold period). Such cohorts can be graphically illustrated relative to one or more a customizable monetary thresholds q1, q2. For example, a first monetary threshold 510 (q1) can correspond to a quantity of money spent (e.g., first spending threshold (q1) can correspond to a 66% threshold where customers spending over the evaluation duration (n) is in the top 66% of customers identified within the determined cohort of customers), while a second monetary threshold 512 (q2) can correspond to a 33% threshold where customers can be identified having spending above or below the bottom 33% of customers identified within the determined cohort of customers. Other monetary thresholds can additionally or alternatively be evaluated and/or presented.

[0068] FIGS. 6A-6B illustrate a simplified flow diagram of an exemplary process 601 of implementing an exemplary hierarchical clustering or agglomerative scoring algorithm model, in accordance with some embodiments. The agglomerative scoring algorithm model, in some embodiments, divides an evaluation duration and evaluates data over multiple sub-durations or chucks of time, such as sub-sets of days or weeks. Some or all of the resulting sub-evaluations over the sub-durations can be agglomerated. In step 602 a user can customize one or more agglomerative parameters at one or more one hierarchical parameter levels. These agglomerative parameters may be dependent on the previous parameter values specified, such as but not limited to the type of cohorts being evaluated, evaluation duration and/or other such parameters. For example, a user may specify an evaluation duration parameter (n), a number of rankings (r) or groupings of a segment or chunk of time, and/or other parameters. These parameters may have default values if not customized and/or minimum values may be defined for these parameter (e.g., the evacuation duration (n) may have a minimum of one month, three weeks, four weeks, two months, and/or other such minimums, and the number of rankings r minimum may be 3 rankings, 5 rankings, or other ranking minimum). The minimums may be predefined based on one or more factors, such as but not limited to the type of cohorts being evaluated, when a last evaluation was performed, a desired level of accuracy, etc. Additionally or alternatively, some embodiments set limits on these agglomerative parameters, which can be dependent on one or more factors such as but not limited to type of cohorts, level of precision desired, response time, processing limits, and/or other such factors. In other instances, the limits are predefined and not dependent on other factors.

[0069] Some embodiments include step 603 that enables the user to customize a sub-duration windows parameter (W) to define how many windows or sub-durations a user wants to create for consideration or groupings of a segment or chunk of time. Again, some embodiments set minimums and/or maximums for this windows parameter (W). The minimum may be predefined or dependent on one or more factors such as but not limited to the evaluation duration (n), the number of ranks (r), the type of cohort, and/or other such factors. For example, the minimum may be set as the minimum of the evaluation duration divided by two (e.g., (n/2) rounded to a nearest integer (int(n/2)) such that there is at least two sub-durations.

[0070] In step 604, the customer data is accessed and partitioned or divided into multiple sub-durations as a function of time based on the evaluation duration (n) and the number of windows (W) to be evaluated over the evaluation duration. As described above and further below, the customer data typically includes historic shopping data, historic purchase data, historic customer patterns, types of payment methods used, other such data, and typically a combination of such data. For example, the customer data can include a frequency of purchases parameter (F) defining dates of purchases and/or number of purchases over a period of time (e.g., the evaluation duration of time (n)), an average spending (M) (which may be an average over a period of time (e.g., per week, per month, etc.), an average per purchase event or other spending parameter), and other relevant customer data. Based on the division of sub-durations, the customer data can be divided into W blocks 605a-605W of customer data corresponding to the different sub-durations. For example, the evaluation duration (n) may be specified as ten (10) weeks with a number of windows (W) defined as five. This would divide the ten weeks of customer data into five sub-windows of two-week sub-divisions. Some embodiments restrict the customer data with the subdivisions of customer data to relevant information used for the particular cohort being determined.

[0071] In step 606, a first ranking is performed of customers based on the individual customer data evaluated according to the customer data relevant to the intended cohort of interest to associate each customer with a ranking from 1 to r according to the partitioned window of time. For example, when considering a loyalty cohort, the customer data evaluated may include and/or be focused on the frequency of purchases (F) and the average weekly spending (M). Each customer being evaluated is then ranked or scored within respective W groupings or buckets $607_a$-$607_W$ corresponding to the sub-durations (determined by n/W). Each customer is assigned a ranking from 1 to the customized number of rankings (r). For example, for each sub-duration (n/W) if the number of rankings (r) is customized to be seven (7) each customer will be assigned a "1", a "2", a "3", a "4", a "5", a "6", or a "7" ranking depending on the customer data being evaluated (in this example, frequency of purchases (F) and the average weekly spending (M)).

[0072] In step 608, two consecutive groupings or buckets 607a-607W are aggregated (e.g., windows 1 and 2 are aggregated, windows 3 and 4 are aggregated, ..., windows W-1 and W are aggregated) and an aggregate ranking $610_a$ - $610_{W/2}$ for each customer is re-evaluated based on an aggregate of the two rankings for that customer. This aggregating of the ranking can be a summing of the multiple rankings, an average of the multiple rankings, a weighting and/or prioritization of the multiple rankings (e.g., summing or averaging with different weighting to different windows), or other such aggregating. For example, a weighting may be applied based on a time of year (e.g., holiday season(s), purchase patterns (e.g., back to school), etc.), a frequency of purchases, a quantity of purchases, categorize of purchases, other such weightings, or a combination of two or more of such weightings and/or prioritization factors. This aggregation in part tracks changes in customer behavior over time.

[0073] In some embodiments, the process 601 can implement in step 611 an addition aggregating of two consecutive aggregated groupings $610_a$ - $610_{W/2}$ of the aggregate rankings, and define second-aggregated groupings $612_a$ - $612_{W/r}$ with second-aggregated ranking for each customer based on the aggregated rankings of the two aggregated rankings for that customer (e.g., aggregated ranking of aggregate-$rank_{1,2}$ and aggregate-$rank_{3,4}$, aggregated ranking of aggregate-$rank_{5,6}$ and aggregate-$rank_{7,8}$, ... aggregated ranking of aggregate-$rank_{W-3,W-2}$ and aggregate-$rank_{W-1,W}$). The process 601 can implement repeated re-aggregation steps (e.g., 613 ... 614) consecutive re-aggregated rankings for each customer providing subsequent re-aggregated groupings $615_a$ – $615_b$ of customer re-aggregated rankings until a final re-aggregation in step 616 to aggregate customer rankings to a final re-aggregated ranking 620 for each relevant customer in a single re-aggregated grouping 617. Some embodiments further include step 621 where customers are associated with one or more cohorts based on correlations and/or differences between final re-aggregated rankings. For example, one or more ranking thresholds can be defined by the user and/or one or more default ranking thresholds may be utilized to associate a customer with a respective cohort based on a predefined relationship with one or more of the ranking thresholds (e.g., associate a user with a first cohort when the re-aggregated ranking is between two boundary ranking thresholds corresponding to the first cohort).

[0074] The agglomerative model, for example, can be beneficial based on purchase cycle of customers that can account for a minimum time after which a customer revisits not included in single aggregated view, and/or no seasonality aspect in single aggregated view. Further, the agglomerative model can provide a representative purchase cycle of customers in a generalized manner, which in some implementations can use a minimum time (purchase cycle) as least count (e.g., 8 weeks) agglomeration can be done to reach segmentation of higher time period (e.g., quarterly or yearly). The agglomerative algorithm model can capture seasonal aspect of the data, can remove biasing in higher time period segmentation if a customer is only a particular season shopper, can consider smaller period of time to combine to make a larger representation (e.g., for yearly cohort calculation) to detect identification of customer engagement over a longer period (e.g., a year), can track customers (e.g., chronological smaller durations) that can give estimates of points or durations of where the customer behavior changes relative to one or more thresholds, and/or other such benefits. As illustrated above, the agglomerative algorithm model can identify the shopping events and/or spending distribution of customers in a customized period (e.g., 8 weeks (purchase cycle)) over rolling duration windows (e.g., two weeks), and obtain a rank score for the customers in those windows. Some embodiments can sum, for any level of aggregation, individual window scores (e.g., for quarterly sum two consecutive windows, for half yearly sum 4 consecutive windows), and classify or otherwise associate to different cohorts according to a final score (e.g., different loyalty/engagement cohort buckets).

[0075] FIGS. 7A-7B illustrate a simplified flow diagram of an exemplary process 701 of implementing an exemplary parametric hyperboloid algorithm model, in accordance with some embodiments. In step 702, a customization of a number of classes, divisions or cohort quantities (CQ) is specified, selected or otherwise defined by the customer or a default number of cohort quantities (CQ) is used at a first hierarchical parameter level. In step 703 features and/or a number of features to be evaluated are specified. In some embodiments, the features are defined at a subordinate hierarchical parameter level. As described above, these parameters may be dependent on one or more superior level customizable parameters specified. For example, when a cohort type being evaluated are purchase price preference cohorts and a number of features specified is four, the four features may be dependent on the superior level designation of purchase price preference cohort and include for example price per units (K1), price per equalized units of measurement, shipping costs (K3), offer value (K4), and/or other such features. As another example, when the cohorts being evaluated are loyalty cohorts and the number of features (K) is specified as three, the features may include for example recency (K1), frequency (K2), monetary (K3) or amount paid, and/or other such monetary features. In some implementations, the number of parameters corresponds to a number of dimensions or hyperboloid divisions defined by K-1 hyperboloids.

[0076] In step 704, an aggregation time window (W) 713 or granularity of sub-duration time periods can be specified and/or a default may be used. In some implementations, the aggregation time windows divides the evaluation duration (n) into multiple sub-durations (as a non-limiting example, divide a year evaluation duration (n) into daily 705, weekly 706, bi-weekly, monthly, per transaction 707, or some other sub-duration). For example, a user may specify a daily 705 aggregation time window to be evaluated relative to an evaluation duration (n) of one year would result in 365 aggregate data points per customer. As another example, if a user selects weekly 706 aggregation time window to be evaluated relative to an evaluation duration (n) of one year would result in 52 aggregate data points per customer. As one simple example, a user may select in step 708 a weekly option and in step 710 choose a number of weeks.

[0077] In step 711, the customer data is evaluated relative to the multiple parameters (K) based on the number of cohort quantities (CQ) or cohorts for the selected aggerate time windows 713. This provides, for each customer, multiple sets ($X_1$ - $X_K$) of customer data points ($X_{K,W}$) corresponding to each feature (K) for each aggregation time window (W)). For example, for a first customer, when evaluating K features, the aggregation in step 711 would provide K sets of W aggregated data points $X_1$ - $X_W$ (i.e., aggregate $X_1 = (X_{11},X_{12},X_{13}, \ldots X_{1W})$, aggregate $X_2 = (X_{21},X_{22},X_{23}, \ldots X_{2W})$ ... aggregate $X_K = (X_{K1},X_{K2},X_{K3}, \ldots X_{KW})$). The aggregation may be a summing, an averaging, a weighted averaging or other such aggregation.

[0078] In step 714, a sub-set of customers' are sampled. For example, 5%, 10%, 20% or some other sub-set of customers are sampled. Typically, the sampling is a random sampling, however, in some implementations, the sampling may be weighted by one or more factors. In step 715, a statistical evaluation is applied for each feature (K) and the corresponding aggregated data points, such as through one or more power transforms (e.g., box cox transformation, Yeo–Johnson transformation, other such transformation, or a combination of two or more such transformations). The sampled resulting transform 719 and/or optimal lambda 716 are provided. Some embodiments include step 717 where some or all of the sampled resulting transformations 719 are processed, and in step 720 minimum and maximum processed aggregated data points are determined from the sampled customer data. In step 718 may be included where statistical outliers of sampled transformations 719 and/or maximum and/or minimums identified, disregarded and/or removed.

[0079] In step 721, a K-dimensional hyperboloid is constructed, to provide equations 722 of one less than the number of cohort quantities (i.e., CQ-1) of hyperplanes that divides the determined K-dimensional hyperplanes into CQ number of cohorts or sub-groups in relation to the sampled customer group. In step 724 customer data for the relevant customers being considered are accessed, instead of just the sample customer data obtained in step 714. In step 725, the power transform is performed on the aggregated data points 712 based on the optimal lambda 716 determined in step 715 to provide transformations 729 of the relevant customers being evaluated (again, in some embodiments, a subset of customers may be evaluated based on customized values set for one or more superior level parameters). Some embodiments include step 723 where statistical outliers of transformations 729 based on maximums and/or minimums identified and/or removed. In step 726, customers are associated with a respective one of the classes or cohorts (CQ) based on the determined CQ-1 hyperplane equations 722.

[0080] The box-cox transformation can be applied to relevant transaction data of the category for u and v separately using the $\gamma_u$ and $\gamma_v$. This can be all customer data or a subset of customer data based on filtering consistent with one or more superior level parameters customized and/or set by default. With a final value of u* and v*, some embodiments use an exemplary classification of transactions according to: if $f_1(u^*,v^*)-1 <= 0$ then “low”; if $f_2(u^*,v^*)-1 > 0$ then “high”; else “medium”. According to a {u,v} pair, transactions should fall inside a unit square.

[0081] FIG. 8 illustrates a simplified, two-dimensional representation of an exemplary hyperplane 801, with two hyperboloids 802a, 802b (CQ-1, where CQ = 3) dividing the two-dimensional hyperplane into three (CQ = 3) representative classes or cohorts 803a, 803b ... 803n (in this example 803c), in accordance with some embodiments. Customer data are evaluated relative to the CQ-1 hyperboloids 802a, 802b (defined by the determined equations 722) to identify the areas or cohorts 803a-803n into which their respective aggregated data fall. For example, the transaction relative to a category can be evaluated relative to price per unit (PPU), price per equalized unit of measure (PPEUOM), offer value (OV) or promotions (PMs), shipping costs (SC), etc. The hyperboloid boundaries 802a, 802b can be determined (e.g., in a two-dimensional plane) using a sample of transactions. Transactions can be classified (e.g., low, medium and high) according to positioning within the two-dimensional space and their relationship to the one or more hyperboloid boundaries 802a, 802b. Some embodiments can calculate how many transactions fall into the different cohort buckets 803 for each customer within customized sub-categories (e.g., customers having at least a threshold number of transactions in the considered period of time).

[0082] As one non-limiting example, for each categories, a sample number of transactions can be identified. Values for one or more parameters can be considered (e.g., PPU (u) and PPEUOM (v)), and upper and lower boundaries can be calculated to determine outliers for u and v (e.g., LB

= (Q1-1.5IQR) and UB = (Q3+1.5IQR)). Identified outlier values can be removed from u and v. A lambda (γ) can be determined for a box cox transformation of both u and v:

$$z = \begin{cases} (x^{\gamma} - 1)/\gamma, & \text{if } \gamma \neq 0; \\ \log x, & \text{if } \gamma = 0. \end{cases}$$

Values for $\gamma_u$ and $\gamma_v$ can be stored. Based on these values the box-cox transformation can be performed on the u and v to find the min(u), max(u), min(v) and max(v). Some embodiments can apply a min-max scaler on u and/or v to scale between 0 to 1.

[0083] The box-cox transformation can be applied to relevant transaction data of the category for u and v separately using the $\gamma_u$ and $\gamma_v$. This can be all customer data or a subset of customer data based on filtering consistent with one or more superior level parameters customized and/or set by default. With a final value of u* and v*, some embodiments use an exemplary classification of transactions according to: if $f_1(u^*,v^*)-1 <= 0$ then "low"; if $f_2(u^*,v^*)-1 > 0$ then "high"; else "medium". According to {u,v} pair, it should fall inside the unit square (as illustrated in FIG. 8) and hyperboloids 802a, 802b can be identified (e.g., using rectangular hyperbola equation (e.g., u=x, v=y: (1=ax)(1-ay)=1 (Eq. 1)), with (x,1) and (1,y) being points in equation 1, and providing x=1/(a-1) (Eq. 2) for $f_1$: $a=a_1$ and $f_2$: $a=a_2$. A representative area can be defined by x=0, y=0, x=1 and y=1, with areas under the curves provided according to $f_1$ from $x=x_1$ to x=1 by $= (x_1 * 1) + \int_{x_1}^{1} y dx$, using $y=x/(a_1x - 1)$ from Eq. 1 and Eq. 2, providing $= \frac{2}{a_1} + \frac{2}{a_1^2} ln|a\_1 - 1|$ (Eq. 3). Similarly, area created by $f_2$ by (0,0), (0,1), $(x_2,1)$ and (x2,0) provided according to f1 from $x=x_2$ to x=1 by $= \frac{2}{a_2} + \frac{2}{a_2^2} ln|a_2 - 1|$ (Eq. 4). The box-cox transformation and min-max scaling can spread the points, with $f_1(u,v)=(1 - a_1u)(1 - a_1v)$ and $f_2(u,v)=(1 - a_2u)(1 - a_2v)$.

[0084] Some embodiments may alternatively find non-rectangular hyperbolas by splitting the y=1 and x=1 in parts (e.g., three parts). Hyperbola equation: (1 – ax)(1 – by)=1 (Eq. 5). With $(x_1,x_2)$ and $(y_1,y_2)$, solving to find a and b for curves provides: a=(1-xy)/(x(1-y)), and b=(1-xy)/(y(1-x)). Exemplary curves would be as follows $f_1(u,v)=(1-a_1u)(1-b_1v)$, where $a=(1-x_1y_1)/(x_1(1-y_1))$ and $b=(1-x_1y_1)/(y_1(1-x_1))$; and $f_2(u,v)=(1-a_2u)(1-b_2v)$, where $a=(1-x_2y_2)/(x_2(1-y_2))$ and $b=(1-x_2y_2)/(y_2(1-x_2))$. As introduced above, the number of parameters and/or features to evaluate can be configured by the user in determining one or more cohorts and/or correlations between customers. For example, some embodiments can evaluate three features (e.g. a=Price Per Unit (PPU), b=Price per equalized unit of measure (PPEUOM), and c=percentage of discount price of the product

(Percentage Offer)) to create two boundary (surfaces) in (n-1)-dimensional plane 820 using a sample of transactions, and classify transactions (e.g., low, medium, or high according to where it belongs with respect to the determined hyperbolic curves).

[0085] FIGS. 9A-9B illustrate a simplified flow diagram of an exemplary process 901 of implementing an exemplary sequential repetition scoring algorithm model, in accordance with some embodiments. Typically, one or more customizable parameters are defined and/or specified at one or more hierarchical parameter levels. In step 902 a user can optionally specify one or more features X in a hierarchical parameter level that is typically a subordinate level to the cohort type parameter level and can be dependent on and/or limited by the superior hierarchical level selected cohort type. For example, with purchase price preference cohorts, an input feature of price quantiles can be specified. As another example with brand affinity cohorts, an input feature may be a particular brand of a product. In some implementations, the process 901 can be directed to a single dimension (e.g., with purchase price preference the process focuses on prince quantiles without consideration of other purchase price preference parameters that might additionally or alternatively be considered with other processes). Some embodiments include step 903 where a user can specify a threshold number of instances (m) of the feature X being evaluated.

[0086] In step 904, an aggregation time window (W) or granularity of sub-duration time periods can be specified and/or a default may be used. In some implementations, the aggregation time windows divides the evaluation duration (n) into sub-durations (as a non-limiting example, divide a year evaluation duration (n) into daily 905, weekly 906, bi-weekly, monthly, per transaction 907, or some other sub-duration). As one simple example, a user may select in step 908 a weekly option and in step 910 choose a number of weeks.

[0087] In step 911, the customer data in some embodiments can be divided based on the sub-durations and is evaluated relative to the feature (X) for the different aggregation time windows (W) over the evaluation duration (n). For example, when the aggregate time window (W) is specified as weekly, the evaluation can identify for each customer whether there is the presence of the feature X within each respective time window (W1, W2, ... Wn). This aggregation of customer data provides a sequence in step 912 over the evaluation duration (n) for each window relative to the instance of the feature X. In step 913, the sequence of data with the presence of the feature X in the chosen aggregation is stored.

