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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 March 2022 (17.03.2022)

WIPO | PCT

(10) International Publication Number
WO 2022/055615 A1

(51) International Patent Classification:
G06Q 30/02 (2012.01)

(21) International Application Number:
PCT/US2021/041773

(22) International Filing Date:
15 July 2021 (15.07.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
17/016,825 10 September 2020 (10.09.2020) US

(71) Applicant: **ORACLE INTERNATIONAL CORPORATION** [US/US]; 500 Oracle Parkway, Mail Stop 5OP7, Redwood Shores, California 94065 (US).

(72) Inventors: **LEI, Ming**; 10324 Royal Ascot Ct., Ellicott City, 21042 (US). **POPESCU, Catalin**; 3186 Alton Road, Atlanta, Georgia 30341 (US).

(74) Agent: **GOLDSMITH, Barry**; Potomac Law Group, PLLC, 8229 Boone Boulevard, Suite 430, Vienna, Virginia 22812 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

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

(54) Title: SHORT LIFE CYCLE SALES CURVE ESTIMATION

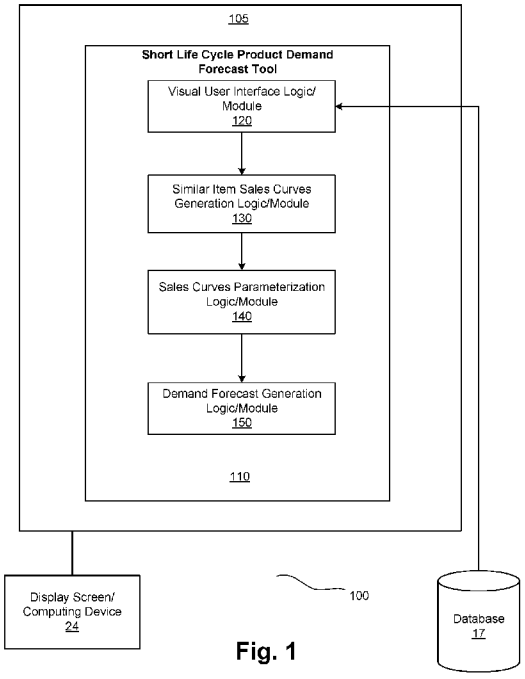


Fig. 1

(57) Abstract: Embodiments generate a short life cycle sales curve for a short life cycle item. Embodiments generate a plurality of similar sales curves corresponding to at least one similar item that is similar to the short life cycle item. Embodiments parametrize each of the similar sales curves, including estimating, for each similar sales curve, a coefficient of innovation parameter, a coefficient of imitation parameter, and an error parameter, and determine a weight for each error parameter. Embodiments combine the coefficient of innovation parameters and the coefficient of imitation parameters using the weights, and generate the short life cycle sales curve using the combined coefficient of innovation parameters and the combined coefficient of imitation parameters.

WO 2022/055615 A1

SHORT LIFE CYCLE SALES CURVE ESTIMATION

FIELD

[0001] One embodiment is directed generally to a computer system, and in particular to a computer system that forecasts demand for retail items.

BACKGROUND INFORMATION

[0002] In general, sales forecast systems encounter problems in producing a week-by-week forecast of sales units for retail items, referred to as a "sales curve" or "demand curve." The sales of retail items in a given week is affected by many factors, such as seasonal factors, whether a discount has been applied to a retail item during the week, and at what point in the lifecycle of a merchandise the week falls. One common approach to forecasting weekly sales units involves building a "causal demand model" for retail items. This demand model is a mathematical model that describes weekly sales units in terms of factors such as the ones listed above. The factors are known as the "demand variables" or "demand features" that form a demand model.

[0003] The demand model specifies mathematically how the demand variables affect sales units. For example, if the amount of discount is a demand variable, historical data may show that a 50% price cut resulted in a 4-fold increase in sales units (i.e., related to price elasticity). In this example, the demand variable is a 50% price cut and the historical sales data is the 4-fold increase in sales. In order for the causal demand model to be of use in forecasting sales units, it is necessary to determine the relationship of the demand variable (50% price cut) to the sales units (4-fold increase). This relationship is referred to as the "demand parameter" associated with the demand variable.