[0088] Some embodiments include one or more steps 914, 915 where the sequence of aggregation data for each customer is evaluated. For example, when evaluating a brand affinity cohort, some embodiments in step 914 where a number of actual repeated purchases (N) for the feature X where actually taken by the respective customer (e.g., the number of separate instances the customer actually purchased brand X). As another example with evaluating brand affinity, some embodiments include step 915 where a number of instances (D) of potential repeat purchase opportunities of a product corresponding to feature X (e.g., a particular brand) are observed for the respective customer (e.g., number of times the respective customer was shopping and could have purchased a product corresponding to feature X). In step 916, a cohort relationship (A) relative to the feature X is calculated. For example with brand affinity, some embodiments calculate an affinity relationship (A) to the feature X as a ratio of number of instances of repeat purchase opportunities (N) by the number of actual purchases (D) (e.g., N/D). As one non-limiting example, a customer may make a series of five purchases as follows: (1) Cola_1, (2) Cola_1, (3) Cola_1, (4) Cola_2, and (5) Cola_1. In this example, after the first purchase (1) the number of actual repeat purchases (N) of Cola_1 is two (repeated purchase numbers (2) and (3) of Cola_1, because purchase (4) interrupted the repeat purchases), while the number of repeat purchase opportunities (D) is three (opportunity numbers (2), (3) and (4), and again opportunity number (4) interrupted the potential repeat opportunities of Cola_1). As such, a cohort relationship (A) of the affinity in this example toward Cola_1 in step 916 can be determined as 2/3, or about 67%. In continuing this example, the number of repeat purchases (N) of Cola_2 is zero (there were no repeat purchases following purchase (4) because purchase (5) interrupted a repeat purchase of Cola_2), while the number of repeat purchase opportunities (D) of Cola_2 was one (i.e., purchase number (5)). As such, the cohort relationship (A) of the affinity in this example toward Cola_2 in step 916 can be determined as 0/1, or 0%.

[0089] Some embodiments include step 917, where a ranking for the customer being evaluated affinities (A) toward cohorts (e.g., rank different brand affinities to different brands of different products). Again, this process can provide such correlations to different cohorts multiple different customers. In step 918 a threshold (m) number of instances of features (X) can be selected according to the ranking performed in step 917 for each customer (e.g., top m brands for each customer when the process 901 is applied to brand affinity). Some or all of the process 901 can be repeated any number of times for some or all of the customers to provide evaluations for different cohort evaluations. In step 920, customers can be classified into one or more cohorts based on the affinity relationships. In some embodiments, one or more ranking thresholds may be established to determine which cohort a customer is to be associated. The cohorts may be associated with a single brand and/or with multiple brands. Some embodiments can use different rankings to provide a general ranking. For example, multiple customers may each be associated with multiple different brand affinity rankings, and the multiple affinity rankings can be processed to determine an overall affinity ranking corresponding to a customer's tendencies to purchase the same brands and/or same products over time. Similarly, some embodiments can track rankings over time to detect changes in behaviors, affinities and/or tendencies.

[0090] It can be highly beneficial to understand customer behaviors, wants, expected changes over time, and how these behaviors and/or changes can affect sales. By identifying such behaviors, detecting changes over time and/or predicting changes, controls can be implemented to modify operation of multiple different aspects of retail, control product allocation through a retail entity and within retail facilities, control data distribution, enable true hyper personalization, track ever changing brand, price, flavor preferences of customer, improve customer experiences, provide targeting that is expected to be more forward-looking and focused on potential high-valued or flight-risk customers, and other such beneficial actions. Customer behavior and shopping patterns can often be transient in nature with both short and long-term impacts to retail entities (e.g., retail stores, online retail entities, distribution centers, suppliers, shipping entities, manufacturers, etc.). It can be beneficial to obtain 360-degree temporal views of customer behaviors to control product distribution, product allocation, data distribution and/or others such aspects of controlling one or more retail entities to be on top of consumer needs and desires. For example, understanding and predicting the prospects for marketing campaigns can help maintain and growth customer purchases, repeat customers, reduce customer churn, increase life-time value and/or other such benefits. Some embodiments create customer cohorts of different behavioral aspects from raw transaction data that are transient in nature. The cohorts can be cohorts for substantially any set of customers, and for substantially any given time period. Such cohorts can be determined for in-store purchases, e-comm purchases, other channels, or combination of two or more channels.

[0091] Customer behavior and shopping patterns can be captured in form of customer cohorts or segments that can be refreshed over time and tracked. Customer behavior portfolios can be created using multiple transaction related behavioral features based on multiple aspects that are associated with a customer shopping in retail. Example cohorts can include, but are not limited to, loyalty cohorts that can be based in part on spend and frequency of purchase (example labels of multiple loyalty cohorts: Platinum, Gold, Silver, Bronze, New, Lapsing, Inactive, etc.); price preference cohorts providing methods to classify customers' price preference in a category, departments, strategic business units, specifies preferred price band for categories, etc.; brand/product affinity cohorts that can categorizes customers according to their affinity towards different brands; geographic cohorts based on locations of purchases; timing cohorts related to timing within a period of time (e.g., year, month, week, holiday season, etc.); locations and/or product placement cohorts related to product placement within areas of one or more retail facilities; discount purchase cohorts related to customer actions in association with marketing; and other such cohorts. Some embodiments can further combine two or more cohorts to obtain correlations and/or distinctions between behaviors identified through different cohorts. Some embodiments can track customer cohorts, changes to cohorts ad/or track migration of customers among cohorts across different time periods. For example, in some embodiments the retail control system 100 can detect and/or track movements over n-many time windows (user-configurable) and is flexible to accommodate different type of cohorts. The retail control system 100 can provide functionality to incorporate n-numbers of static customer features (e.g., channel, demographics, marketing vehicle, etc.) that can enable analytical views to generate insights. Some embodiments can implement Markov property of movement of one or more customers behaviors and/or cohorts providing a predictive algorithm that uses historical transitions to predict the future cohort of one or more customers and/or groups of customers. The framework can provide functionality to incorporate n-many customer related features, and can provide store, E-commerce and Omni views.

[0092] Some embodiments provide a multi-layer system to analyze customer data to identify cohorts of customers. The system can include a backend (extract, transform, load (ETL)) layer that can ingest data from one or more transaction tables, and process data. Some embodiments can create static features to show different cuts across the cohort migration. One or more user interface layers can cooperate with the backend layer and database layers to provide integration taking inputs from other layers. Customer level categorization can be appended into one or more cohort databases, customer databases and/or other relevant data. As described above, some embodiments employ one or more machine learning models and/or artificial intelligence engines in identifying one or more cohorts. User configurable parameters can be specified by users and/or configurable rules used to obtain desired relevant data analysis and cohort information. Similarly, different cohort algorithm models can be applied to obtain similar cohorts. Similarly, different cohort algorithm models can be applied to obtain similar cohorts. Further, some embodiments evaluate cohorts over time. For example, a set of cohorts for a cohort type can be recreated across multiple (e.g., 4, 6, 10, etc.) different time points, which may utilize time gaps (e.g., with one quarter gap) between two consecutive time points. Again, in some embodiments some or all of the time-points and intervals can be user configurable to focus the evaluation on intended data, provide relevant cohorts, and/or identify changes over desired time periods.

[0093] In some embodiments, the retail control system 100 can process customer information to provide a multi-dimensional cross-temporal multi-channel view. FIG. 10 illustrates a simplified graphic representation of an exemplary multi-dimensional table 1000, in accordance with some embodiments. In some implementations, this view can be presented through a single table that presents one or more customer behavior-based cohorts across one or more shopping channels together in a single view along with the temporal view across n-many time points. The multi-dimensional table 1000, in some implementations, includes multiple customer entries 1002 each corresponding to a particular customer, and can include, for example, customer identifiers 1004 (e.g., unique customer number, user name, etc.), timing information 1006, time window 1008 corresponding to the window of time in which the reported information for an instance is relevant, one or more cohort identifiers 1010-1012 with which the particular customer is associated for a respective one of one or more particular cohort types 1014-1016 evaluation (e.g., with a loyalty cohort evaluation 1014 the respective customer can be associated with a loyalty cohort type 1010 corresponding to the evaluated time window; with a price preference cohort evaluation 1015 the respective customer can be associated with a respective price preference cohort 1011; and with a brand loyalty evaluation 1016 the respective customer can be associated with a brand cohort 1012 corresponding to the evaluated time window), other relevant information, and typically a combination of two or more of such information. The table view, in some instances, can show transitions of customer behavior between cohorts of one or more cohort types between time windows (e.g., a particular customer (ID: 1415101) can be identified in a first time window (e.g., July' 22) for the loyalty cohort type 1010 as being associated with a "Lapsed" cohort 1020, and in a second time window (e.g., Oct.' 22) as being associated with a "Silver" cohort 1022. Some embodiments enable the customer to filter the table and/or perform other operations through the table. This can further support downstream use-case from one single touchpoint, such as customer data platform for ingestion to a dashboard that enables real time analysis.

[0094] The retail control system 100, in some embodiments, can utilize the processed customer information and customer identified cohort associations to implement one or more controls to modify operation of one or more retail facilities and/or retail entities. Some examples can include the identification of more relevant information to be supplied to a particular customer based on the association with one or more cohorts. For example, with the association of a price preference cohort at customer and/or category levels, in combination with an association with one or more brand preference cohorts, the system can prioritize products returned from a product search (e.g., through on-line channel searching), show more personalized advertising, banners and/or other such information to more frequent customers, identify products more likely to be considered by a customer in making product suggestions and/or recommended substitutions (e.g., when a desired product is out of stock), identify complimentary products a customer is more likely to purchase in cross-selling, and/or other such information identification. Such information is more consistent with the customer's current state of interaction with the retail entity and can be more personalized, which can lead to less rejection and increased customer satisfaction. As another example, using cohort association, customers identified as being associated with a threshold loyalty can be provided priority for one or more aspects of a retail entity (e.g., prioritize in queues, discounts and/or advertising based on loyalty, etc.). Similarly, product stocking and/or placement within a retail facility can be adjusted based on an improved understanding of customer brand preference and price preference for one or more categories for sets of customers. Still further, causes of changes in customer behaviors can often more readily be identified and/or adjustments can be identified based on the customer changes between cohorts over time.

[0095] The exemplary processes and flowcharts are provided herein for illustrative purposes and are non-limiting examples of methods. One of ordinary skill in the art will recognize that exemplary methods may include more or fewer steps than those illustrated in the exemplary flowcharts, and that the steps in the exemplary flowcharts may be performed in a different order than the order shown in the illustrative flowcharts. Similarly, one or more steps of the exemplary flowcharts may be repeated one or more times.

[0096] The association of customers with different cohorts is not static. Instead, the associations are continuously varying over time as customers' behaviors change. For example, a customer A may be associated with cohort_1 at a first time T1, but be associated with cohort_2 at a second time t2 based on changing behaviors and/or actions by the user relative to at least the time interval between the first time T1 and the second time T2. Accordingly, it can be highly advantageous to identify these changes or migrations between cohorts over time. Further, such migration information can be utilized in controlling one or more aspects of the operation of retail entities, supplies, manufacturers and/or other such entities.

[0097] FIG. 11 illustrates a simplified flow diagram of an exemplary cohort migration identification process 1100, in accordance with some embodiments. In step 1102, the migration prediction control circuit 105 accesses previously identified, historic customer cohort information. The migration between cohorts can be limited to a single cohort type, which may be selected by the user, but in some implementations two or more cohorts types can be evaluated at the same time and/or cooperatively evaluated to potentially identify interdependence of causes and resulting corresponding migrations. As described above, the customer behaviors of numerous different customers can be evaluated over time to identify correlations between different users and associate groups of users into different cohorts for each of multiple different cohort types. The migration tracking is further typically focused over a limited period of time. Some embodiments include step 1103 where the migration prediction control circuit 105 accesses a number of distinct time points (H). In some embodiments, the number of distinct time points (H) are specified by the user and/or a default number of time points (H) are utilized. The number of time points can be substantially any relevant number of time points that are to be evaluated relative to changing cohort associations. For example, the number of time points (H) may be a specific number that a migration duration is to be equally divided into, a time point duration (e.g., weekly, bi-weekly, monthly, bi-monthly, tri-monthly, quarter of a year, yearly, or other such time point duration) and the migration duration is divided into a number of such time point durations, or other such specified time points or default time points. For example, the number of time points (H) may be specified as eight (8) to be applied to a two year migration duration resulting in time point durations of quarterly time points. The customer cohort information is determined relative to time. Accordingly, the customer cohort information be obtained for the respective time points providing time point cohort information 1104a-1104H at each of the time points (H). This provides distinct cohort identification information for the different customers at the respective time points over the migration duration.

[0098] In step 1105, the customer cohort information for respective customers is combined, merged and/or filtered with respective customer level attribute data 1106 obtained from the customer information maintained in the customer databases 108a. The attributes data comprises information such as but not limited to customer age or age group (e.g., 18-20, 20-25, 25-30, 30-

40, "Gen-Z", "Gen-Y", "Baby-Boomer", etc.), income group, gender, household size, marital status, home owner or renting, geographic information (e.g., where the customer lives, one or more store locations where customer shops, work location, etc.), sales channel information, annual income group, annual purchase group, and/or substantially any other relevant customer attributes information. Some embodiments include step 1107 where one or more evaluations are preformed to track over the different time points at least a subset of customers cohort association for one or more cohort types. Such evaluations, in some embodiments, can generate in step 1107 a Sankey plot that shows the migration movements and/or patterns of customers between cohorts. This evaluation can identify the transitions or flows of cohort associations of the different customers. These flows can identify customers that remain within the same cohort over one or more time points as well as those customers that are re-associated with a different cohort of a cohort type between two or more time points. In some embodiments, the cohort migration identification process 1100 can control provide the Sankey plot in step 1108 to be rendered on one or more users' computing systems 122. This evaluation can illustrate the migration of groups and/or individual users between cohorts of one or more cohort types over multiple time periods.

[0099] Still referring back to FIG. 11, in some embodiments, the cohort evaluation control circuit 104 enables the user to focus and/or filter the migration evaluation based on attributes and/or time. For example, in some implementations, the cohort evaluation control circuit 104 can be configured to enable the user computing system 122 to display a filter attribute interface that allows a user to specify and/or select one or more attributes of the customer level attributes data 1106. Additionally, or alternatively, in some embodiments the cohort evaluation control circuit 104 can be configured to control the user computing system 122 to display time filtering options in step 1112 that enable the user to specify one or more migration evaluation durations (e.g., $t_1$-$t_2$) over which the migration is to be identified. This migration evaluation duration is typically focused within the selected time periods (H) (e.g., $1 \leq t_1 < t_2 \leq H$). In other implementations, the time designations can be substantially any time duration, and time points can be determined within the specified time duration. The selected attributes can be accessed in step 1110 and can be applied in step 1114 in relation to the selected time period to filter the customer migration data focusing the customer evaluation to customers corresponding to the relevant selected attributes and the values set for those attributes. As one non-limiting example, a user may specify a first cohort (e.g., a "Silver" cohort of a loyalty cohort type), specify a timing (e.g., at time point T3 to time point T4), an age range (e.g., "Gen-Z" customers), and an income range (e.g., $50K-

$75K annually) to obtain migration information for those filtered customers that are associated with the "Silver" cohort at time point T3, and the migration of those customers to the same or different cohorts at time point T4. In some embodiments, the migration prediction control circuit 105 can control a user computing system 122 to render a filter attribute data interface that can be configured to enable the user to specify a set of one or more of the multiple attributes. This filtering provides a filtered customer migration information at step 1116 that identifies the customers' cohort associations for the filtered set of customers filtered based on the attributes, as well as corresponding migration information for this set of customers.

[0100] The cohort migration identification process 1100 identifies a subset of cohorts of customer filtered based on one or more attributes with respect to specified time periods, and the identification of customer migration or movements between associated cohorts of a cohort type. Some embodiments can further provide reports and/or graphical representations of some or all of the migrations for one or more cohorts of the filtered set of customers in relation to the one or more specified time periods. This information can be further considered to determine causes of the migrations of one or more customers and/or sub-groups of customers.

[0101] FIG. 12 illustrates a simplified flow diagram of an exemplary cause detection process 1201, in accordance with some embodiments. The cause detection process 1201 can be implemented in part based on the filtered customer migration information obtained based on the filtered customer data and resulting migration information obtained at step 1116 in the cohort migration identification process 1100, and try to determine one or more causes of changes in one or more particular behaviors of one or more customers in terms of the cohort migration. Some embodiments attempt to identify one or more customers that did not change cohorts that are similar to each customer that did change cohorts, and use customer information and/or historical differences as part of the process of identifying a cause of a migration to a different cohort. In step 1202, based on the evaluation of historic customer information, the migration prediction control circuit 105 can identify a subset of customers that were associated with one or more cohorts (e.g., cohort X) at the initial time $t_1$ of the migration evaluation duration. In some embodiments, migration prediction control circuit 105 and/or the cohort evaluation control circuit 104 provides a graphical user interface that enables the user to select the one or more cohorts of interest that are to be evaluated in identifying reasons for customers' migrations between two or more cohorts. In other embodiments, the detection process 1201 continues from the cohort migration identification process 1100 and continues the evaluation based on the one or more cohorts (X) identified in the cohort migration identification process 1100.

[0102] In step 1203, a set of migrated customers are identified that were associated with the cohort X at the initial time $t_1$, but are not associated with the cohort X at a subsequent time $t_2$, and instead are associated with a different cohort of the cohort type at time $t_2$ (e.g., no longer associated with the "Platinum" cohort of the loyalty cohort type at the subsequent time $t_2$). In step 1204, this first subgroup or identified set of migrated customers can collectively be associated and in some instances can be considered a causation group or a "focus customer base" of migrated customers. In step 1205, for each customer of the focus customer base, one or more statistical evaluations, one or more first machine learning causation models, and/or one or more algorithms are applied relative to one or more features 1206, attributes and/or other such customer information. For example, when a cohort type being evaluated is a loyalty type cohorts, parameters may include a recency parameter ($x_1$), a frequency parameter ($x_2$), a monetary parameter ($x_3$) and/or other such parameters (e.g., number of items purchased within an order, number of page views and visits, preferences to a specific store or channel, preference to a specific delivery or pickup mechanism, etc.). Other parameters may additionally or alternatively be used with loyalty cohorts and/or other cohorts. As further non-limiting examples, brand affinity cohort evaluation may consider number of purchases, number of sites or page views, number of clicks of that product, number of searches, number of times product was in cart, etc.; purchase price preference cohort provide information about price ranges customers typically and/or are comfortable buying one or more products and can consider parameters such as price per unit, promotions or discounts, buying in bulk, shipping charges or free shipping, etc.; (e.g., purchase price preference, price sensitivity,). In some embodiments, one or more of the features evaluated can be user specified through user customization as described above.

[0103] In step 1207, one or more direct causes (e.g., causation feature $x_c$) are identified that are predicted to have contributed to the respective customer's change in behavior resulting in the change in cohort at time $t_2$. Cohort identification typically is a multi-dimensional analysis of multiple different features and/or parameters. The direct cause evaluation in step 1207 can in some embodiments attempt to identify the one or more dimensions, in which the cohort is determined, that caused the change in customer behavior between time $t_1$ and time $t_2$. Such detection can be based on respective threshold changes relative to one or more features, a value associated with a feature crossing one or more thresholds, detected based on machine learning algorithms applied to some of the customer information, other such detections, or a combination of two or more of such methods of detection. Further, some embodiments may prioritize or weight the causation features to identify one or more features predicted to have had the greatest effect in changing the respective customer's behavior. One or more of the most relevant or prioritized features may be identified as the one or more primary causes of behavior change. For example, larger variations in changes may be prioritized over other changes, particular features may be prioritized over other features, machine learning algorithms or models may be applied to prioritize based on training of these algorithms using historic customer information, other such methods, or a combination of two or more of such methods.