[0004] In this example, the demand parameter may be determined to specify that for every 25% price reduction, sales of a particular retail item will increase by 2-fold. With the demand parameter determined, it is then possible to forecast sales units by specifying the future values of the demand variables. To continue the price cut example, the retailer might know that next season it will be running a 40% price cut during some weeks. The demand model will then forecast sales units for those

weeks accounting for the 40% price cut.

[0005] The demand parameter is determined by examining historical retail sales data (known as “retail panel data”) containing price cuts for the retail item itself, or for similar retail items. However, as noted above, several demand variables affect the sales of retail items. These several demand variables apply simultaneously. For example, a retailer may have performed the 50% price cut during the summer for a summer item, in which case the 4-fold increase in sales may be partially due to an increase in seasonal demand for summer retail items during summer. To separate the effects of the several demand variables on sales, a regression is performed on the demand model to determine values for demand parameters that cause the demand model to best fit retail panel data.

[0006] Further, some merchandise such as fashion products have very short life-cycles and the available prior sales data usually are very sparse. Estimating the sales curve of such short life cycle products is very challenging because of the short sales history and sparse sales cycle.

SUMMARY

[0007] Embodiments generate a short life cycle sales curve for a short life cycle item. Embodiments generate a plurality of similar sales curves corresponding to at least one similar item that is similar to the short life cycle item. Embodiments parametrize each of the similar sales curves, including estimating, for each similar sales curve, a coefficient of innovation parameter, a coefficient of imitation parameter, and an error parameter, and determine a weight for each error parameter. Embodiments combine the coefficient of innovation parameters and the coefficient of imitation parameters using the weights, and generate the short life cycle sales curve using the combined coefficient of innovation parameters and the combined coefficient of imitation parameters.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Fig. 1 illustrates a computer system having a computing device configured with a short life cycle product demand forecast tool in accordance to

embodiments.

[0009] Fig. 2 is a block diagram of a computer server/system in accordance with an embodiment of the present invention.

[0010] Fig. 3 is a flow diagram of the functionality of short life cycle product demand tool of Fig. 1 when estimating promotional effects that can be used for a demand forecast in accordance with one embodiment.

[0011] Figs. 4A-B and 5 illustrate an example of an implementation of embodiments of the invention, and how each example corresponds to the functionality of Fig. 3.

[0012] Fig. 6 is a flow diagram of the functionality of the short life cycle product demand forecast tool of Fig. 1 when generating a sales/demand curve for the short life cycle item (i.e., a given item/location or SKU) in accordance with one embodiment.

[0013] Figs. 7 and 8 illustrate an example of an implementation of embodiments of the invention.

[0014] Fig. 9 illustrates an integrated manufacturing, inventory and logistics system that includes demand forecasting as disclosed herein in accordance with one embodiment.

DETAILED DESCRIPTION

[0015] Embodiments generate demand forecast for short life cycle products/items, such as fashion items or technology items. Embodiments generate sales curves at different intersection levels for products that are similar to the short life cycle product, using demand forecasting systems/methods. Embodiments then combine the sales curves at different intersections using parameterization of the sales curves and error and error estimations. The aggregate sales curve is then used as a demand forecast for the short life cycle product.

[0016] Short life cycle products are characterized by a demand that only occurs during a short time period after which they become obsolete, which leads to in some cases a very short demand time series sales curve. Examples of short life cycle products include technology products (e.g., computers, consumer electronics, video games) and fashion products (e.g., toys, apparel, text books). The period of

demand for short life cycle products can vary from a few years to a few weeks.

[0017] The dynamic of new products demand is generally characterized by a relatively slow growth in the introduction stage, followed by a stage of rapid growth, after which the demand is stabilized and the product enters in a stage of maturity. Finally, the demand declines and the product is then typically replaced by another product, which is likely fairly similar to the preceding product.

[0018] In many industries, particularly in the technology area, short life cycle products are increasingly common. Technology companies are motivated by a continuous introduction of new products as a consequence of highly competitive markets. The competitive advantage of a company is determined largely on its ability to manage frequent entries and exits of products.

[0019] In the fashion industry, fashion products are characterized by short product life cycles and high market success uncertainty. An unsuccessful product requires multiple price discounts to clear inventory. The typical life cycle of a fashion product may be as low as 10 weeks.

[0020] As discussed above, in the retail industry, retailers need to predict future demand to better manage their inventory or promotion/markdown planning. To accurately forecast demand, retailers consider all factors that could impact the demand such as promotions, price change, seasonality, weather and so on. Known solutions for retailers have used various algorithms to estimate the promotion or price effects, but these algorithms typically are not suitable for short life cycle products in which sparse historical data for a particular item may be available.

[0021] Specifically, when it comes to products that have a short lifecycle, there is insufficient data early in this process to determine how the future sales of the product will vary over time. In the past, retailers have tried to find the closest pre-existing item in their historical data as proxy for predicting in-season sales. However such methods may be error prone for two reasons. First, the customer response to a newly introduced item is unknown, and second, the newly introduced item interacts with the rest of the assortment, fundamentally altering the remainder of the lifecycle sales of all items.

[0022] Currently, the retail industry is witnessing a proliferation of short life cycle products, including apparel and fashion retailers, and high-end electronics

consumer product retailers, among others. New product designs are introduced in the market even as older versions or prior short life cycle products are cleared from the inventory or phased out. The time-span of the lifecycles themselves are getting shorter. For example, a fast-fashion product lifecycle can last no more than 12-13 weeks, leaving very little time for a retailer to adapt to changing customer preferences at various locations and over time.

[0023] More accurate short life cycle products demand forecasting allows retailers to better allocate and manage short life cycle items in an assortment. However, predicting sales for newly introduced products is a challenge due to zero sales history and different store locations launching at different dates. Typically, short life cycle products exhibit their characteristic sales curve of an initial slow increase in sales at from the point of introduction in the market, followed by increasing sales to reach a peak value, and then a gradual decline until either all inventory is sold out or the product is removed from the market.

[0024] The term "item" or "retail item", as used herein, refers to merchandise sold, purchased, and/or returned in a sales environment. The terms "particular item" and "single item" are used interchangeably herein and refer to a particular item type (e.g., to a particular type of cellular telephone such as an iPhone 8), not to a unit item.

[0025] The terms "period", "time period", "retail period", or "calendar period", as used herein, refer to a unit increment of time (e.g., a 7-day week) which sellers use to correlate seasonal periods from one year to the next in a calendar for the purposes of planning and forecasting. The terms may be used interchangeably herein.

[0026] The term "sales channel" or "location" or "retail location", as used herein, may refer to a physical store where an item is sold, or to an online store via which an item is sold.

[0027] The term "sales data", as used herein, refers to historical sales and promotion information that has been recorded for an item that has been sold in past retail periods (e.g., over 52 weeks of the past year). Sales data may include, for example, a number of units (or a monetary amount) of an item sold in each retail period, along with data characterizing one or more types of promotions for the item.

Sales data may be stored in a database, for example.

[0028] The terms "promotion" and "sales promotion" are used interchangeably herein and refer to a particular type of promotion for an item. Some examples of promotion components may include a price discount promotion component, a television advertisement component, a radio advertisement component, a newspaper advertisement component, an internet advertisement component, an email advertisement component, and an in-store advertisement component.

[0029] The term "promotion effect" refer to a numerical value that characterizes the effect (e.g., the effect on sales and profitability) of promoting an item. For example, an estimated promotion effect of 2.0 may indicate that a promotion, or combination or promotions, is estimated to result in twice as many sales (a 100% increase) for an item. Promotion effects (i.e., values) may be used in a demand forecast model to forecast a demand for an item. Promotion effects may also be used in a computerized inventory system to control various aspects of inventory for an item.