[0104] In step 1208, a time or time period is identified that corresponds to the sufficient change in behavior corresponding to the change of association with a different cohort. In some embodiments, one or more algorithms and/or second machine learning causation or change point models are applied to the customer information relative to the one or more most relevant change features to identify based on the application of a change point model a change point in time $t_m$ when the customer behavior changed, a threshold change occurred, an inflection point of the change occurred and/or other relevant distinguishing point. This change point $t_m$ is obtained at step 1209. For example, a monetary dimension or feature may be identified as the primary influential change feature for the respective customer that has changed, the change point detection algorithm in step 1213 attempts to identify a time $t_m$ when that monetary change occurred. The one or more change point detection algorithms and/or models can be implemented through third party algorithms (e.g., Ruptures change point Python library, Changefinder algorithm library, Trend Changepoints by Prophet, etc.), and/or similar customized change point detection algorithms.

[0105] In step 1211, a non-migrated set of customers are identified by evaluating the customer cohort information relative to the migration evaluation duration to identify those customers that were identified at the initial evaluation time $t_1$ as being associated with the identified cohort X and were still associated at time $t_2$ with the identified cohort X. In step 1212, this second subgroup or identified set of non-migrated customers can be collectively associated and, in some instances, can be considered a control group or a "control customer base". For example, the cohort X can be a "Platinum" cohort of a customer loyalty cohort type, and customers can be identified that were associated with the "Platinum" cohort at time $t_1$ and are still associated with the "Platinum" cohort at time $t_2$. In step 1213, the one or more algorithms and/or machine learning change point models applied in step 1208 are similarly applied to the customer information of the second sub-group control customer base. As described above, the customers of the control customer base did not change cohorts at time $t_2$, and as such the change point model typically provide a potential change point in time $t_m$* for respective customers. The set of potential change points in time $t_m$* can subsequently be used as a check against customers of the focus customer base in confirming there are no or less than threshold fluctuations in scale. In some instances, one or more potential change points in time $t_m$* can equate to time $t_2$ (or some other time greater than $t_2$, e.g., defaulted to $t_2$+1) because of the lack of change that would be expected to otherwise cause the change in behavior. Because the association with a cohort is based on ranges relative to two or more factors, it is possible that a change point for a non-migrated customer occurred that resulted in a change within the given cohort while not enough change to cause the customer to be associated with a different cohort. In such instances a potential change point in time $t_m$* may be detected corresponding to the change within the cohort during the migration evaluation duration (e.g., $t_1$ to $t_2$).

[0106] In step 1215, customers from the first sub-group of migrated customers are evaluated relative to customers of the second sub-group of non-migrated customers in identifying corresponding or similar customers between the two sub-groups. In some embodiments, this evaluation in part attempts to identify "look-a-like" or correspondingly similar customers between the two sub-groups. For example, some embodiments evaluate, for each customer of the focus customer base migrated customers identified in step 1204 and based on the different determined direct causes $X_c$, the determined change point in time $t_m$ to the potential change points in time $t_m$* of the customers of the control customer base non-migrated customer and confirm the respective change points in time $t_m$ for a customer (e.g., customer Z) of the migrated group is less than the potential change points in time $t_m$* of one or more customers of the control customer base non-migrated customers. In some embodiments, when a threshold number of change points in time $t_m$ fail the corresponding direct cause $X_c$ may be disregarded. The step 1215 can, in some embodiments, be repeated any number of times to repeatedly evaluate the multiple migrated customers relative to multiple different customers of the non-migrated group to identify a set of non-migrated customers that have a potential change point in time $t_m$* that is later in time than the change point in time $t_m$ of the migrated customer being evaluated.

[0107] In step 1218, the process performs a control selection to identify one or more non-migrated customers that most closely correlate to and/or are similar to the migrated customer being evaluated based on similarities. In some embodiments, the control selection can be based on an evaluation of each customer of the focus customer base migrated customers relative to customers of the control customer base non-migrated customers relative to the potential change points in time $t_m$* as well as multiple different customer data factors 1216 Y1-Ym (e.g., one or more demographics factors Y1, one or more geographic factors Y2, one or more shopping frequencies Y3 (e.g., average number of visits over a period), preferred department of purchases Y4, channel of purchases factor, age factor, income factor, other such factors, and typically a combination of two or more of such factors) in identifying in step 1220, for the customer being evaluated from the migrated customer set, one or more similar or a most similar customer (e.g., control customer $Z_c$) of the control customer base non-migrated customers, identified based on the filtering. This control selection, in some implementations, identifies a non-migrated customer ($Z_c$) that has similar features and/or behaviors to those of a migrated customer being evaluated. Based on the control selection, some embodiments obtain a mapping or correlations between the migrated customers and the non-migrated customers.

[0108] In step 1221, one or more sets of attributes (f1, f2, f3, ..., fm) and/or features are identified relative to the root cause of the migration of a migrated customer to the second cohort. These attributes can be selected by the user and/or one or more attributes may be automatically selected through one or more algorithms and/or causation models (e.g., based on the cohort and/or cohort type being evaluated, timing of the evaluation, the evaluation duration, migration duration, one or more specified time points (H), etc.), the selected filtering attributes of step 1110, and/or other such information). These attributes can include substantially any number of attributes associated with customer shopping and/or purchases. Some embodiments evaluate tens, hundreds or thousands of attributes. Examples of these attributes can include but are not limited to number of returns over a period of time, number of items added to a virtual cart but not purchased over a period of time, number of items left in a virtual cart for more than one or more thresholds, number of virtual carts left without checkout for more than one or more thresholds, spending amounts, locations of purchases, departments from which products are purchased, prices of products added to a virtual cart but not purchased, number of times over a period that a customer visited a physical store, number of items purchased in a visit, color of items purchased, types of items purchased, sizes of items purchased, quantities of items purchased, number of different items considered prior to a purchase, other such attributes, and typically a combination of multiple of such attributes. These attributes can be previously specified, while in other instances a user can define attributes. The attributes considered, in some embodiments, are filtered and/or limited to those associated with the identified cause or causes associated with the change in behavior determined in step 1207.

[0109] In step 1222, a first preceding time period n1 is identified that is after the initial time t1 and is prior in time to the change point in time $t_m$ of the migrated customer being evaluated. In step 1223, a second preceding time period n2 is identified that is prior in time to the first preceding time period and the change point in time $t_m$. These preceding times can be user defined, defaults durations, determined based on an evaluation duration t1 to t2, based on a timing of the change point in time $t_m$ relative to t1 and/or t2, other such factors, or a combination of such factors. Similarly, when one or more of the first and second preceding times n1, n2 are user defined, some embodiments set limits and/or thresholds in the setting of the first and/or second preceding times n1, n2, such as but not limited a threshold duration after t1, a threshold duration prior to t2, a threshold duration prior to $t_m$, n1 is later in time than n2, other such limits, or a combination of such limits. In some implementations, the first and second preceding times can be defined as days, weeks, months or other relevant timing. For example, the first preceding time period n1 may be two, three or some other number of weeks prior to the change point in time $t_m$, and the second preceding time period n2 may be 12 weeks prior to the change point in time $t_m$.

[0110] In step 1224a, each attribute for the migrated customer of interest (Z) is identified in step 1221 is evaluated over the first time period n1 to $t_m$ to obtain a first period attribute result. For example, with an attribute of virtual cart abandonments (i.e., products virtually added to a virtual cart, but the cart never checked out), the number of cart abandonments that occurred during the time period n1 to $t_m$. As another example, with an attribute of number clothing items purchased over the first time period n1 to $t_m$ can be determined. In step 1224b, each of the identified attributes are further evaluated for the identified similar non-migrated customer ($Z_c$) of the control group is evaluated over the first time period n1 to $t_m$ to obtain a first period attribute result for the non-migrated customer. Similarly, in step 1225a, each attribute identified in step 1221 for the migrated customer of interest (Z) is evaluated over the second time period n2 to $t_m$ (or n2 to n1) to obtain a second period attribute result (e.g., the number of cart abandonments over the second time period n2 to $t_m$, number clothing items purchased over the second time period n2 to $t_m$, etc.), and in in step 1225b, each of the attributes for the non-migrated customer of interest ($Z_c$) is evaluated over the second time period n2 to $t_m$ (or n2 to n1) to obtain a second period attribute result for the non-migrated customer.

[0111] In step 1226a, for each attribute being considered in relation to the migrated customer (Z), the first period attribute result is evaluated relative to the second period attribute result corresponding customer from the non-migrated customers to determine a migrated customer attribute ratio (R_Z) 1227. This result evaluation can be a ratio, a subtraction, a statistical evaluation, other such evaluation or combination of two or more evaluations to obtain an attribute change result that defines a relationship of the change in the attribute being evaluated over the two different time periods. For example, a ratio can be generated (e.g., the first period attribute result (numerator) over the second period attribute result (denominator)). This evaluation can in part be used to determine whether the change is increase or decrease, a rate of change and/or other such conditions. In step 1226b, for each attribute being considered in relation to the non-migrated customer ($Z_c$) of the control group, the first period attribute result for the non-migrated customer is evaluated relative to the second period attribute result for the non-migrated customer producing a non-migrated customer attribute ratio ($R\_Z_c$) 1228.

[0112] Some embodiments include step 1229 where the migrated customer attribute ratio 1227 is confirmed as reasonable root cause of the change of behavior for the particular migrated customer being evaluated. In some implementations, for example, the confirmation of a reasonable root cause can be determined by determining a relationship between the migrated customer attribute ratio 1227 and the non-migrated customer attribute ratio 1228. For example, a difference between the migrated customer attribute ratio 1227 and the non-migrated customer attribute ratio 1228 may be determined providing an attribute difference factor (DF) for the respective migrated customer being evaluated. In other embodiments, the confirmation can additionally or alternatively be based on a ratio determined and/or other relationship determined between the migrated customer attribute ratio 1227 and the non-migrated customer attribute ratio 1228. Different attributes are going to affect both the migrated customers and the non-migrated customers. By determining the attribute difference factor DF relationship between the migrated customers and the non-migrated customers, the process can confirm an intensity or affect level of the attribute.

[0113] Some embodiments include step 1230 where the attribute difference factors DF for the different attributes (f1, f2, f3, ..., fm), for each migrated customer having a similar attribute identified in step 1221 that is determined to be at least a partial cause of the change in behavior, can be aggregated (e.g., averages, summed, a median determined, etc.) to provide an aggregated causation factor 1231 or attribute factor indicating a level of influence that the respective attribute had on the changed behavior of the group or a subset of migrated customers resulting in the change in cohort association. These steps can be repeated for two or more attributes, as well as repeated for multiple different migrated customers (and identified similar non-migrated customers) to produce multiple different aggregated causation factors 1231 indicating respective expected levels of influence that the different attributes had over the set of migrated customers. The process in some embodiments can further include step 1232, where the set of aggregated attribute factors can be ranked. This allows a determination of which attributes have greater influence on customer behaviors for this filtered subset of customers being evaluated.

[0114] FIG. 13 illustrates a simplified flow diagram of an exemplary migration prediction process 1300 in predicting customer migration between cohorts over one or more time windows, in accordance with some embodiments. In step 1301, for a particular customer and for a particular cohort of a cohort type (e.g., a Platinum cohort of the loyalty cohort type), the cohort information for multiple (S) different series 1302a-1302n of previously determined historic cohorts over multiple consecutive time instances can be accesses by the migration prediction control circuit 105 and divided into evaluation groups 1302a-1302n. For example, each evaluation group can correspond to a set number of a series of sequentially determined cohorts. As a non-limiting example, the number of cohort associations in a consecutive series of cohort associations within an evaluation group can be five, such that each evaluation group 1302a-1302n includes a series of five consecutive historic cohort associations. The consecutive times can, in some embodiments, correspond to the time points 1104a-1104H applied based on the H time points from the cohort migration identification process 1100 of FIG. 11. In continuing the above example with each series being based on five consecutive time points, customer cohort information can be grouped based on time points $t_1$-$t_5$, $t_2$-$t_6$, $t_3$-$t_7$ and can continue until $t_{H-5}$ through $t_{H-1}$. The number of time points in the evaluation groups is not restricted to five, and other implementations may group more or less than five sequential time points. This grouping provides multiple windows of time defined by the number of time points combined (e.g. five time points, or 3 time points, or 8 time points, or 20 time points).

[0115] In step 1303, for each of the evaluation groups 1302-1302n, the migration prediction control circuit 105 can designate a latest time point or sub-window in time as an evaluation sub-window 1304 (illustrated in FIG. 13 as a “Y” variable), while the earliest set of sub-windows are designated as training or raw data 1305 sub-windows (illustrated in FIG. 13 as an “X” raw data variable, or feature engineering data). As a non-limiting example with five sequential time points being evaluated, the fifth time point of the sequence is specified at the evaluation Y variable, while the first through fourth time points are the raw data X variable (e.g., for $t_1$-$t_5$: t5 is Y, while $t_1$-$t_4$ is evaluation; for $t_2$-$t_6$: $t_6$ is the Y, while $t_2$-$t_5$ are the evaluation; and for $t_{H-5}$ to $t_{H-1}$: $t_{H-1}$ is the Y, while $t_{H-5}$ through $t_{H-2}$ are the evaluation). This can be repeated throughout the time points and repeated for each customer.

[0116] Table 1 below is a representative example of multiple customers (c1-cn) relative to cohort identifications for each of multiple time points $t_1$-$t_H$ for a particular cohort type (in this example "Loyalty" cohort type, with cohorts: Platinum (P), Gold (G), Silver (S), Bronze (B), New (N), and Lapsing (L)).

| | $t_1$ | $t_2$ | $t_3$ | $t_4$ | $t_5$ | $t_6$ | $t_7$ | . . . | $t_{H-5}$ | $t_{H-4}$ | $t_{H-3}$ | $t_{H-2}$ | $t_{H-1}$ | $t_H$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c1 | P | P | G | G | P | P | G | | P | P | G | P | G | P |
| c2 | S | S | VS | G | G | G | B | | S | S | B | S | B | B |
| c3 | B | B | B | L | L | L | L | | L | L | L | L | L | L |
| . . . | | | | | | | | | | | | | | |
| cn | G | S | G | S | S | S | S | | B | B | B | B | S | B |

Table 1

[0117] For customer 1 (c1) the first Y variable is P ($t_5$) while a first raw data X variable ($t_1$-$t_4$) is the set (P, P, G, G); a second Y variable is P ($t_6$) while a second raw data X variable ($t_2$-$t_5$) is the set ( P, G, G, P); a third Y variable is G ($t_7$) while a third raw data X variable ($t_3$-$t_6$) is the set (G, G, P, P); ... an $n^{th}$ Y variable is G ($t_{H-1}$) while an $n^{th}$ raw data X variable ($t_{H-5}$ - $t_{H-2}$) is the set (P, P G, P). This is repeated for each customer and for each of the relevant sets of time points. Table 2 below illustrates the set of X variables, and Y variable for the time points $t_1$ - $t_5$ from Table 1 above:

| | X | Y |
|---|---|---|
| c1 | P, P, G, G | P |
| c2 | S, S, S, G | G |
| c3 | B, B, B, L | L |
| . . . | | |
| cn | G, S, G, S | S |

Table 2

[0118] In step 1306 the raw data X for the multiple different customers over the multiple different time points or windows is evaluated to identify or extract characteristics 1307 (W1, W2, W3, ..., Wp) from the raw data X and relevant to changing migration of respective customers over the multiple (S) different series 1302a-1302n of previously determined historic cohorts over multiple consecutive time instances. Some embodiments implement one or more known feature engineering techniques. In step 1308, one or more machine learning first predictive models or algorithms are developed based on the extracted characteristics 1307 in relation to the Y variable data 1304 as a predictive factor, from the multiple different customers over the multiple different time points, and the characteristics 1307. For example, some embodiments build and/or train one or more Cat-Boost model, as known in the art. In step 1310, the one or more built and/or trained models are stored.

[0119] In step 1311, a set of multiple of the most recent cohort designations for each customer being evaluated are accessed (e.g., from the customer database, a cohort database, or other relevant memory storage). For example, the most recent 4, 3, 5 or other set of cohort designations (e.g., cohort designations at time points $t_{H-3}$ through $t_H$). Some embodiments include step 1312 where this set of most recent cohort designations are used as input data to one or more known feature engineering techniques to extract predictive characteristics. In step 1313, the developed machine learning is applied to the extracted predictive characteristics, to obtain predicted cohort designations 1314 for the different customers being evaluated. In step 1315, the predicted cohort associations for the multiple customers and/or the aggregated causation factors 1231 are utilized to control and modify the operation of the retail entity. For example, such modifications can include but not limited to modifying product geographic distribution to retail facilities, modifying product placement within one or more retail facilities, modifying information distribution to one or more customers, modifying marketing campaigns and/or information distributed to different customers and/or sets of customers, controlling point of sale systems in presenting information, modifying discounts at one or more facilities, controlling the operation of one or more applications implemented on customer computing devices to present modified information, controlling customer on-line queueing to prioritize some customers, controlling on-line shopping servers regarding product availability and/or presentation to customers, other such controls, or a combination of two or more different such controls in relation to the operation of the retail system. The aggregated causation factor 1231 can be used in relation to the predicted cohort associations in attempts to prevent or alter cohort migration and/or limit cohort migration. For example, based on the predicted change of a first customer from a first cohort to a second cohort, one or more aggregated causation factors 1231 associated with the first customer and/or other customers (e.g., causes that affected other customers' behavior away from and/or toward the first and/or second cohorts) can be used to predict how changes to the retail operation may affect the first customer's future behavior in a way that would prevent the migration to the second cohort. Some embodiments apply one or more machine learning retail modification engines based on the determined root causes, and typically the predicted migration. Such retail modification models can be trained over time using historic customer data, including historic cohort associations, historic migrations and identified root causes associated with those historic migrations, as well as feedback of resulting migrations and/or lack of migrations based on historic actions taken to modify retail operations.

[0120] In some embodiments a migration prediction control circuit 105 of the cohort migration prediction system 103 can communicatively coupled over the one or more distributed communication network 106 with the customer database 108a, and is accessible via the distributed communication network by a plurality of user computing systems 122. The migration prediction control circuit 105 typically can further couple with and/or include memory storing migration prediction code, that when executed can process customer data to predict migration of customers between cohorts. In some embodiments, the migration prediction control circuit 105 can identify migration of a first set of customers between multiple cohorts of a first cohort type. One or more machine learning causation models can be applied to at least a subset of the customer information to identify, based on the application of the machine learning causation model, one or more causation factors 1231 predicted to at least partially cause a change in behavior resulting in the migration of the set of customers previously associated with a first cohort, of a first cohort type, to a second cohort of the first cohort type. Some embodiments apply one or more machine learning prediction models to at least a subset of the customer information relative to the first cohort and corresponding to a first time period, and predict future migration at a subsequent second time period of a second set of customers between the first cohort and at least the second cohorts of the first cohort type. The migration prediction control circuit 105, in some embodiments, can be configured to divide multiple consecutive cohort instance associations identified over time for each of the first set of customers into multiple series of consecutive cohort associations, and generate the first machine learning prediction model as a function of the multiple series of consecutive cohort associations. Each of the series of consecutive cohort associations includes a subset of the multiple consecutive cohort associations, with each cohort association of each of the series of consecutive cohort associations associated with a respective one multiple time points over a period of time and consecutive in time based on the association with a respective one of the time points. That each, each cohort association, as described above, is determined based on customer information occurring within and/or corresponding to a time window prior to the time point with which the respective cohort association is determined and thus associated (e.g., time divisions and/or subdivisions, in relation to an evaluation duration (n)).