[0030] Embodiments, in general, utilize the following demand model or function for demand forecasting ("equation (1)"):

Demand = base demand * seasonality * promo effects (* additional features effects) (1)

Where "base demand" is the historical demand without taking account any effects or other factors, seasonality is the impact on demand based on the season (i.e., time of year), and promo effects are the effects on demand based on one or more promotions offered during a time period. Any or all additional features/variables that impact demand can be added to the model as appropriate. However, the number of features could exceed 100 in some situations.

[0031] Many demand models take into account additional effects, such as the weather. For example, if the current year's weather differs significantly from the previous year, and from two years ago, the forecast may need corrections. For example, if the hot weather during Summer is longer this year, the forecast for steaks and ice cream may need to be increased. Another additional effect can be the inventory. If a popular fashion is out of some sizes and/or colors, the forecast may need to be adjusted downward to account for the missing articles. Yet another effect can be store count. If the retailer plans to aggressively expand, and increase

the number of stores by 10% in the following year, the forecast needs to be adjusted accordingly. However, for purposes of embodiments of the invention, it is assumed that seasonality and promotion effects have the overwhelmingly largest impact on the sales forecast.

[0032] Fig. 1 illustrates a computer system 100 having a computing device 105 configured with short life cycle product demand forecast tool 110 in accordance to embodiments. In one embodiment, short life cycle product demand forecast tool 110 may be part of a larger computer application (e.g., a computerized inventory management and demand forecasting application), configured to forecast and manage sales, promotions, and inventory for retail items at various retail locations. Short life cycle product demand forecast tool 110 is configured to computerize the process of forecast demand for a short life cycle product, particularly when there is none or limited sales history for the product due to its short life cycle nature.

[0033] In one embodiment, system 100 is a computing/data processing system including an application or collection of distributed applications for enterprise organizations. The applications and computing system 100 may be configured to operate with or be implemented as a cloud-based networking system, a software-as-a-service ("SaaS") architecture, or other type of computing solution.

[0034] The demand forecast, in the form of a sales or demand curve, is an important driver of the supply chain. If a forecast is inaccurate, allocation and replenishment perform poorly, resulting in financial loss for the retailer. Improvements in forecast accuracy for promoted items may be achieved by the embodiments disclosed herein. Further, a better understanding of the impact a promotion has on demand may be achieved. This helps the retailer to plan promotions more effectively with respect to channel, pricing, and customer segments, for example.

[0035] Specifically, in order to plan future activities, many retailers rely on merchandise sales and lifecycle plans for different item categories. These plans typically consist of weekly values reflecting the selling pattern during the future lifecycle of a product. The plans are frequently viewed as time series in the form of a sales curve, since they have values which are spaced at uniform time intervals/periods (e.g., weekly). For example, the time series can assess the

predicted sales of an item for the next 26 weeks. The time series are typically in the form of a graph/curve of sales vs. periods.

[0036] The lengths of the plans can vary in time, which provides difficulties in adapting a plan for one product to another. Often the shape of a plan/curve is common to multiple groups of items, but the products may have different lifecycles or selling seasons. It is common for a retail planner to create one sales plan, and then uses a tool to stretch or shrink it to fit a season of different length. For example, an established sales curve may extend for 13-weeks, but a retailer may wish to use the same curve for a high fashion item that will only sell for 6-weeks. However, the known tools tend to distort the leading and trailing edges of the curve.

[0037] Further, the shape of the sales curve towards the end of the lifecycle directly influences the need for price markdowns. The shape of the curve drives the forecast of the item. If the values are higher than they should be, the forecast for the item is higher. The result is that the retailer risks having unsold items by not aggressively marking down prices. If the shape is lower than expected, the forecast is lower, and the retailer tries to increase demand by marking down the prices. The result is margin erosion due to unnecessary markdowns.

[0038] In one embodiment, short life cycle product demand forecast tool 110 is implemented on computing device 105 and includes logics or modules for implementing various functional aspects of short life cycle product demand forecast tool 110. In one embodiment, short life cycle product demand forecast tool 110 includes visual user interface logic/module 120, similar item sales curves generation logic/module 130, sales curves parameterization logic/module 140, and demand forecast generation logic/module 150.

[0039] Other embodiments may provide different logics or combinations of logics that provide the same or similar functionality as seasonality prediction model tool 110 of Fig. 1. In one embodiment, short life cycle product demand forecast tool 110 is an executable application including algorithms and/or program modules configured to perform the functions of the logics. The application is stored in a non-transitory computer storage medium. In one embodiment, the logics of short life cycle product demand forecast tool 110 are implemented as modules of instructions stored on a computer-readable medium.

[0040] Computer system 100 also includes a display screen 24 operably connected to computing device 105. In accordance with one embodiment, display screen 24 is implemented to display views of and facilitate user interaction with a graphical user interface ("GUI") generated by visual user interface logic 120 for viewing and updating information associated with generating the short life cycle product demand forecast (e.g., seasonality curves, sales data, etc.). The graphical user interface may be associated with a demand forecast application and visual user interface logic 120 may be configured to generate the graphical user interface.

[0041] In one embodiment, computer system 100 is a centralized server-side application that provides at least the functions disclosed herein and that is accessed by many users via computing devices/terminals communicating with the computer system 100 (functioning as the server) over a computer network. Therefore, display screen 24 may represent multiple computing devices/terminals that allow users to access and receive services from short life cycle product demand forecast tool 110 via networked computer communications.

[0042] In one embodiment, computer system 100 further includes at least one database 17 operably connected to computing device 105 and/or a network interface to access database 17 via a network connection. For example, in one embodiment, database 17 is operably connected to visual user interface logic 120. In accordance with one embodiment, database 17 is configured to store and manage data structures (e.g., records of sales data) associated with short life cycle product demand forecast tool 110 in a database system (e.g., a computerized inventory management and demand forecasting application).

[0043] In one embodiment, visual user interface logic 120 is configured to generate a graphical user interface ("GUI") to facilitate user interaction with seasonality prediction model tool 110. For example, visual user interface logic 120 includes program code that generates and causes the graphical user interface to be displayed based on an implemented graphical design of the interface. In response to user actions and selections via the GUI, associated aspects of generating sales curves and prediction models for short life cycle retail items may be manipulated.

[0044] For example, in one embodiment, visual user interface logic 120 is configured to facilitate receiving inputs and reading data in response to user actions.

For example, visual user interface logic 120 may facilitate selection, reading, and inputting of sales data (seasonality information and unit sales data or monetary sales data) associated with retail items sold at retail locations. The sales data may reside in at least one data structure (e.g., within database 17) associated with (and accessible by) a demand forecast application (e.g., short life cycle product demand forecast tool 110) via the graphical user interface.

[0045] Sales data may include, for example, data representing past sales and promotions of an item that is similar to the short life cycle item for which demand is to be determined across a plurality of past retail periods. The sales data may be segmented into retail periods of past weeks, with each past week having numerical values assigned to it to indicate the number of items sold (or monetary amount acquired for items) for that week. The sales data may also include numerical values representing price discounts and values of other promotion components across the retail periods, and seasonality information (which may also be separate from the sales data) in accordance with one embodiment. The sales data for an item may be accessed via network communications, in accordance with one embodiment.

[0046] In one embodiment, similar item sales curves generation logic/module 130 is configured to generate sales curves at different intersection levels for an item that is similar to the short life cycle product for which the demand forecast is being determined, as disclosed below.

[0047] In one embodiment, sales curves parameterization logic/module 140 is configured to parametrize each of the sales curves generated by logic 130. In one embodiment, demand forecast generation logic/module 150 generates a forecast of demand of the short life cycle product by aggregating the parameters generated at 140 using error metrics to arrive at a final sales curve for the short life cycle product that provides a demand forecast.

[0048] In one embodiment, the generated prediction of demand predicts an amount of needed inventory (e.g., for an item at a single store) which is then used for orders to a computerized inventory system (e.g., by a computerized inventory management and demand forecasting system). The prediction of demand may also control an amount of inventory (e.g., for an item at a single store) to be allocated by the computerized inventory system. The prediction of demand may further control

adjustment of an amount inventory (e.g., for an item at a single store) by the computerized inventory system.