[0121] In applying the prediction machine learning model, the migration prediction control circuit can in some implementations apply the first prediction model to a most recent set of multiple cohort associations for each of the first set of customers, wherein the most recent set of multiple cohort associations is more recent in time than each of the multiple series of consecutive cohort associations. One or more feature engine algorithms can be applied to each of the multiple series of consecutive cohort associations to identify characteristics relevant to migration between multiple cohorts, and generate the first machine learning prediction model directly dependent on the characteristics. In some embodiments, the migration prediction control circuit 105 can be configured to apply the feature engine algorithm to a most recent set of multiple cohort associations for each of the first set of customers. The most recent set of multiple cohort associations is more recent in time than each of the multiple series of consecutive cohort associations. Based on the application of the feature engine algorithm, predictive characteristics can be identified relevant to the most recent set of multiple cohort associations. The application of the prediction machine learning model can include applying the prediction model to the predictive characteristics associated with the first set of customers. Additionally or alternatively, the migration prediction control circuit, in identifying migration of the second set of customers, can be configured to control a user computing system 122 to render a filter attribute data interface configured to enable the user to specify a set of multiple attributes, and filter the numerous different customers as a function of the set of multiple attributes and historic customer cohort association data defining for each of the numerous different customers at multiple different time points an association with one of the multiple different cohorts of the first cohort type to identify the first set of customers in relation to the first cohort, where the first set of customers is a subset of and less than the numerous different customers.

[0122] The migration prediction control circuit 105, in some embodiments, can further be configured to identify, for each customer of the second set of customers, a corresponding customer of a set of non-migrated customers based on similarities (e.g., step 1218), and determine the first causation factor of the behavior change as a function of a correlation between customer attributes corresponding to the first customer relative to customer attributes corresponding to a second customer from the set of non-migrated customers. Some embodiments include a retail allocation control system is configured to control the distribution of information to modify retail operations, such as modify a distribution of a set of information to a subset of customers of the second set of customers predicted to move between the first cohort and the second cohort based on the predicted cohort associations for the second set of customers and the respective causation factors of the behavior change determined for each of the second set of customers.

[0123] FIG. 14 illustrates a simplified flow diagram of an exemplary process 1400 of predicting customer categorization and controlling retail entities, in accordance with some embodiments. In step 1401 the cohort migration prediction control circuit 105 is configured to control a user computing system 122 to render a filter attribute data interface that can enable a user to select and/or specify a set of multiple attributes 1110 and/or time parameters that can be utilized to focus subsequent evaluation to a filtered set of customers. The attributes correspond to attributes can include demographic attributes, customer attributes (e.g., age, geographic location, etc.), behavior attributes, shopping history attributes and/or other such attributes. For example, in some embodiments, the attributes can be and/or include attributes identified in step 1110. In step 1402, the customers and/or customer information for the numerous different customers (e.g., hundreds of thousands, millions, tens of millions, etc.) can be evaluated relative to the attributes to filter the customers as a function of the set of multiple attributes and historic customer cohort association data defining for each of the numerous different customers at multiple different time points an association with one of the multiple different cohorts of one or more cohort types. In step 1403, a set of customers can be identified based on the filtering in relation to the one or more cohorts (e.g., a first cohort). This set of customers typically is a subset of and less than the numerous different customers identified in the one or more customer databases.

[0124] In step 1404, for each customer of a set of customers, a corresponding customer of a set of non-migrated customers can be identified based on similarities. For example, the non-migrated customers can include customers identified in step 1218. In step 1405, one or more machine learning causation models can be applied to at least a subset of the customer information stored in customer database corresponding to numerous different customers and acquired over time, and typically acquired in association one or more retail entities and/or other such entities with which the customer interacts. In step 1406, at least one causation factor can be identified based on the application of the machine learning causation model. The causation factors typically are predicted to at least partially cause a change in behavior resulting in a migration of the set of customers previously associated with a first cohort, of a first cohort type, to a second cohort of the first cohort type. In some embodiments, the one or more causation factors can be determined as a function of a correlation between customer attributes corresponding to the first customer relative to customer attributes corresponding to a second customer from the set of non-migrated customers.

[0125] In step 1407, multiple consecutive cohort instance associations identified over time for each of the set of customers can be divided into multiple series of consecutive cohort associations. In step 1408, a feature engine algorithm can be applied to each of the subsets of consecutive cohort associations to identify characteristics relevant to cohort migration between multiple cohorts. Some embodiments include step 1409 where the feature engine algorithm (e.g., step 1312) can be applied to a most recent set of multiple cohort associations for each of the set of customers. The most recent set of multiple cohort associations is more recent in time than each of the multiple series of consecutive cohort associations.

[0126] In step 1410, a machine learning prediction model can be generated as a function of the multiple series of consecutive cohort associations and typically directly dependent on the identified characteristics. In step 1411, predictive characteristics can be identified relevant to the most recent set of multiple cohort associations. In step 1412, the generated machine learning prediction model can be applied to at least a subset of the customer information relative to the first cohort and corresponding to a first time period. Some embodiments in applying the prediction model apply the prediction model to the predictive characteristics associated with the set of customers. In step 1413, future migration at a subsequent second time period of a second set of customers is predicted to occur based on the application of the machine learning prediction model. The migration, for example, may be between a first cohort and one or more other cohorts of the cohort type. Further, in some embodiments, the first machine learning prediction model to a most recent set (e.g., step 1311) of multiple cohort associations for each of the set of customers, the most recent set of multiple cohort associations is more recent in time than each of the multiple series of consecutive cohort associations. Some embodiments include step 1414 where one or more retail entities and/or retail system components are controlled based on the predicted migration of customers between cohorts. Some embodiments, for example, control the distribution of information to modify a distribution of a set of information to a subset of customers of the set of customers predicted to move between a first cohort and a second cohort based on the predicted cohort associations for the set of customers and the respective one or more causation factors of the behavior change determined for each customer of the set of customers. One or more steps of the process can be repeated. Similarly, the process can be repeated any number of times for different cohorts, different cohort types, different customer sets, different times, etc.

[0127] Further, the circuits, circuitry, systems, devices, processes, methods, techniques, functionality, services, servers, sources and the like described herein may be utilized, implemented and/or run on many different types of devices and/or systems. FIG. 15 illustrates an exemplary system 1500 that may be used for implementing any of the components, circuits, circuitry, systems, functionality, apparatuses, processes, and/or other above or below mentioned systems or devices, or parts of such circuits, circuitry, functionality, systems, apparatuses, processes, or devices. For example, the system 1500 may be used to implement some or all of the cohort processing system 102, the cohort evaluation control circuit 104, cohort migration prediction systems 103, cohort migration prediction control circuits 105, model training system 112, purchase system 114, inventory system 116, user computing systems 122, delivery control systems 120, automated product storage and retrieval systems, point-of-sale (POS) systems, and/or other such systems, components, circuitry, functionality and/or devices. However, the use of the system 1500 or any portion thereof is certainly not required.

[0128] By way of example, the system 1500 may comprise one or more control circuits or processor modules 1512, one or more memory 1514, and one or more communication links, paths, buses or the like 1518. Some embodiments may include one or more user interfaces 1516, and/or one or more internal and/or external power sources or supplies 1540. The control circuit 1512 can be implemented through one or more processors, microprocessors, central processing unit, logic, local digital storage, firmware, software, and/or other control hardware and/or software, and may be used to execute or assist in executing the steps of the processes, methods, functionality and techniques described herein, and control various communications, decisions, programs, content, listings, services, interfaces, logging, reporting, etc. Further, in some embodiments, the control circuit 1512 can be part of control circuitry and/or a control system 1510, which may be implemented through one or more processors with access to one or more memory 1514 that can store instructions, code and the like that is implemented by the control circuit and/or processors to implement intended functionality. In some applications, the control circuit and/or memory may be distributed over a communications network (e.g., LAN, WAN, Internet) providing distributed and/or redundant processing and functionality. Again, the system 1500 may be used to implement one or more of the above or below, or parts of, components, circuits, systems, processes and the like.

[0129] The user interface 1516 can allow a user to interact with the system 1500 and receive information through the system. In some instances, the user interface 1516 includes a display 1522 and/or one or more user inputs 1524, such as buttons, touch screen, track ball, keyboard, mouse, etc., which can be part of or wired or wirelessly coupled with the system 1500. Typically, the system 1500 further includes one or more communication interfaces, ports, transceivers 1520 and the like allowing the system 1500 to communicate over a communication bus, a distributed computer and/or communication network 106 (e.g., a local area network (LAN), the Internet, wide area network (WAN), etc.), communication link 1518, other networks or communication channels with other devices and/or other such communications or combination of two or more of such communication methods. Further the transceiver 1520 can be configured for wired, wireless, optical, fiber optical cable, satellite, or other such communication configurations or combinations of two or more of such communications. Some embodiments include one or more input/output (I/O) ports 1534 that allow one or more devices to couple with the system 1500. The I/O ports can be substantially any relevant port or combinations of ports, such as but not limited to USB, Ethernet, or other such ports. The I/O interface 1534 can be configured to allow wired and/or wireless communication coupling to external components. For example, the I/O interface can provide wired communication and/or wireless communication (e.g., Wi-Fi, Bluetooth, cellular, RF, and/or other such wireless communication), and in some instances may include any known wired and/or wireless interfacing device, circuit and/or connecting device, such as but not limited to one or more transmitters, receivers, transceivers, or combination of two or more of such devices.

[0130] In some embodiments, the system may include one or more sensors 1526 to provide information to the system and/or sensor information that is communicated to another component, such as the cohort processing system, control circuits, inventory systems, user computing systems, automated product storage and retrieval systems, etc. The sensors can include substantially any relevant sensor, such as optical-based scanning sensors to sense and read optical patterns (e.g., bar codes), radio frequency identification (RFID) tag reader sensors capable of reading RFID tags in proximity to the sensor, distance measurement sensors (e.g., optical units, sound/ultrasound units, etc.), accelerometers, gyroscopes, weight sensors, imaging systems, and/or other such sensors. The foregoing examples are intended to be illustrative and are not intended to convey an exhaustive listing of all possible sensors. Instead, it will be understood that these teachings will accommodate sensing any of a wide variety of circumstances in a given application setting.

[0131] The system 1500 comprises an example of a control and/or processor-based system with the control circuit 1512. Again, the control circuit 1512 can be implemented through one or more processors, controllers, central processing units, logic, software and the like. Further, in some implementations the control circuit 1512 may provide multiprocessor functionality.

[0132] The memory 1514, which can be accessed by the control circuit 1512, typically includes one or more processor-readable and/or computer-readable media accessed by at least the control circuit 1512, and can include volatile and/or nonvolatile media, such as RAM, ROM, EEPROM, flash memory and/or other memory technology. Further, the memory 1514 is shown as internal to the control system 1510; however, the memory 1514 can be internal, external or a combination of internal and external memory. Similarly, some or all of the memory 1514 can be internal, external or a combination of internal and external memory of the control circuit 1512. The external memory can be substantially any relevant memory such as, but not limited to, solid-state storage devices or drives, hard drive, one or more of universal serial bus (USB) stick or drive, flash memory secure digital (SD) card, other memory cards, and other such memory or combinations of two or more of such memory, and some or all of the memory may be distributed at multiple locations over the computer network 106. The memory 1514 can store code, software, executables, scripts, data, content, lists, programming, programs, log or history data, user information, customer information, product information, and the like. While FIG. 15 illustrates the various components being coupled together via a bus, it is understood that the various components may actually be coupled to the control circuit and/or one or more other components directly.

[0133] In some embodiments, retail facility control systems are provided comprising: a customer database storing customer information corresponding to numerous different customers; an algorithm database storing a plurality of evaluation algorithm models; a cohort processing system comprising a cohort evaluation control circuit communicatively coupled over one or more distributed communication networks with the customer database and the algorithm database, and wherein the cohort evaluation control circuit is accessible via the distributed communication network by a plurality of remote user computing systems, wherein the cohort evaluation control circuits is configured to: control a first user computing system to render a cohort parameter user interface enabling the user to specify one or more customer cohort type parameters of a plurality of customer cohort types, and receive based on user interaction a specification of at least a first customer cohort type of the plurality of cohort types; identify multiple different first level customizable parameters each corresponding to the first customer cohort type; cause the first user computing system to render a hierarchical parameter interface configured to enable the user to specify for each of multiple hierarchical parameter levels one or more respective values for each of one or more different levels of customizable parameters, and receive a respective value for at least one of the customizable parameters of at least two different levels of the levels of customizable parameters dependent on the first customer cohort type; apply one or more algorithm models, of the plurality of algorithm models, to the customer information stored in the customer database based on the customized parameter values for at least one of the customizable parameters of at least two different levels of the customizable parameters, and identify through the application of the one or more algorithm models multiple cohorts of customers based on multi-dimensional correlations consistent with the parameter values for the at least one of the customizable parameters of the at least two different levels of the levels of customizable parameters.

[0134] Some embodiments provide methods of controlling retail facility comprising: controlling, by a cohort evaluation control circuit coupled over one or more distributed communication networks with a customer database and an algorithm database, a first user computing system to render a cohort parameter user interface enabling the user to specify one or more customer cohort type parameters of a plurality of customer cohort types, and receive based on user interaction a specification of at least a first customer cohort type of the plurality of cohort types; identifying multiple different first level customizable parameters each corresponding to the first customer cohort type; causing the first user computing system to render a hierarchical parameter interface configured to enable the user to specify, for each of multiple hierarchical parameter levels, one or more respective values for each of one or more different levels of customizable parameters, and receiving a respective value for at least one of the customizable parameters of at least two different levels of the levels of customizable parameters dependent on the first customer cohort type; applying one or more algorithm models, of the plurality of algorithm models, to the customer information stored in the customer database based on the parameter values for at least one of the customizable parameters of at least two different levels of the customizable parameters; and identifying through the application of the one or more algorithm models multiple cohorts of customers based on multi-dimensional correlations consistent with the parameter values for the at least one of the customizable parameters of the at least two different levels of the levels of customizable parameters.

[0135] Some embodiments provide retail facility control systems, comprising: a customer database; an algorithm database; a cohort processing system comprising a cohort evaluation control circuit configured to: control a user computing system to render a cohort parameter user interface enabling the user to specify customer cohort type parameters, and receive based on user interaction a specification of at least a first customer cohort type of the plurality of cohort types; identify multiple different first level customizable parameters each corresponding to the first customer cohort type; cause the first user computing system to render a hierarchical parameter interface configured to enable the user to specify respective values for different levels of customizable parameters dependent on the first customer cohort type; apply one or more algorithm models to the customer information based on the customized parameter values and identify cohorts of customers.

[0136] Some embodiments provide systems to predict reallocation of customer categorization, comprising: a customer database storing customer information corresponding to numerous different customers; and a cohort migration prediction system comprising a migration prediction control circuit communicatively coupled over a distributed communication network with the customer database, and wherein the migration prediction control circuit is accessible via the distributed communication network by a plurality of user computing systems, and the migration prediction control circuit is configured to: identify migration of a first set of customers between multiple cohorts of a first cohort type; apply a first machine learning causation model to at least a first subset of the customer information and identify, based on the application of the machine learning causation model, at least a first causation factor predicted to at least partially cause a change in behavior resulting in the migration of the first set of customers previously associated with a first cohort, of the first cohort type, to a second cohort of the first cohort type; and apply a first machine learning prediction model to at least a first subset of the customer information relative to the first cohort and corresponding to a first time period, and predict future migration at a subsequent second time period of a second set of customers between the first cohort and at least the second cohorts of the first cohort type.

[0137] Some embodiments provide methods to predict customer categorization, comprising: identifying migration of a first set of customers between multiple cohorts of a first cohort type; applying a first machine learning causation model to at least a first subset of the customer information and identify, based on the application of the machine learning causation model, at least a first causation factor predicted to at least partially cause a change in behavior resulting in the migration of the first set of customers previously associated with a first cohort, of the first cohort type, to a second cohort of the first cohort type; and applying a first machine learning prediction model to at least a first subset of the customer information relative to the first cohort and corresponding to a first time period, and predict future migration at a subsequent second time period of a second set of customers between the first cohort and at least the second cohorts of the first cohort type.

[0138] Those skilled in the art will recognize that a wide variety of other modifications, alterations, and combinations can also be made with respect to the above described embodiments without departing from the scope of the invention, and that such modifications, alterations, and combinations are to be viewed as being within the ambit of the inventive concept.

# Claims

1. A system to predict reallocation of customer categorization, comprising:

a customer database storing customer information corresponding to numerous different customers; and

a cohort migration prediction system comprising a migration prediction control circuit communicatively coupled over a distributed communication network with the customer database, and wherein the migration prediction control circuit is accessible via the distributed communication network by a plurality of user computing systems, and the migration prediction control circuit is configured to:

identify migration of a first set of customers between multiple cohorts of a first cohort type;

apply a first machine learning causation model to at least a first subset of the customer information and identify, based on the application of the machine learning causation model, at least a first causation factor predicted to at least partially cause a change in behavior resulting in the migration of the first set of customers previously associated with a first cohort, of the first cohort type, to a second cohort of the first cohort type; and

apply a first machine learning prediction model to at least the first subset of the customer information relative to the first cohort and corresponding to a first time period, and predict future migration at a subsequent second time period of a second set of customers between the first cohort and at least the second cohorts of the first cohort type.

2. The system of claim 1, wherein the migration prediction control circuit is configured to divide multiple consecutive cohort instance associations identified over time for each of the first set of customers into multiple series of consecutive cohort associations, and generate the first machine learning prediction model as a function of the multiple series of consecutive cohort associations.

3. The system of claim 2, wherein the migration prediction control circuit in applying the first machine learning prediction model comprises applying the first prediction model to a most recent set of multiple cohort associations for each of the first set of customers, wherein the most recent set of multiple cohort associations is more recent in time than each of the multiple series of consecutive cohort associations.

4. The system of claim 2, wherein the migration prediction control circuit is further configured to apply a feature engine algorithm to each of the multiple series of consecutive cohort associations to

identify characteristics relevant to migration between multiple cohorts, and generates the first machine learning prediction model directly dependent on the characteristics.

5. The system of claim 4, wherein the migration prediction control circuit is further configured to apply the feature engine algorithm to a most recent set of multiple cohort associations for each of the first set of customers, wherein the most recent set of multiple cohort associations is more recent in time than each of the multiple series of consecutive cohort associations, and identify predictive characteristics relevant to the most recent set of multiple cohort associations; and

wherein the migration prediction control circuit in applying the first machine learning prediction model comprises applying the first prediction model to the predictive characteristics associated with the first set of customers.

6. The system of claim 5, wherein the migration prediction control circuit, in identifying migration of the second set of customers, is configured to control a user computing system to render a filter attribute data interface configured to enable a user to specify a set of multiple attributes, and filter the numerous different customers as a function of the set of multiple attributes and historic customer cohort association data defining for each of the numerous different customers at multiple different time points an association with one of the multiple different cohorts of the first cohort type to identify the first set of customers in relation to the first cohort, wherein the first set of customers is a subset of and less than the numerous different customers.

7. The system of claim 5, wherein the migration prediction control circuit is further configured to identify, for each customer of the second set of customers, a corresponding non-migrated customer of a set of non-migrated customers based on similarities, and determine the first causation factor of the behavior change as a function of a correlation between customer attributes corresponding to a receptive first customer relative to customer attributes corresponding to a second customer from the set of non-migrated customers.

8. The system of claim 7, further comprising:

a retail allocation control system is configured to control the distribution of information to modify a distribution of a set of information to a subset of customers of the second set of customers predicted to move between the first cohort and the second cohort based on the predicted cohort associations for the second set of customers and respective causation factors of the behavior change determined for each of the second set of customers.

9. The system of claim 1, wherein the migration prediction control circuit is further configured to identify, for each customer of the second set of customers, a corresponding non-migrated customer of a set of non-migrated customers based on similarities, and determine the first causation factor of the behavior change as a function of a correlation between customer attributes corresponding to a first customer relative of the second set of customers to customer attributes corresponding to a second customer from the set of non-migrated customers.

10. The system of claim 1, wherein the customer information comprises customer identifying information, customer location, purchase history information, and shopping pattern information.

11. A method to predict customer categorization, comprising:

identifying migration of a first set of customers between multiple cohorts of a first cohort type;

applying a machine learning causation model to at least a first subset of customer information stored in customer database corresponding to numerous different customers and acquired over time;

identifying, based on the application of the machine learning causation model, at least a first causation factor predicted to at least partially cause a change in behavior resulting in the migration of the first set of customers previously associated with a first cohort, of the first cohort type, to a second cohort of the first cohort type;

applying a first machine learning prediction model to at least the first subset of the customer information relative to the first cohort and corresponding to a first time period, and

predicting future migration at a subsequent second time period of a second set of customers between the first cohort and at least the second cohort of the first cohort type.

12. The method of claim 11, further comprising:

dividing multiple consecutive cohort instance associations identified over time for each of the first set of customers into multiple series of consecutive cohort associations; and

generating the first machine learning prediction model as a function of the multiple series of consecutive cohort associations.

13. The method of claim 12, wherein the applying the first machine learning prediction model comprises applying the first prediction model to a most recent set of multiple cohort associations for each of the first set of customers, wherein the most recent set of multiple cohort associations is more recent in time than each of the multiple series of consecutive cohort associations.

14. The method of claim 12, further comprising:

applying a feature engine algorithm to each of the multiple series of consecutive cohort associations to identify characteristics relevant to migration between multiple cohorts, and generates the first machine learning prediction model directly dependent on the characteristics.

15. The method of claim 14, further comprising:

applying the feature engine algorithm to a most recent set of multiple cohort associations for each of the first set of customers, wherein the most recent set of multiple cohort associations is more recent in time than each of the multiple series of consecutive cohort associations; and

identifying predictive characteristics relevant to the most recent set of multiple cohort associations;

wherein the applying the first machine learning prediction model comprises applying the first prediction model to the predictive characteristics associated with the first set of customers.

16. The method of claim 15, wherein the identifying migration of the second set of customers further comprises:

controlling a user computing system to render a filter attribute data interface configured to enable a user to specify a set of multiple attributes; and

filtering the numerous different customers as a function of the set of multiple attributes and historic customer cohort association data defining for each of the numerous different customers at multiple different time points an association with one of the multiple cohorts of the first cohort type to identify the first set of customers in relation to the first cohort, wherein the first set of customers is a subset of and less than the numerous different customers.

17. The method of claim 15, further comprising:

identifying, for each customer of the second set of customers, a corresponding customer of a set of non-migrated customers based on similarities; and determining the first causation factor of the behavior change as a function of a correlation between customer attributes corresponding to a respective first customer relative to customer attributes corresponding to a second customer from the set of non-migrated customers.

18. The method of claim 17, further comprising:

controlling distribution of information to modify a distribution of a set of information to a subset of customers of the second set of customers predicted to move between the first cohort and the second cohort based on the predicted cohort associations for the second set of customers and a respective causation factors of the behavior change determined for each of the second set of customers.

19. The method of claim 11, further comprising:

identifying, for each customer of the second set of customers, a corresponding customer of a set of non-migrated customers based on similarities; and

determining the first causation factor of the behavior change as a function of a correlation between customer attributes corresponding to a first customer relative of the second set of customers to customer attributes corresponding to a second customer from the set of non-migrated customers.

20. The method of claim 11, wherein the customer information comprises customer identifying information, customer location, purchase history information, and shopping pattern information.

FIG. 1

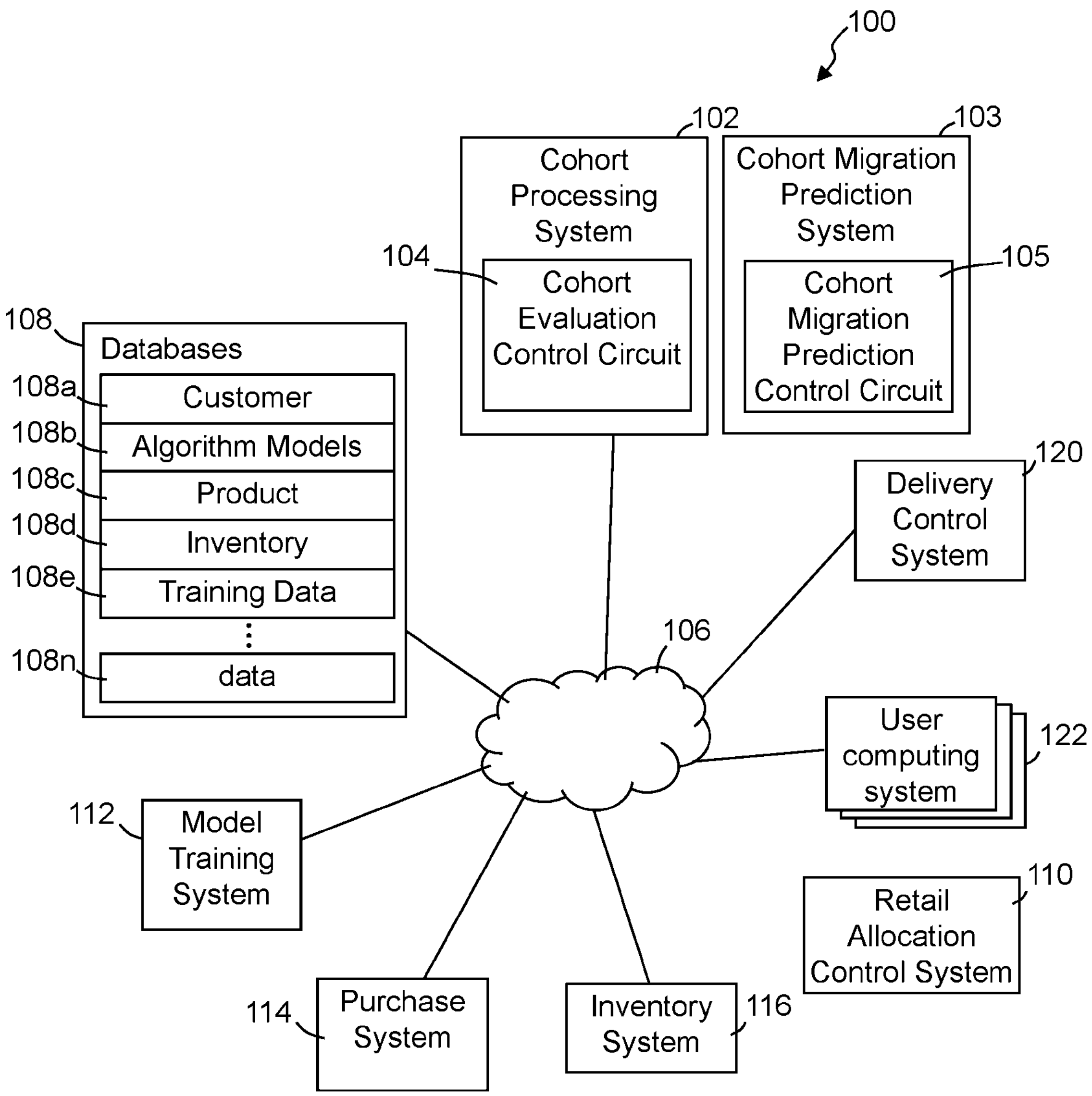

100
102
Cohort Processing System
104
Cohort Evaluation Control Circuit
103
Cohort Migration Prediction System
105
Cohort Migration Prediction Control Circuit
108
Databases
108a
Customer
108b
Algorithm Models
108c
Product
108d
Inventory
108e
Training Data
108n
data
106
120
Delivery Control System
122
User computing system
110
Retail Allocation Control System
112
Model Training System
114
Purchase System
116
Inventory System

*FIG. 2*

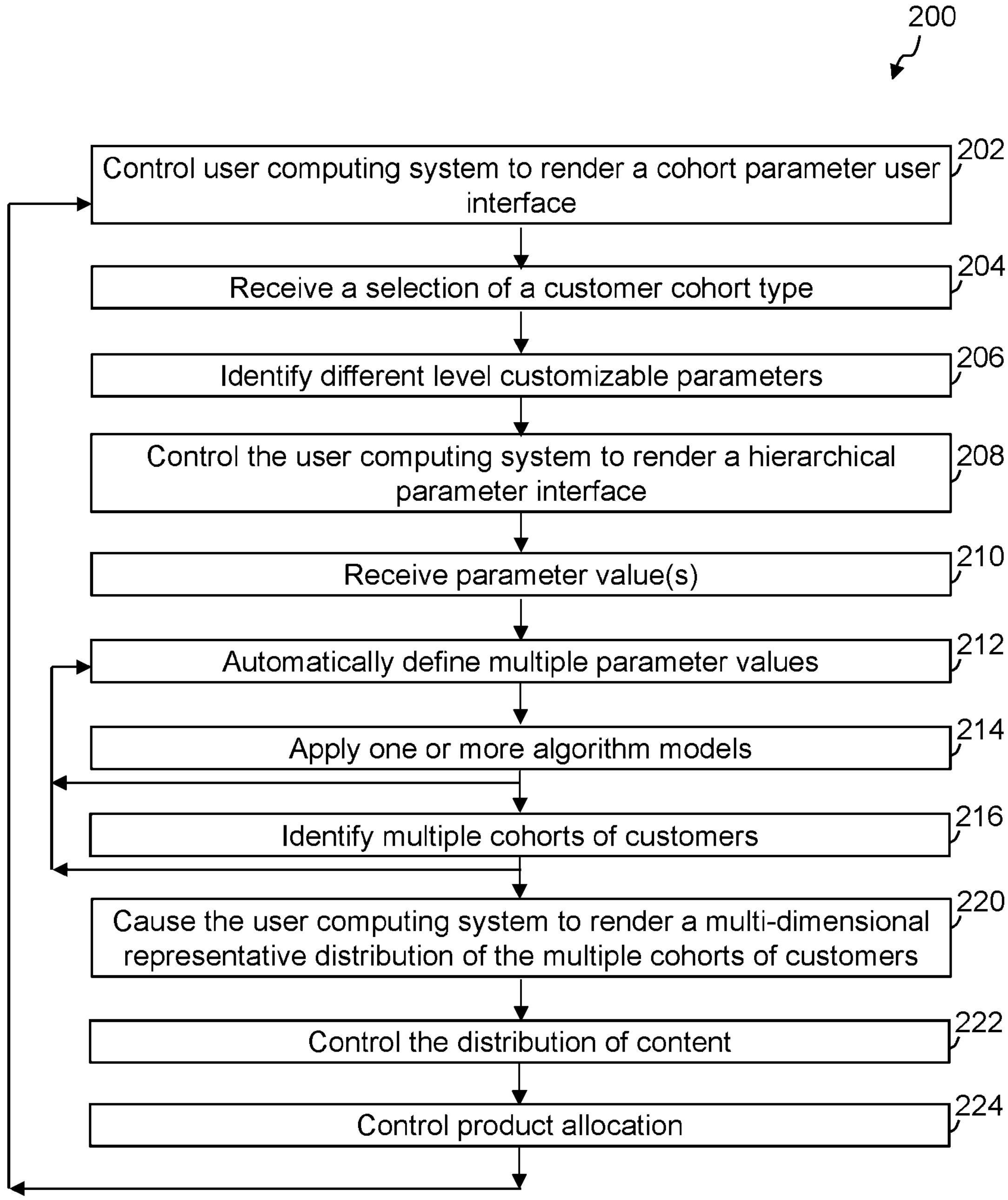

200
Control user computing system to render a cohort parameter user interface
202
Receive a selection of a customer cohort type
204
Identify different level customizable parameters
206
Control the user computing system to render a hierarchical parameter interface
208
Receive parameter value(s)
210
Automatically define multiple parameter values
212
Apply one or more algorithm models
214
Identify multiple cohorts of customers
216
Cause the user computing system to render a multi-dimensional representative distribution of the multiple cohorts of customers
220
Control the distribution of content
222
Control product allocation
224

*FIG. 3*

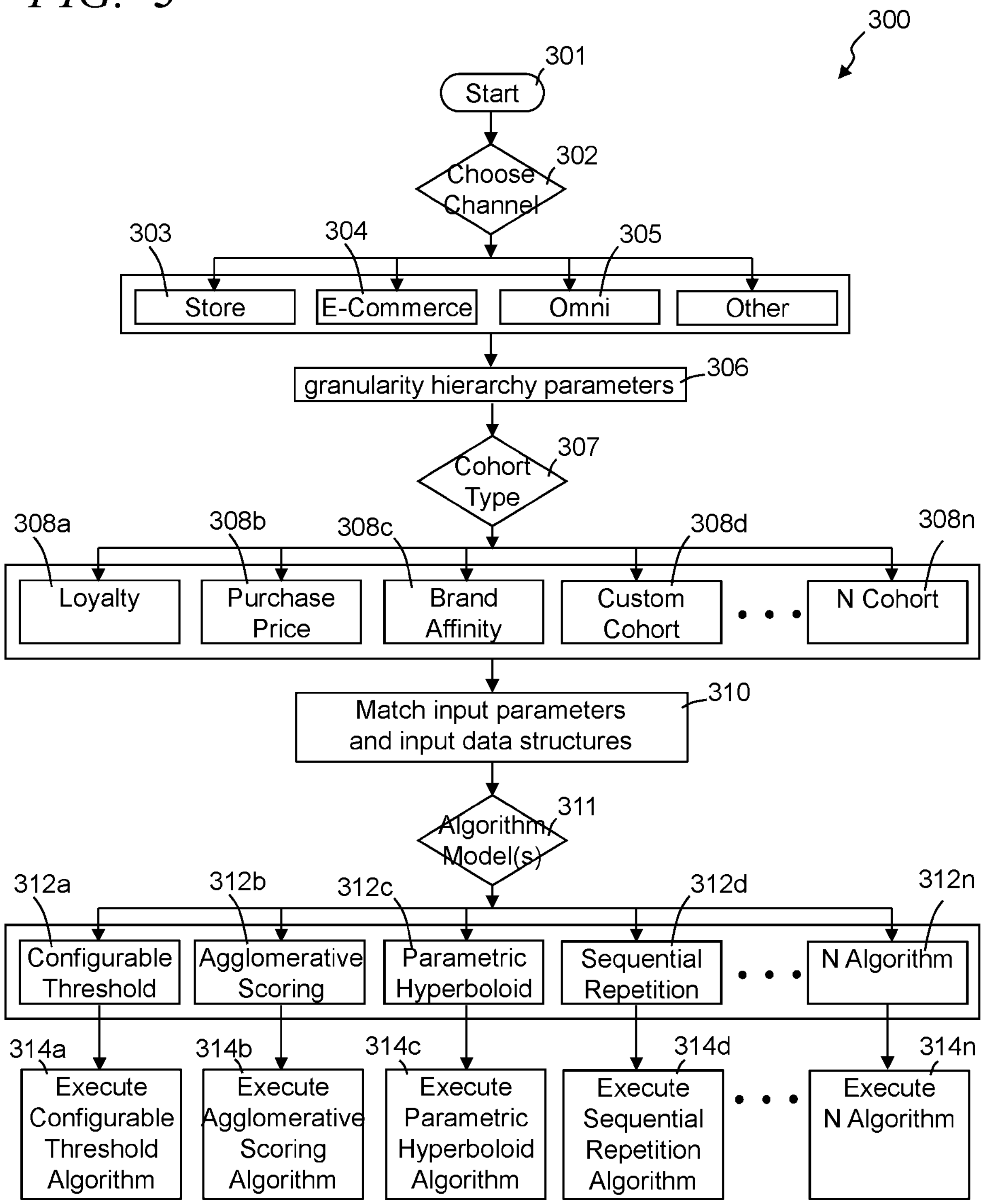

300
301
Start
302
Choose Channel
303
304
305
Store
E-Commerce
Omni
Other
granularity hierarchy parameters
306
307
Cohort Type
308a
308b
308c
308d
308n
Loyalty
Purchase Price
Brand Affinity
Custom Cohort
N Cohort
Match input parameters and input data structures
310
311
Algorithm/ Model(s)
312a
312b
312c
312d
312n
Configurable Threshold
Agglomerative Scoring
Parametric Hyperboloid
Sequential Repetition
N Algorithm
314a
314b
314c
314d
314n
Execute Configurable Threshold Algorithm
Execute Agglomerative Scoring Algorithm
Execute Parametric Hyperboloid Algorithm
Execute Sequential Repetition Algorithm
Execute N Algorithm

*FIG. 4A*

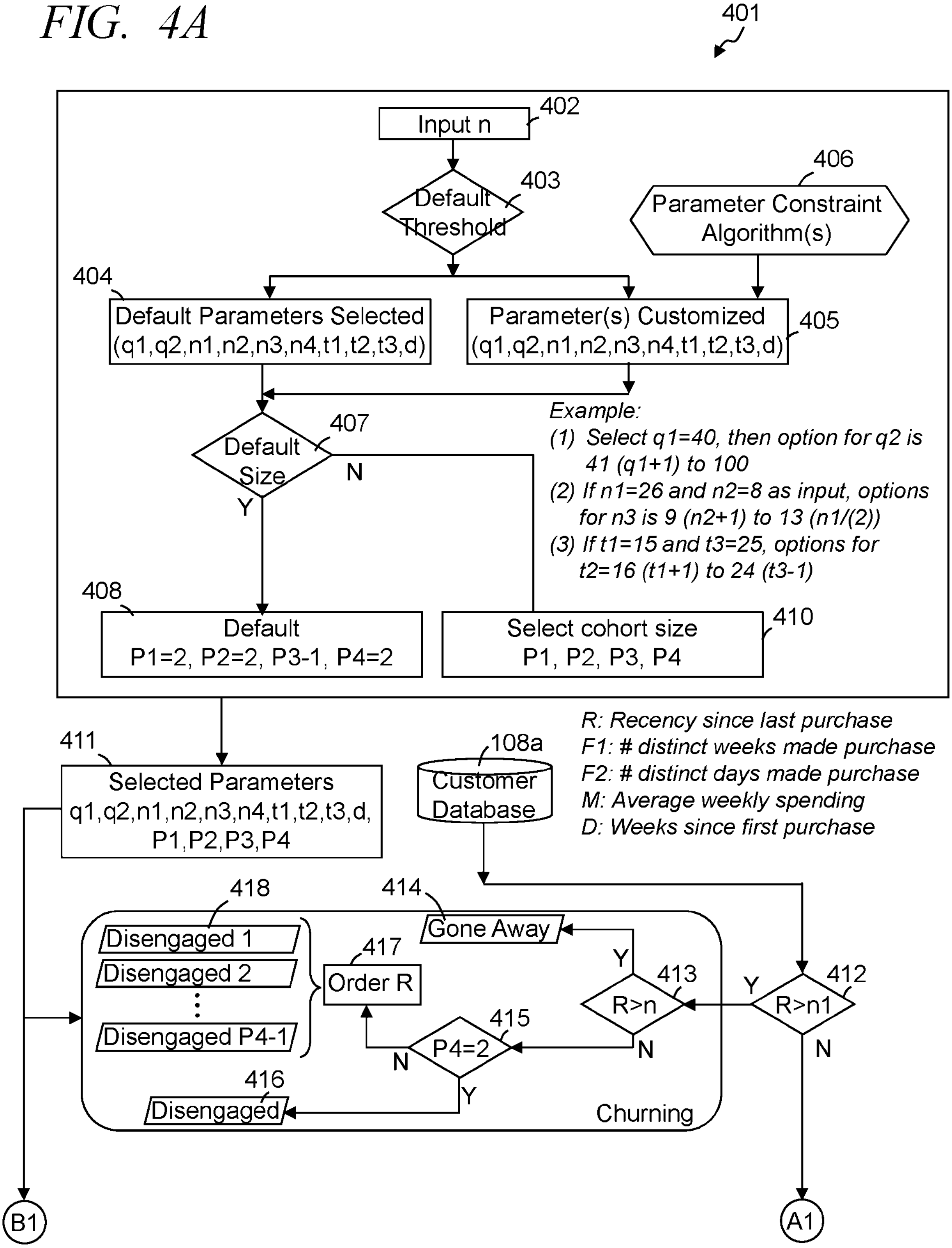

401
Input n
402
Default Threshold
403
Parameter Constraint Algorithm(s)
406
404
Default Parameters Selected (q1,q2,n1,n2,n3,n4,t1,t2,t3,d)
Parameter(s) Customized (q1,q2,n1,n2,n3,n4,t1,t2,t3,d)
405
Default Size
407
N
Y
Example:
(1) Select q1=40, then option for q2 is 41 (q1+1) to 100
(2) If n1=26 and n2=8 as input, options for n3 is 9 (n2+1) to 13 (n1/(2))
(3) If t1=15 and t3=25, options for t2=16 (t1+1) to 24 (t3-1)
408
Default P1=2, P2=2, P3-1, P4=2
Select cohort size P1, P2, P3, P4
410
R: Recency since last purchase
F1: # distinct weeks made purchase
F2: # distinct days made purchase
M: Average weekly spending
D: Weeks since first purchase
411
Selected Parameters q1,q2,n1,n2,n3,n4,t1,t2,t3,d, P1,P2,P3,P4
108a
Customer Database
418
414
Disengaged 1
Disengaged 2
Disengaged P4-1
Gone Away
417
Order R
Y
413
R>n
Y
412
R>n1
415
P4=2
N
N
N
416
Disengaged
Y
Churning
B1
A1