[0049] Fig. 2 is a block diagram of computer server/system 100 in accordance with an embodiment of the present invention. Fig. 2 illustrates further hardware/software details of system 100. Although shown as a single system, the functionality of system 100 can be implemented as a distributed system. Further, the functionality disclosed herein can be implemented on separate servers or devices that may be coupled together over a network. Further, one or more components of system 100 may not be included. For example, for functionality of a server, system 100 may need to include a processor and memory, but may not include one or more of the other components shown in Fig. 2, such as a keyboard or display.

[0050] System 100 includes a bus 12 or other communication mechanism for communicating information, and a processor 22 coupled to bus 12 for processing information. Processor 22 may be any type of general or specific purpose processor. System 100 further includes a memory 14 for storing information and instructions to be executed by processor 22. Memory 14 can be comprised of any combination of random access memory ("RAM"), read only memory ("ROM"), static storage such as a magnetic or optical disk, or any other type of computer readable media. System 100 further includes a communication device 20, such as a network interface card, to provide access to a network. Therefore, a user may interface with system 100 directly, or remotely through a network, or any other method. Some or all of the components of system 100 can implement the entirety

[0051] Computer readable media may be any available media that can be accessed by processor 22 and includes both volatile and nonvolatile media, removable and non-removable media, and communication media. Communication media may include computer readable instructions, data structures, program modules, or other data in a modulated data signal such as a carrier wave or other transport mechanism, and includes any information delivery media.

[0052] Processor 22 is further coupled via bus 12 to display 24, such as a Liquid Crystal Display ("LCD"). A keyboard 26 and a cursor control device 28, such as a computer mouse, are further coupled to bus 12 to enable a user to interface with system 100.

[0053] In one embodiment, memory 14 stores software modules that provide functionality when executed by processor 22. The modules include an operating system 15 that provides operating system functionality for system 100. The modules further include a short life cycle demand forecasting module 16 that implements one or more of modules 120, 130, 140, 150, and all other functionality disclosed herein. System 100 can be part of a larger system. Therefore, system 100 can include one or more additional functional modules 18 to include the additional functionality, such as a retail management system (e.g., the "Oracle Retail Demand Forecasting System" or the "Oracle Retail Advanced Science Engine" ("ORASE") from Oracle Corp.) or an enterprise resource planning ("ERP") or other type of inventory management system. Database 17 is coupled to bus 12 to provide centralized storage for modules 16 and 18 and store customer data, product data, transactional data, etc. In one embodiment, database 17 is a relational database management system ("RDBMS") that can use Structured Query Language ("SQL") to manage the stored data. In one embodiment, a specialized point of sale ("POS") terminal 99 generates the transactional data and historical sales data (e.g., data concerning transactions of each item/SKU (stock-keeping unit) at each retail store) used to forecast demand. POS terminal 99 itself can include additional processing functionality to forecast demand in accordance with one embodiment and can operate as a specialized demand forecasting system either by itself or in conjunction with other components of Fig. 2.

[0054] In one embodiment, particularly when there are a large number of retail stores, a large number of items, and a large amount of historical data, database 17 is implemented as an in-memory database ("IMDB"). An IMDB is a database management system that primarily relies on main memory for computer data storage. It is contrasted with database management systems that employ a disk storage mechanism. Main memory databases are faster than disk-optimized databases because disk access is slower than memory access, the internal optimization algorithms are simpler and execute fewer CPU instructions. Accessing data in memory eliminates seek time when querying the data, which provides faster and more predictable performance than disk.

[0055] In one embodiment, database 17, when implemented as a IMDB, is

implemented based on a distributed data grid. A distributed data grid is a system in which a collection of computer servers work together in one or more clusters to manage information and related operations, such as computations, within a distributed or clustered environment. A distributed data grid can be used to manage application objects and data that are shared across the servers. A distributed data grid provides low response time, high throughput, predictable scalability, continuous availability, and information reliability. In particular examples, distributed data grids, such as, e.g., the "Oracle Coherence" data grid from Oracle Corp., store information in-memory to achieve higher performance, and employ redundancy in keeping copies of that information synchronized across multiple servers, thus ensuring resiliency of the system and continued availability of the data in the event of failure of a server.