FIG. 4B

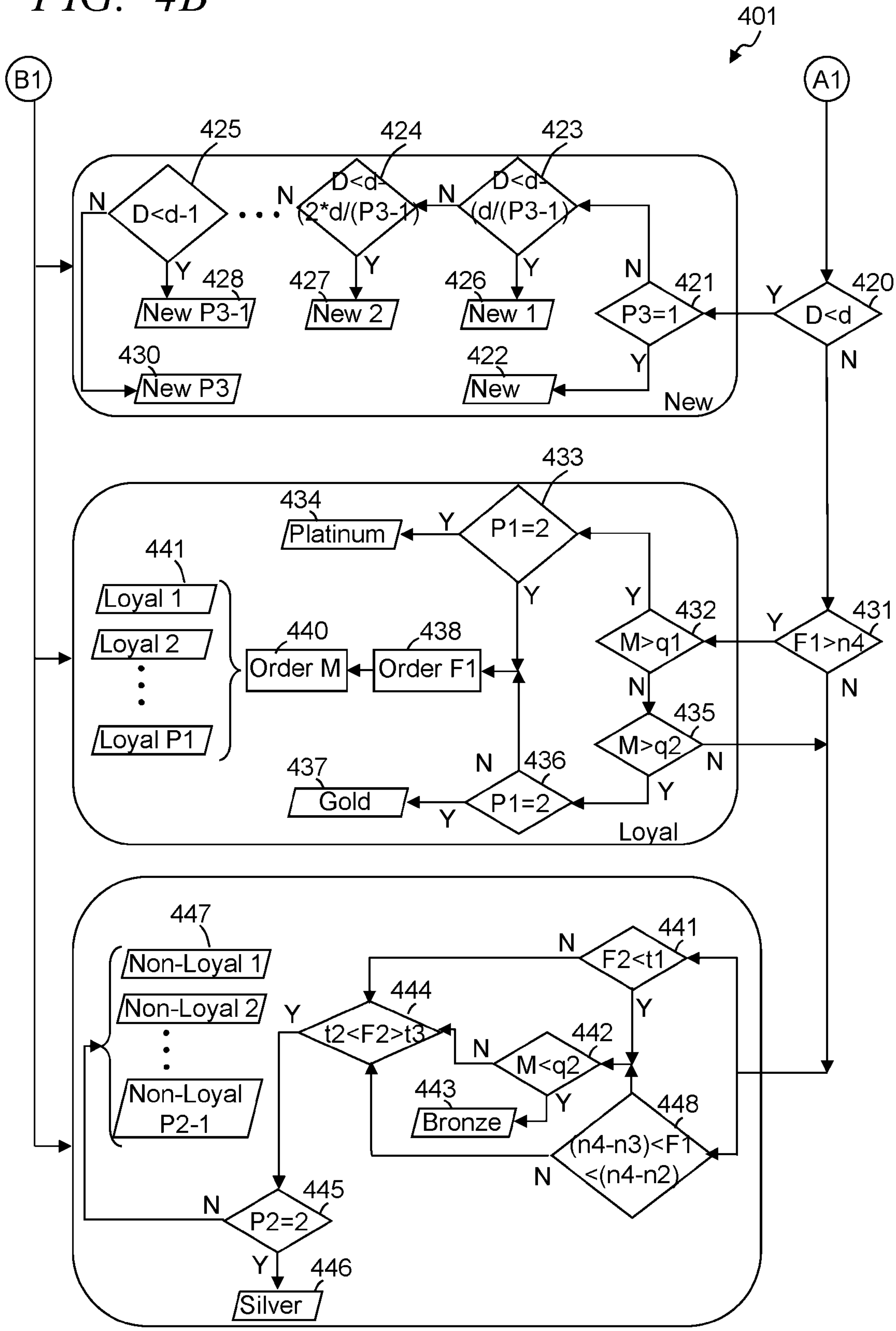
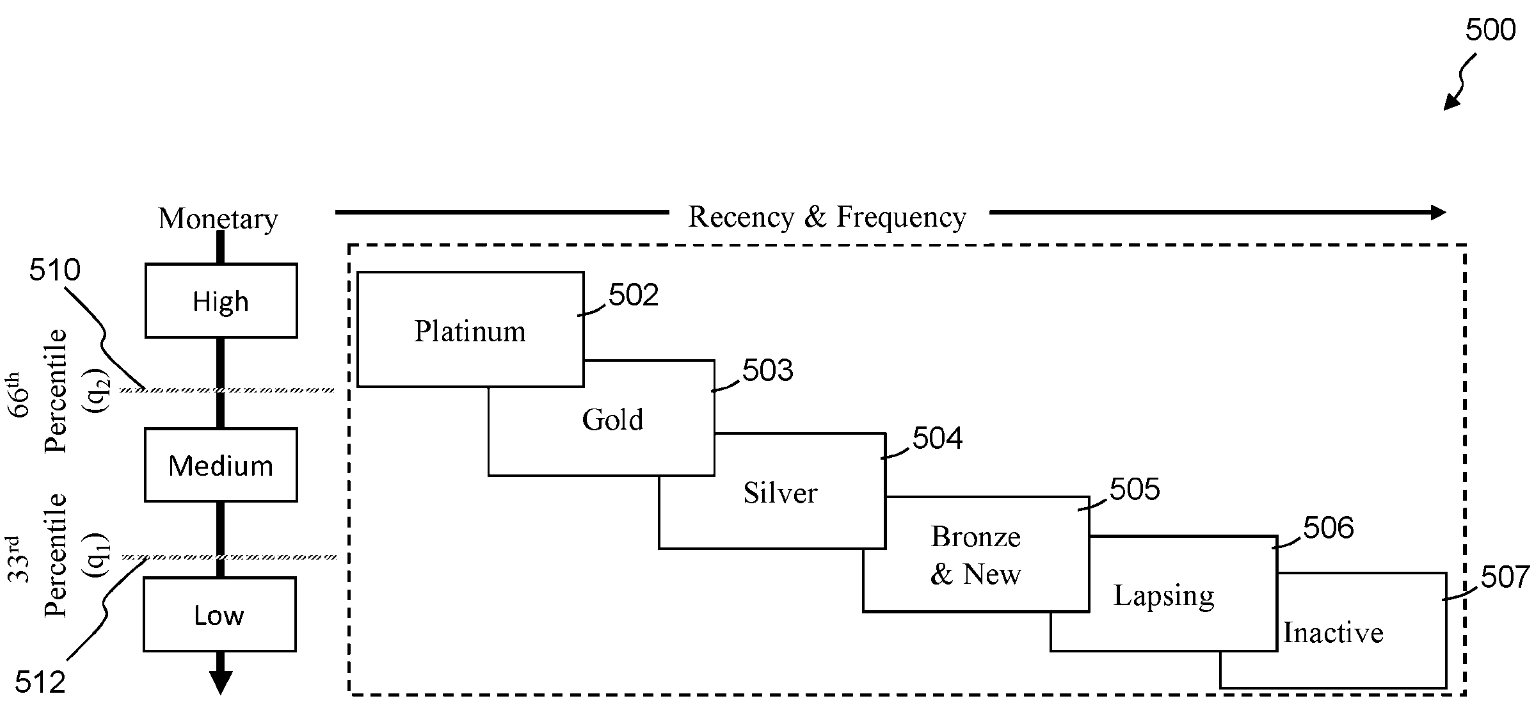

401
B1
A1
425
424
423
D<d-1
D<d-(2*d/(P3-1))
D<d-(d/(P3-1))
N
Y
428
427
426
421
420
New P3-1
New 2
New 1
P3=1
D<d
430
422
New P3
New
New
433
434
Platinum
P1=2
441
Loyal 1
Loyal 2
Loyal P1
440
438
Order M
Order F1
432
431
M>q1
F1>n4
435
M>q2
437
436
Gold
P1=2
Loyal
447
Non-Loyal 1
Non-Loyal 2
Non-Loyal P2-1
441
F2<t1
444
t2<F2>t3
442
M<q2
443
Bronze
448
(n4-n3)<F1<(n4-n2)
445
P2=2
446
Silver

FIG. 5




- *The threshold values (e.g., $q_1$, $q_2$, n, $n_1$, $n_2$, $n_3$, $n_4$, $t_1$, $t_2$, $t_3$) and/or configurable by user with feasibility restrictions put in place; and/or*
- *One or more threshold values (e.g., $q_1$, $q_2$, n, $n_1$, $n_2$, $n_3$, $n_4$, $t_1$, $t_2$, $t_3$) can be automated with the input of n or default n*

*FIG. 6A*

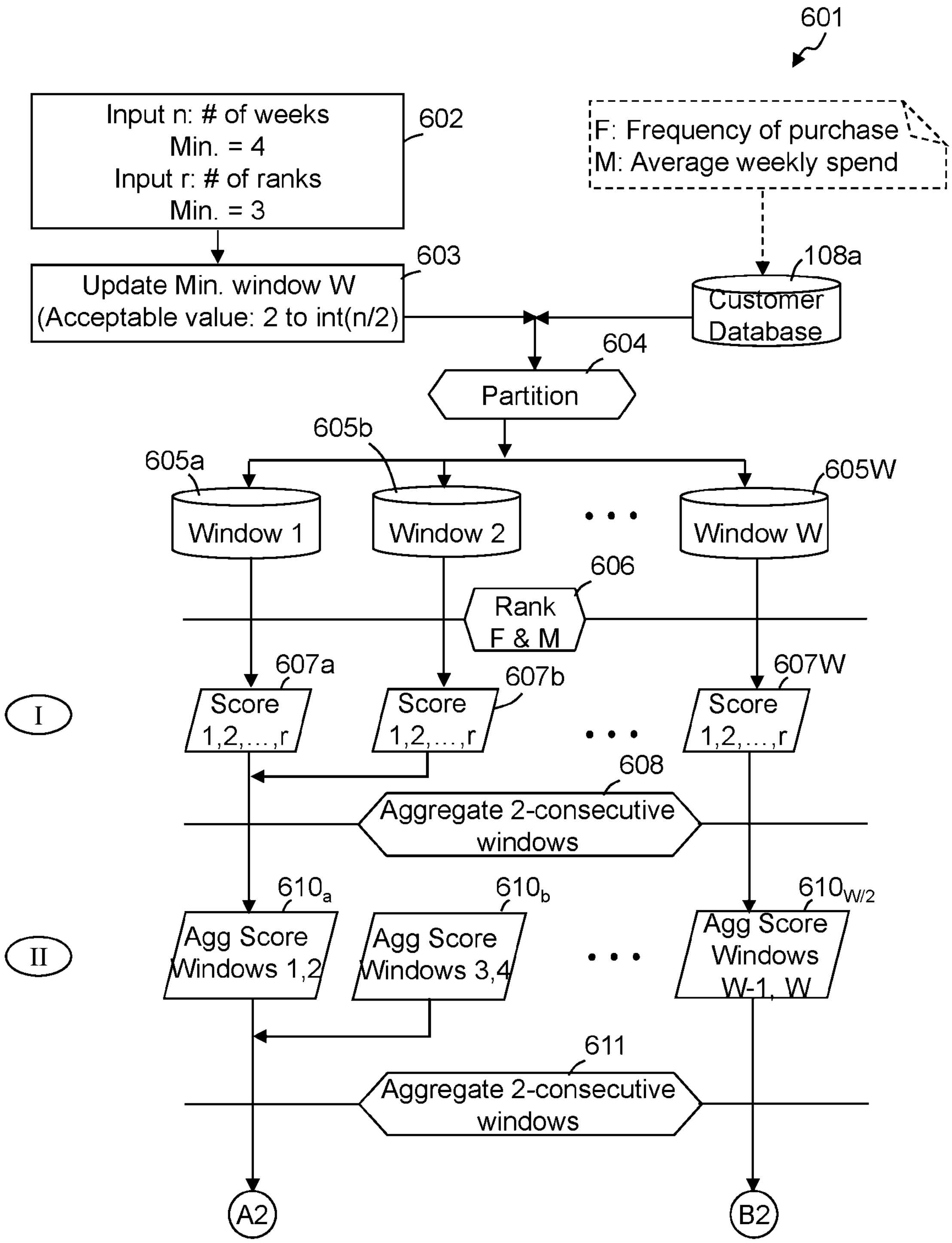

601
Input n: # of weeks
Min. = 4
Input r: # of ranks
Min. = 3
602
F: Frequency of purchase
M: Average weekly spend
603
Update Min. window W
(Acceptable value: 2 to int(n/2))
108a
Customer Database
604
Partition
605a
605b
605W
Window 1
Window 2
Window W
606
Rank
F & M
607a
607b
607W
I
Score
1,2,...,r
608
Aggregate 2-consecutive windows
610a
610b
610W/2
II
Agg Score Windows 1,2
Agg Score Windows 3,4
Agg Score Windows W-1, W
611
Aggregate 2-consecutive windows
A2
B2

FIG. 6B

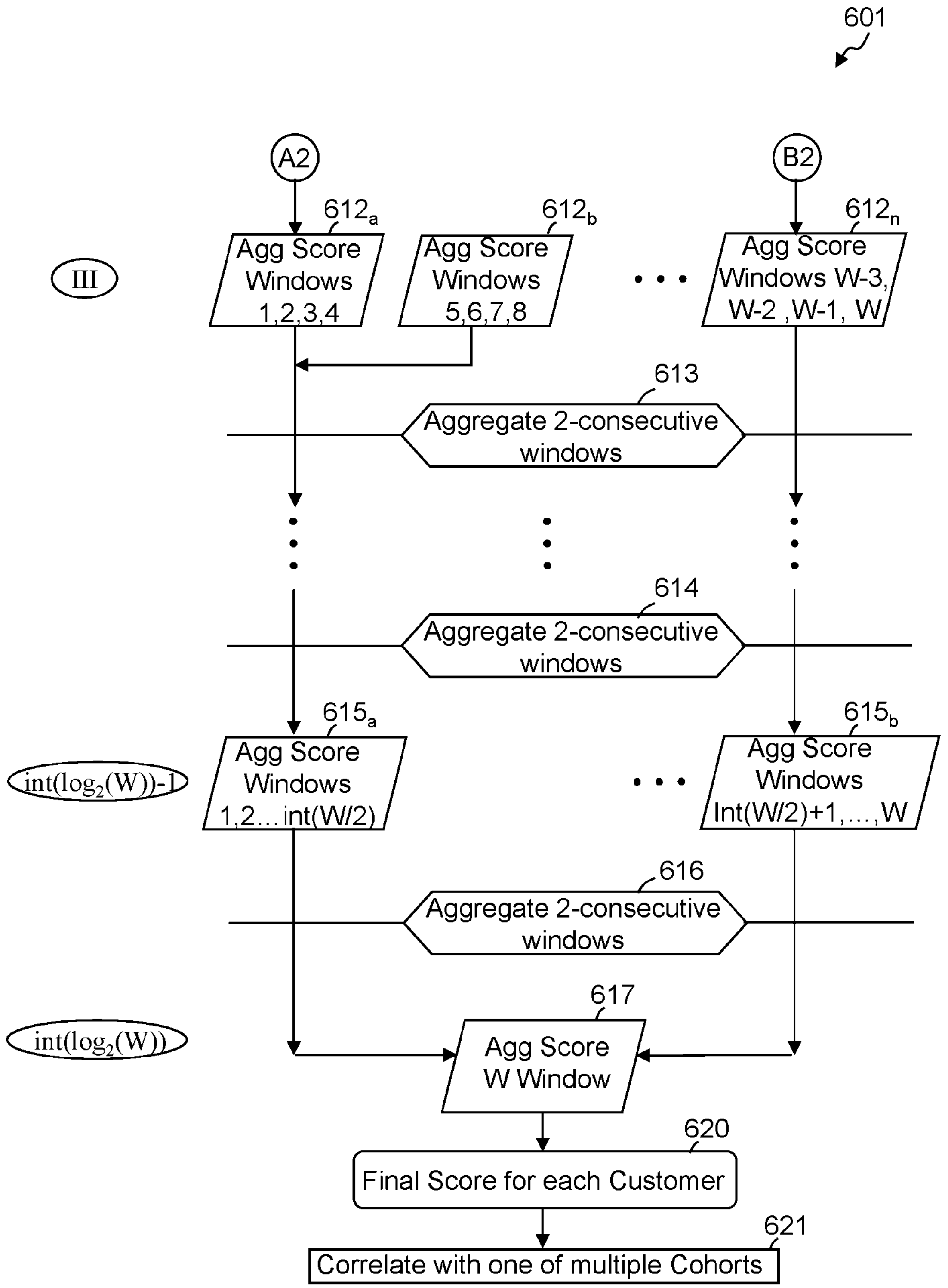

601
A2
B2
612a
612b
612n
III
Agg Score Windows 1,2,3,4
Agg Score Windows 5,6,7,8
Agg Score Windows W-3, W-2 ,W-1, W
613
Aggregate 2-consecutive windows
614
Aggregate 2-consecutive windows
615a
615b
int(log2(W))-1
Agg Score Windows 1,2…int(W/2)
Agg Score Windows Int(W/2)+1,…,W
616
Aggregate 2-consecutive windows
617
int(log2(W))
Agg Score W Window
620
Final Score for each Customer
621
Correlate with one of multiple Cohorts

*FIG. 7A*

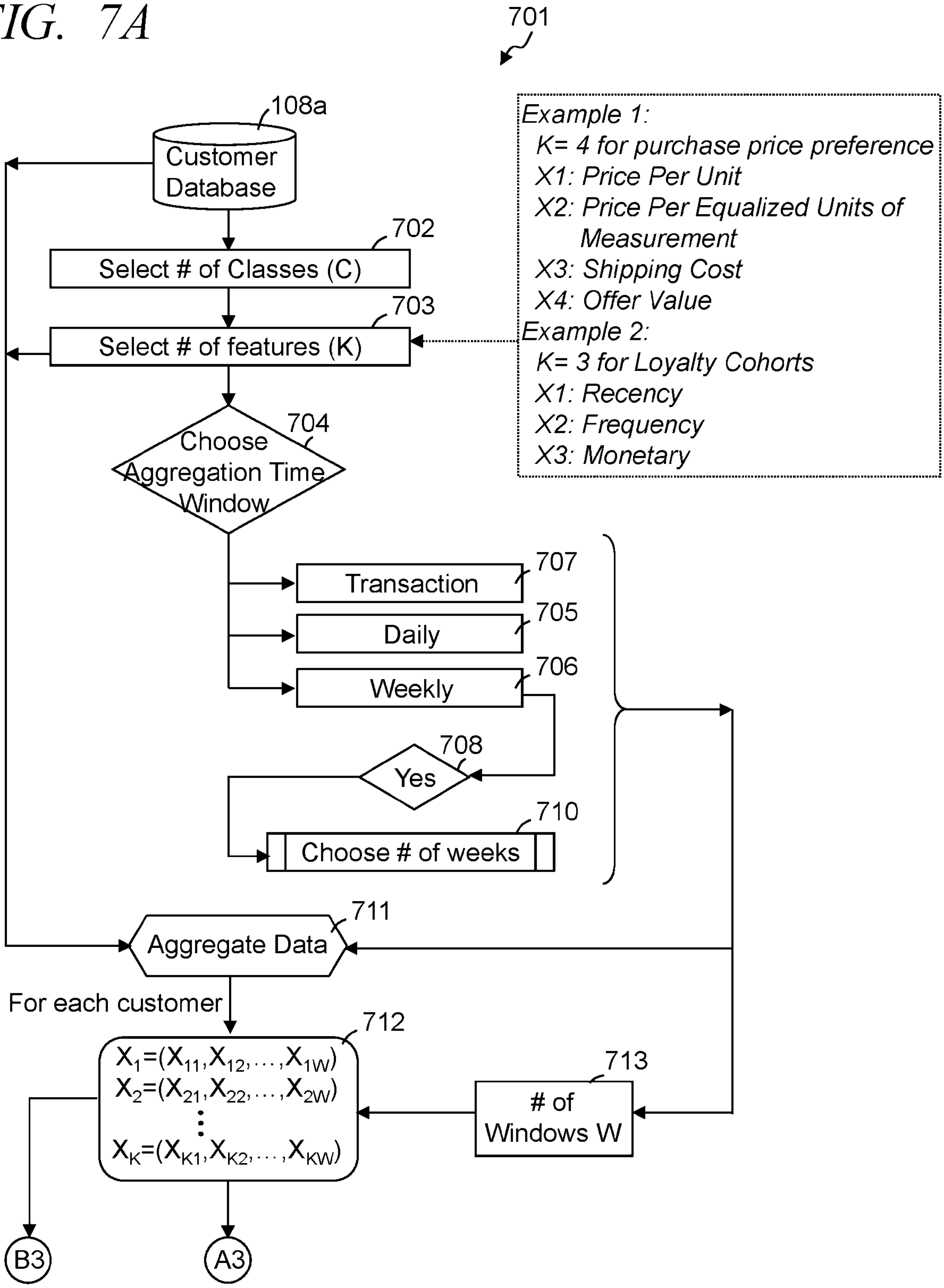
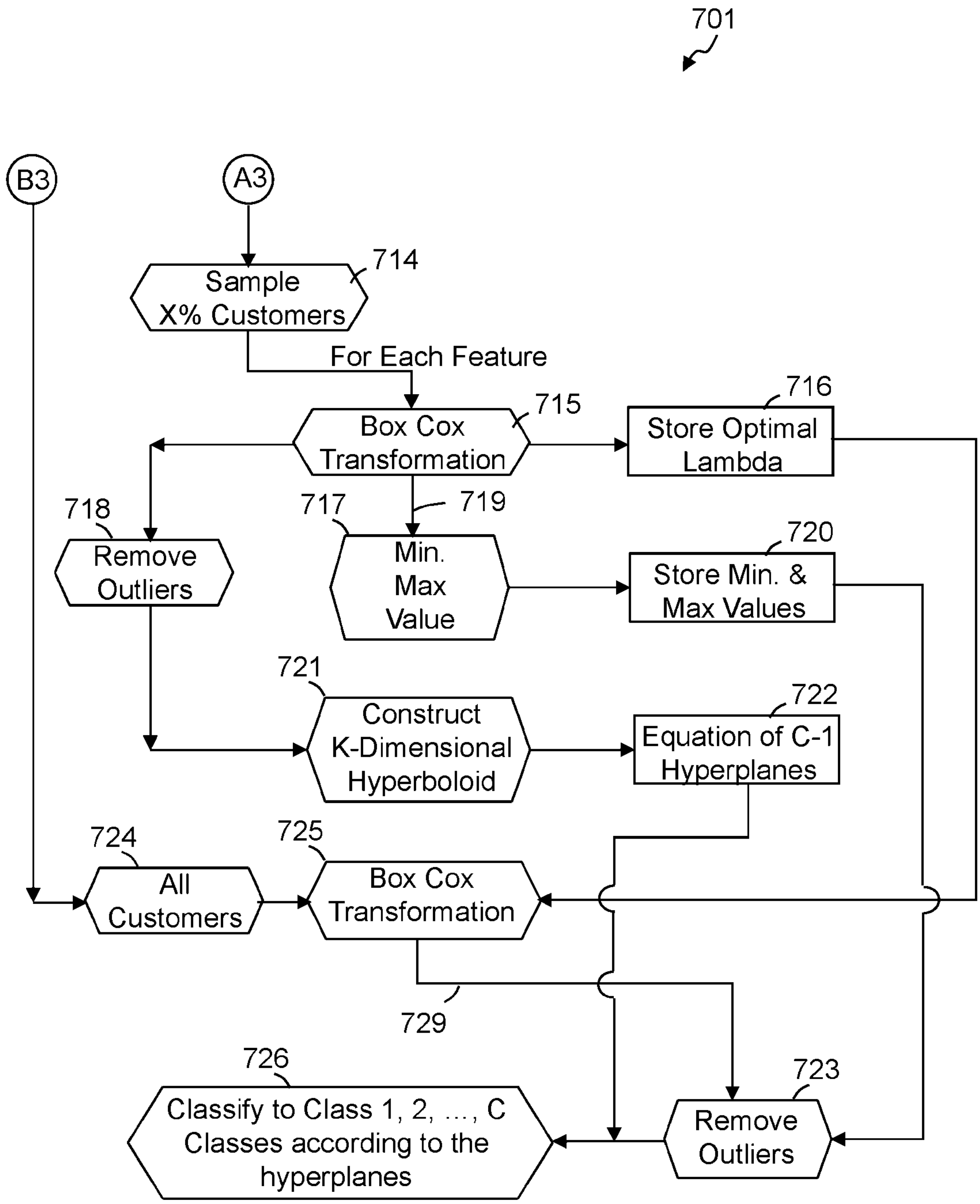

701
108a
Customer Database
702
Select # of Classes (C)
703
Select # of features (K)
Example 1:
K= 4 for purchase price preference
X1: Price Per Unit
X2: Price Per Equalized Units of Measurement
X3: Shipping Cost
X4: Offer Value
Example 2:
K= 3 for Loyalty Cohorts
X1: Recency
X2: Frequency
X3: Monetary
704
Choose Aggregation Time Window
Transaction
707
Daily
705
Weekly
706
708
Yes
710
Choose # of weeks
711
Aggregate Data
For each customer
712
$X_1=(X_{11},X_{12},\ldots,X_{1W})$
$X_2=(X_{21},X_{22},\ldots,X_{2W})$
$X_K=(X_{K1},X_{K2},\ldots,X_{KW})$
713
# of Windows W
B3
A3

*FIG. 7B*


701
B3
A3
Sample X% Customers
714
For Each Feature
Box Cox Transformation
715
Store Optimal Lambda
716
718
Remove Outliers
717
719
Min. Max Value
720
Store Min. & Max Values
721
Construct K-Dimensional Hyperboloid
722
Equation of C-1 Hyperplanes
724
All Customers
725
Box Cox Transformation
729
726
Classify to Class 1, 2, ..., C Classes according to the hyperplanes
723
Remove Outliers

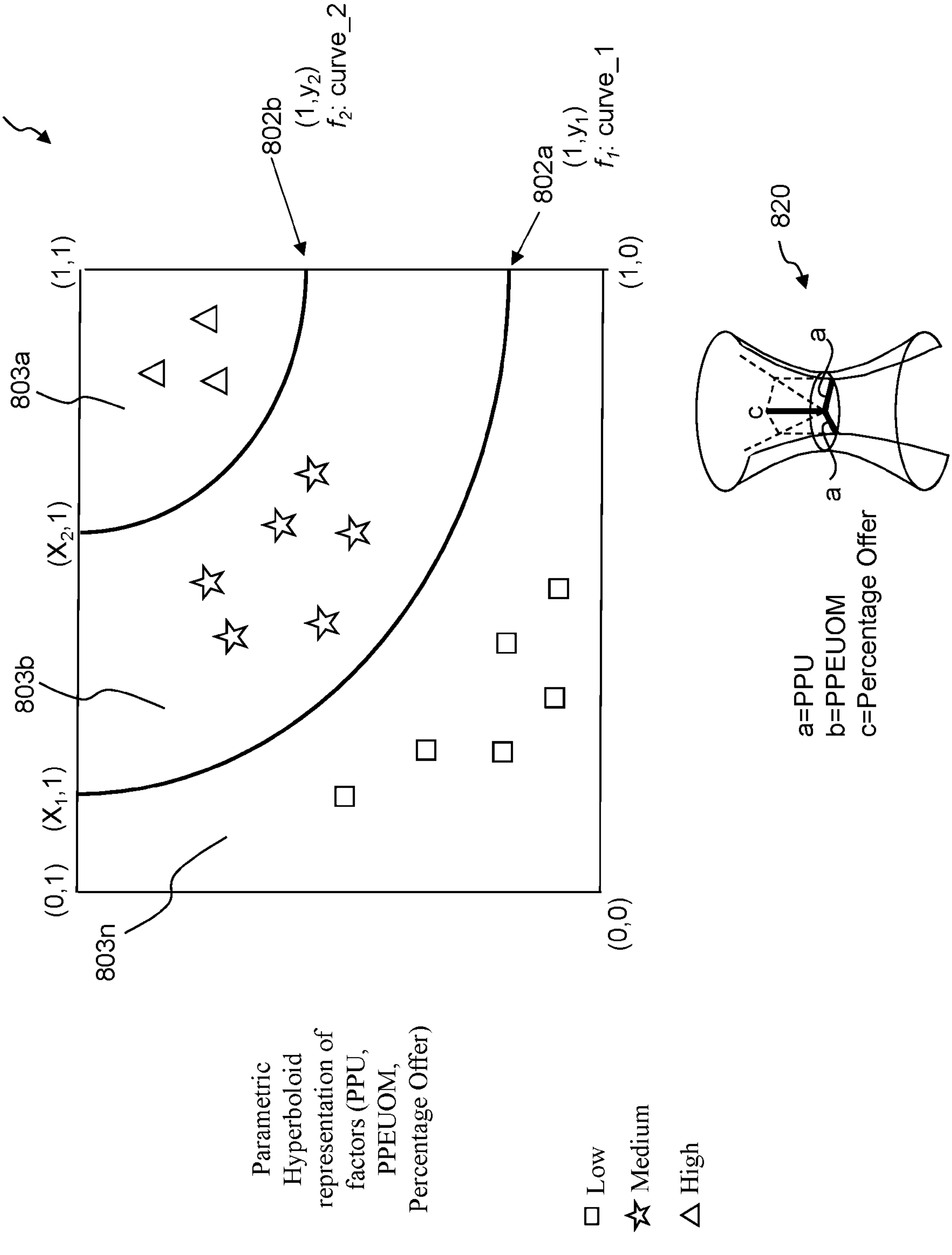
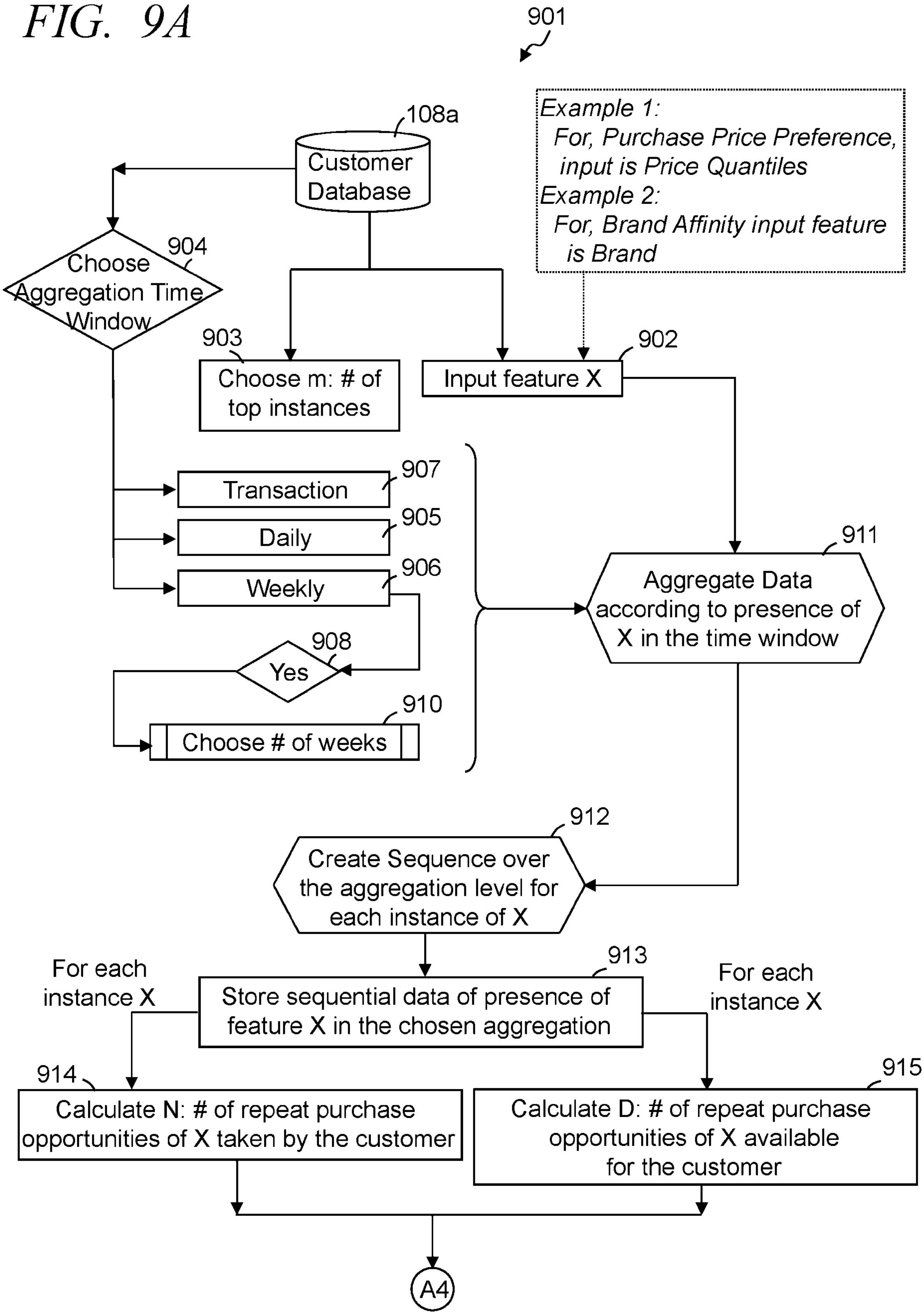

801
802b
(1,y2)
f2: curve_2
802a
(1,y1)
f1: curve_1
803a
803b
803n
(1,1)
(X2,1)
(X1,1)
(0,1)
(0,0)
(1,0)
820
a
c
a
a=PPU
b=PPEUOM
c=Percentage Offer
Parametric Hyperboloid representation of factors (PPU, PPEUOM, Percentage Offer)
Low
Medium
High

*FIG. 8*

FIG. 9A



901
108a
Customer Database
Example 1:
For, Purchase Price Preference, input is Price Quantiles
Example 2:
For, Brand Affinity input feature is Brand
904
Choose Aggregation Time Window
903
Choose m: # of top instances
902
Input feature X
907
Transaction
905
Daily
906
Weekly
908
Yes
910
Choose # of weeks
911
Aggregate Data according to presence of X in the time window
912
Create Sequence over the aggregation level for each instance of X
913
Store sequential data of presence of feature X in the chosen aggregation
For each instance X
For each instance X
914
Calculate N: # of repeat purchase opportunities of X taken by the customer
915
Calculate D: # of repeat purchase opportunities of X available for the customer
A4

*FIG. 9B*

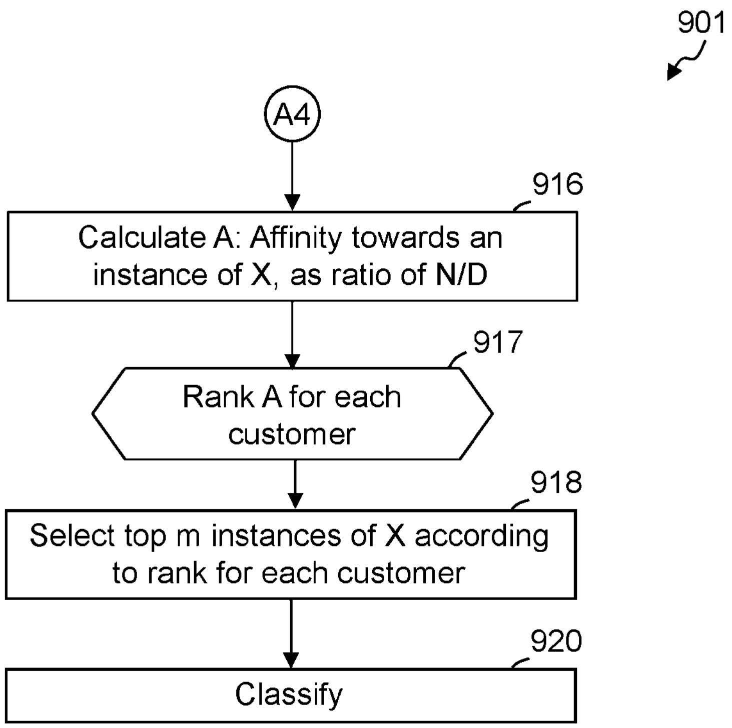

901
A4
916
Calculate A: Affinity towards an instance of X, as ratio of N/D
917
Rank A for each customer
918
Select top m instances of X according to rank for each customer
920
Classify