[0056] In one embodiment, system 100 is a computing/data processing system including an application or collection of distributed applications for enterprise organizations, and may also implement logistics, manufacturing, and inventory management functionality. The applications and computing system 100 may be configured to operate with or be implemented as a cloud-based networking system, a software-as-a-service ("SaaS") architecture, or other type of computing solution.

[0057] Embodiments are disclosed from the perspective that, for an item (i.e., a class of items such as yogurt or men's shirts or an individual SKU) sold at a location (e.g., a retail location), the item may be promoted in various ways at various times (i.e., pre-defined retail periods, such as a day, week, month, year, etc.). A retail calendar has many retail periods (e.g., weeks) that are organized in a particular manner (e.g., four (4) thirteen (13) week quarters) over a typical calendar year. A retail period may occur in the past or in the future. Historical sales/performance data may include, for example, a number of units of an item sold in each of a plurality of past retail periods as well as associated promotion data (i.e., for each retail period, which promotions were in effect for that period) and any other relevant demand features/variables.

[0058] Embodiments generate a short life cycle curve estimation by combining the sales pattern at different intersections and using the bass diffusion model. Embodiments generate more accurate and robust short life cycle curves

compared to known approaches.

[0059] In order to demand the forecast for a short life cycle item, which in general may only be available for a few months and, since it is being newly introduced, there generally is no historical sales history available, historical sales data for similar items must be received and used. For example, if the short life cycle item is a laptop computer, which will be for sale for approximately 10 weeks, and then replaced by a similar laptop computer with maybe an improved processor, or larger amount of solid-state drive ("SSD"), sales history for previous "versions" of that laptop will be received. Similarly, for a fashion item such as a sports-related t-shirt, similar previously sold t-shirts can be considered.

[0060] Embodiments receive sales history for the "similar" sales item at different intersection levels, such as SKU/store, subclass/store, department/region, etc., and then generate different sales/demand curves for the similar item at each of the different intersection levels. The similar sales item can be similar in that it sold for approximately the same time period the previous year, for example, or that it is merely similar based on attributes (e.g., a t-shirt, a computer, etc.) even if the time frame during which it was sold spanned multiple months or years, so it is not even considered a short life cycle item. The latter scenario may be used if there does not exist another similar item previous sold that is also considered a short life cycle item. In general, retailers typically pick the "similar" product based on merchandise properties such as type, size, color, shape, etc.

[0061] In some embodiments, any known solution can be used to determine the sales/demand curves for the similar item at each of the different intersection levels. In one embodiment, instead of generating the sales curves at different intersection levels, auto clustering prediction models are used to generate the sales curves for a cluster of similar items, which can be combined with a sales curve of a similar item. In this embodiment, promotion effects/features are estimated using pooled sales data points at aggregated levels and auto clustering the data points. Each of the multiple cluster models are trained using regression and each saved cluster trained model is then used to predict demand for retail products/items.

[0062] As discussed above, in the retail industry, retailers need to predict future demand to better manage their inventory or promotion/markdown planning.

To accurately forecast demand, retailers consider all factors that could impact the demand such as promotions, price change, seasonality, weather and so on. Known solutions for retailers have used various algorithms to estimate the promotion or price effects.

[0063] Further, in general single product/location combinations do not have enough sales observations to produce robust effects estimation. To account for this, a large number of observations pooled from across different products/locations/periods are typically used to estimate the effects. They are then used to forecast demand for the products/locations that were pooled during estimation phase. Typically, known solutions pool the data based on information related to the product and location hierarchies (e.g., pooling the data within same subclass and region). Then, during forecasting, the promotion effects will be the same for the entire forecast horizon, and also the same for all the products and locations which participated in the estimation.