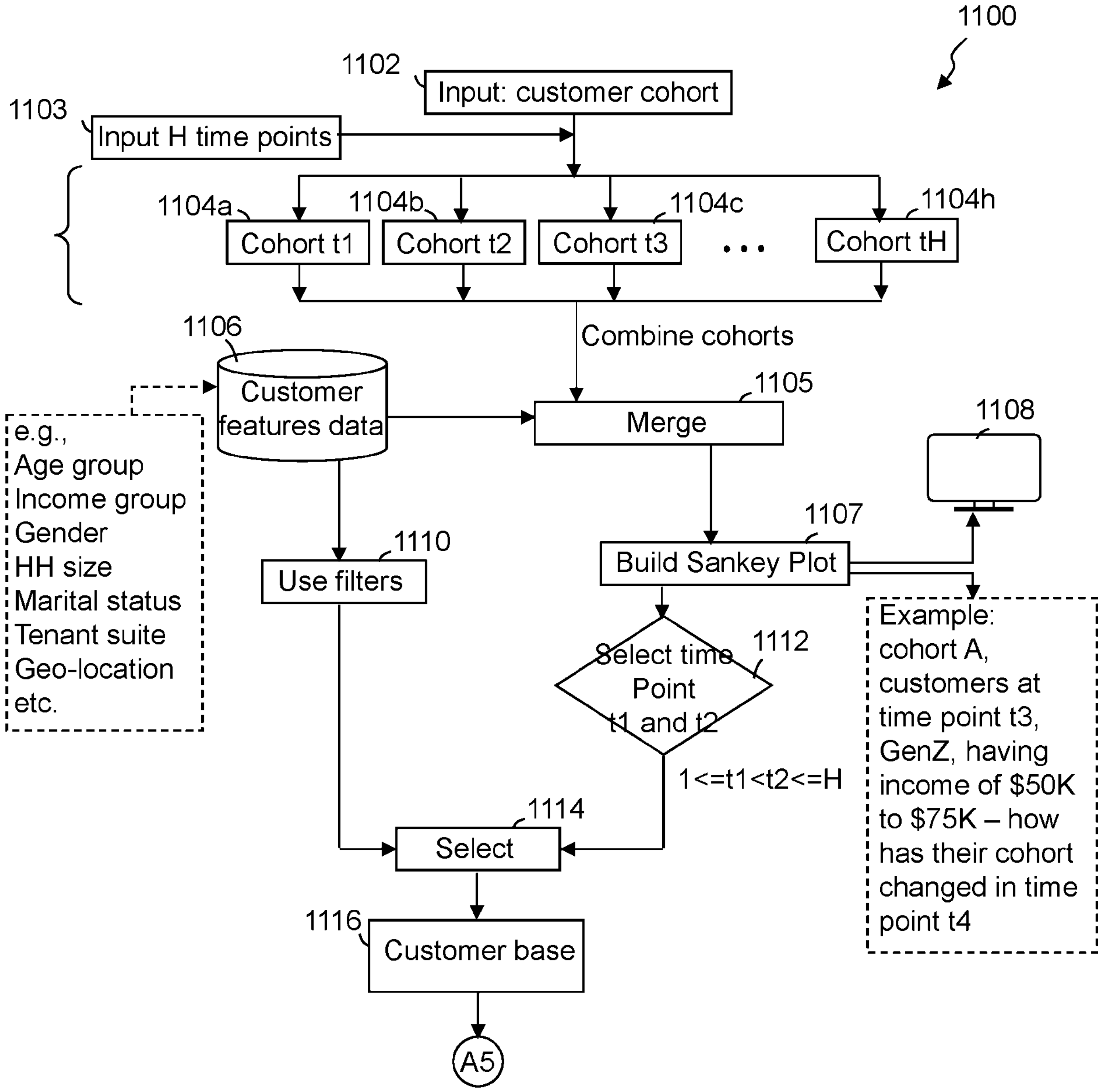
FIG. 10

1000

| Customer ID (1004) | Instance (1006) | Time Window (1008) | Loyalty Cohort (1010) | Next Loyalty Cohort Prediction | Price Preference (1011) | Brand Preference (1012) |
|---|---|---|---|---|---|---|
| 9124124 | July' 22 | Yearly | Platinum | … | High | {B1, B23, …} |
| 1415101 | July' 22 | Yearly | Bronze | … | High | {B12, B24} |
| … | … | … | … | … | … | … |
| 4151591 | July' 22 | Yearly | Gold | …. | Low | {} |
| … | … | … | … | … | … | … |
| 9124124 | Oct' 22 | Yearly | Platinum | Platinum | High | {B23, …} |
| 1415101 | Oct' 22 | Yearly | Silver | Gold | Medium | {B12, B24} |
| … | … | … | … | … | … | … |
| 4151591 | Oct' 22 | Yearly | Platinum | Gold | Low | {B14} |

1002, 1014, 1015, 1016, 1020, 1022

*FIG. 11*



1100
1102
Input: customer cohort
1103
Input H time points
1104a
1104b
1104c
1104h
Cohort t1
Cohort t2
Cohort t3
...
Cohort tH
Combine cohorts
1106
Customer features data
1105
Merge
1108
e.g.,
Age group
Income group
Gender
HH size
Marital status
Tenant suite
Geo-location
etc.
1107
Build Sankey Plot
1110
Use filters
1112
Select time Point t1 and t2
1<=t1<t2<=H
Example: cohort A, customers at time point t3, GenZ, having income of $50K to $75K – how has their cohort changed in time point t4
1114
Select
1116
Customer base
A5

FIG. 12

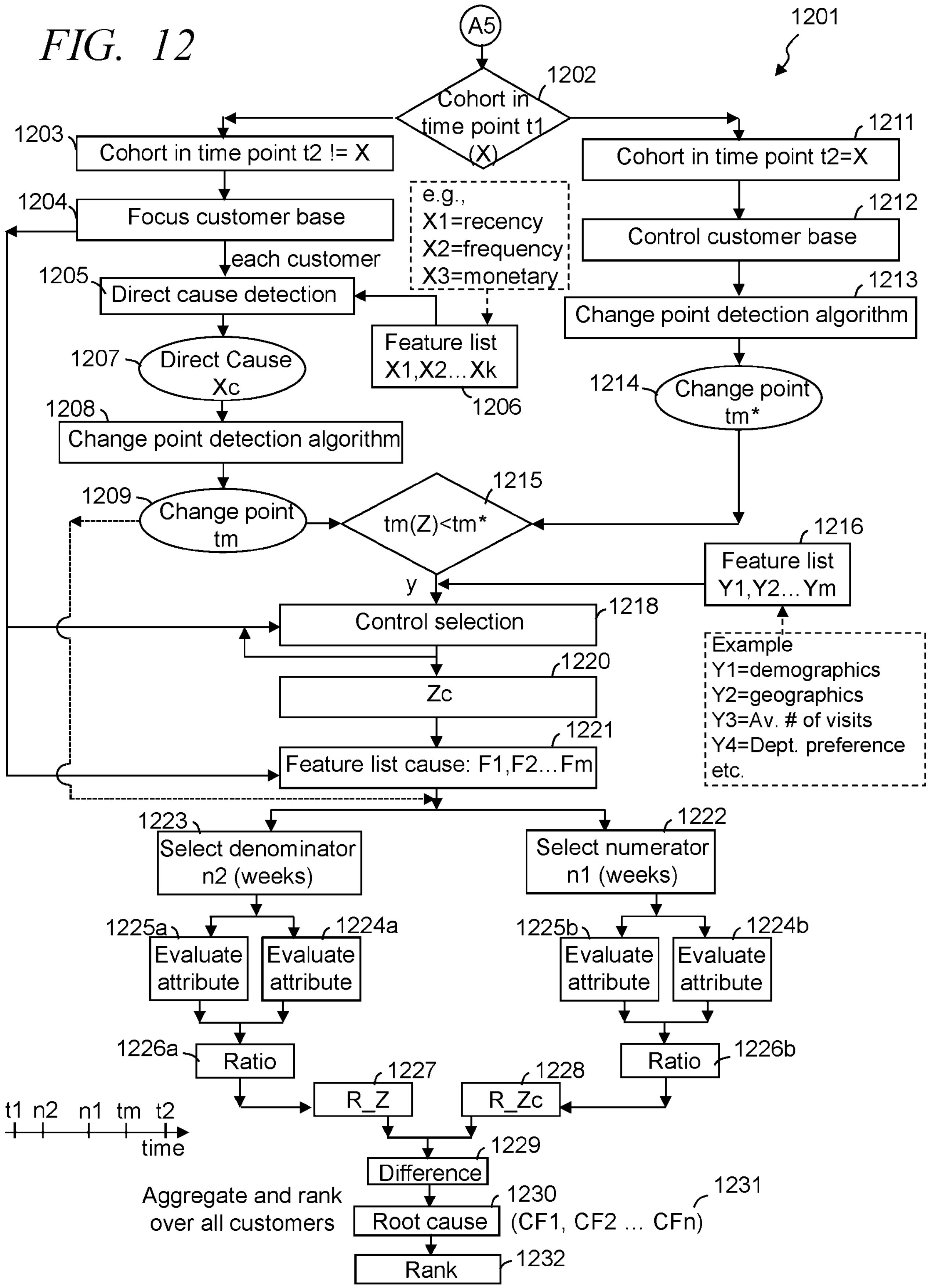
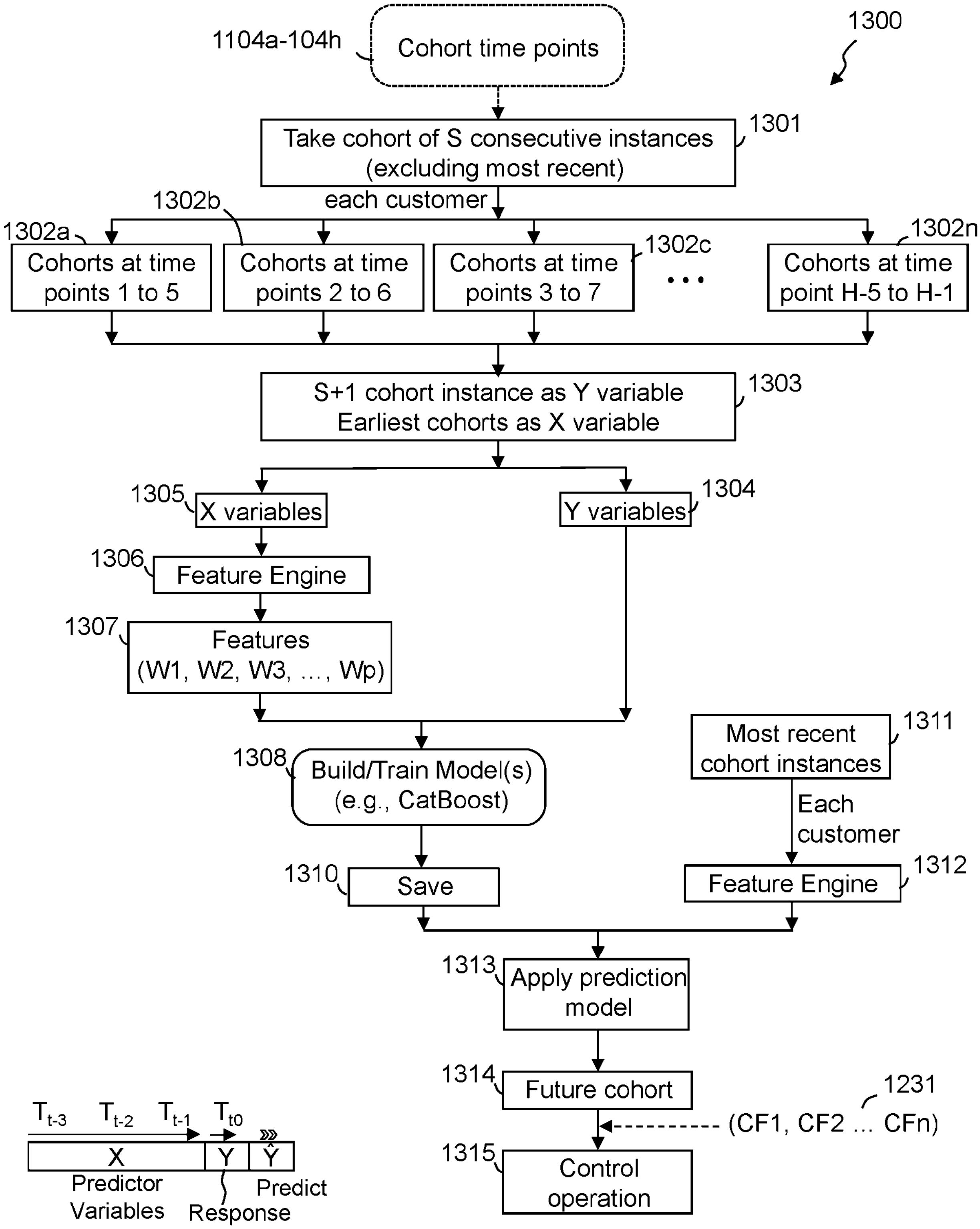

1201
A5
1202
Cohort in time point t1 (X)
1203
Cohort in time point t2 != X
1204
Focus customer base
each customer
1205
Direct cause detection
e.g.,
X1=recency
X2=frequency
X3=monetary
Feature list X1,X2…Xk
1206
1207
Direct Cause Xc
1208
Change point detection algorithm
1209
Change point tm
1211
Cohort in time point t2=X
1212
Control customer base
1213
Change point detection algorithm
1214
Change point tm*
1215
tm(Z)<tm*
y
1216
Feature list Y1,Y2…Ym
Example
Y1=demographics
Y2=geographics
Y3=Av. # of visits
Y4=Dept. preference
etc.
1218
Control selection
1220
Zc
1221
Feature list cause: F1,F2…Fm
1223
Select denominator n2 (weeks)
1222
Select numerator n1 (weeks)
1225a
Evaluate attribute
1224a
Evaluate attribute
1225b
Evaluate attribute
1224b
Evaluate attribute
1226a
Ratio
1226b
Ratio
1227
R_Z
1228
R_Zc
t1 n2 n1 tm t2
time
1229
Difference
Aggregate and rank over all customers
1230
Root cause
1231
(CF1, CF2 … CFn)
1232
Rank

*FIG. 13*



1300
1104a-104h
Cohort time points
Take cohort of S consecutive instances (excluding most recent)
1301
each customer
1302a
1302b
1302c
1302n
Cohorts at time points 1 to 5
Cohorts at time points 2 to 6
Cohorts at time points 3 to 7
Cohorts at time point H-5 to H-1
S+1 cohort instance as Y variable
Earliest cohorts as X variable
1303
1305
X variables
Y variables
1304
1306
Feature Engine
1307
Features (W1, W2, W3, …, Wp)
1308
Build/Train Model(s) (e.g., CatBoost)
1311
Most recent cohort instances
Each customer
1310
Save
Feature Engine
1312
1313
Apply prediction model
1314
Future cohort
1231
(CF1, CF2 … CFn)
1315
Control operation
T t-3
T t-2
T t-1
T t0
X
Y
Ŷ
Predictor Variables
Response
Predict

FIG. 14

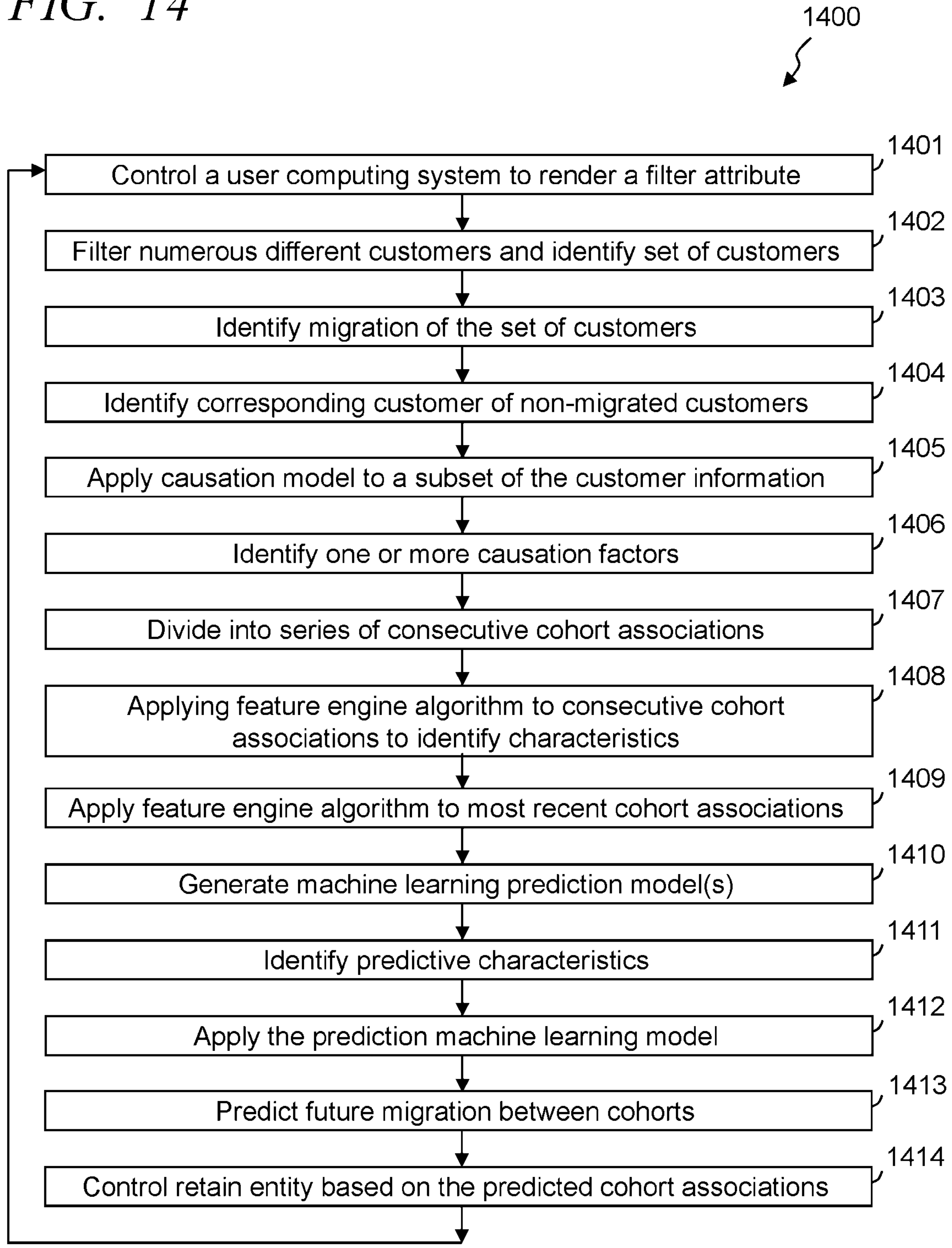
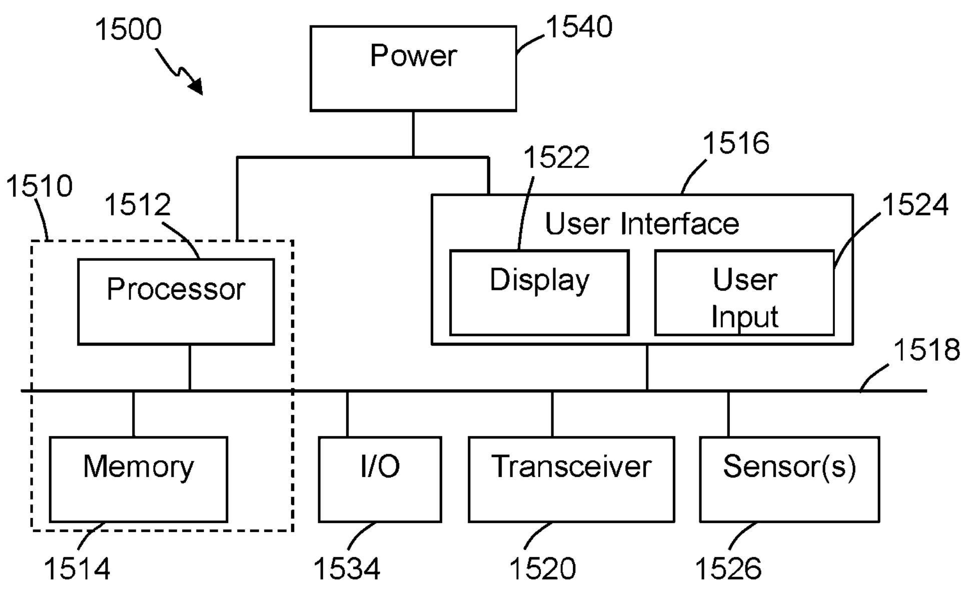

1400
1401
Control a user computing system to render a filter attribute
1402
Filter numerous different customers and identify set of customers
1403
Identify migration of the set of customers
1404
Identify corresponding customer of non-migrated customers
1405
Apply causation model to a subset of the customer information
1406
Identify one or more causation factors
1407
Divide into series of consecutive cohort associations
1408
Applying feature engine algorithm to consecutive cohort associations to identify characteristics
1409
Apply feature engine algorithm to most recent cohort associations
1410
Generate machine learning prediction model(s)
1411
Identify predictive characteristics
1412
Apply the prediction machine learning model
1413
Predict future migration between cohorts
1414
Control retain entity based on the predicted cohort associations

FIG. 15


1500
Power
1540
1522
1516
1510
1512
User Interface
1524
Processor
Display
User Input
1518
Memory
I/O
Transceiver
Sensor(s)
1514
1534
1520
1526