[0064] However, these known solutions ignore at least two facts: (1) Every product/location is affected differently by the same promotion. For example, a low sales product would be affected differently than a high sales product, or different store formats would be affected differently; (2) The customer responds differently to the same event depending on the time of year. In other words, the effects of promotions are time-sensitive. For example, an ice cream promotion would be more effective during a hot summer as opposed to during a cold winter.

[0065] In contrast, embodiments use machine learning and auto clustering to estimate the impact of promotions on demand by taking into account the timing of an event, as well as the details of every product/location during estimation with pooled data. The details and timing, collectively, are referred to as “features”. Examples of features include base sales, price, seasonality, brand, promotions, size, color, pack size, supplier, length, etc. While features such as price and seasonality may be relevant for all types of products, some others are item specific. For example, pack size impacts the demand for yogurts, however the length of the pack is insignificant. Conversely, the brand is very important for fashion items, but is much less important for hardware items, such as nails or hammers.

[0066] Embodiments estimate promotion effects at aggregated levels with

pooling all the data together. The resulting effects are product/location specific and time-sensitive, meaning they can change depending on the time period an event occurs. In general, in embodiments, promotion effects are estimated with pooling data at aggregated levels and different product/location/time are automatically grouped into clusters. The promotion effects are dynamically fetched during the forecast phase, based on the product/location and time-related features. Even for new product/location combinations, with no historical demand, embodiments are able to generate time-sensitive promotion effects.

[0067] Fig. 3 is a flow diagram of the functionality of short life cycle product demand forecast tool 110 of Fig. 1 when estimating promotional effects that can be used for a demand forecast in accordance with one embodiment. In one embodiment, the functionality of the flow diagram of Fig. 3 (and Fig. 6 below) is implemented by software stored in memory or other computer readable or tangible medium, and executed by a processor. In other embodiments, the functionality may be performed by hardware (e.g., through the use of an application specific integrated circuit ("ASIC"), a programmable gate array ("PGA"), a field programmable gate array ("FPGA"), etc.), or any combination of hardware and software.

[0068] At 302, historical item sales data is received for all items/SKUs for all stores, for a particular class/category of products, or for only a single item of interest. For example, the class/category can be "yogurt", "coffee" or "milk." Each class has one or more subclasses, all the way down to the SKU or Universal Product Code ("UPC") level, which would be each individual item for sale. For example, for the class of yogurt, a sub-class could be each brand of yogurt, and further sub-classes could be flavor, size, type (e.g., Greek or regular), down to an SKU which would correspond to every individual different type of yogurt item sold.

[0069] Historical sales and performance data may include, for example, data representing past sales and promotions of each item across a plurality of past retail sales periods. The historical performance data may be segmented into retail periods of past weeks, with each past week having numerical values assigned to it to indicate the number of items sold for that week. The historical performance data may also include numerical values representing price discounts and values of other promotion components across the retail periods, in accordance with one

embodiment. The historical performance data for an item may be accessed via network communications, in accordance with one embodiment, including being accessed from each POS terminal 99 at each retail store and/or accessed from database 17.

[0070] The historical performance data includes sales data associated with the plurality of promotion components across a plurality of time periods (e.g., weeks). Examples of promotion components include, but are not limited to, a price discount component, a television advertisement component, a radio advertisement component, a newspaper advertisement component, an email advertisement component, an internet advertisement component, and an in-store advertisement component. The historical data includes, for each item, a listing of feature/variables/attributes for the item, such as price, promotions, seasonality, brand, color, style, etc.

[0071] The historical sales data is received as multiple data points or a "data set", with a single data point for each sales of an item per store (i.e., at the product/store/week level in embodiments where "week" is the desired time period). In embodiments, the data points can be received by electronically parsing data generated by all POSs 99 at all relevant retail stores.

[0072] At 304, at an aggregate product/location (store) level, different types of features are extracted. Many of the features are extracted from non-sales related data, such as from ERP, merchandising systems and inventory management systems associated with the relevant stores. Some of the features (e.g., sales data, promotions in place during the sales period) are extracted from the sales data of 302. The extracted types of features in embodiments include the following:

a. Product related features $\mathbf{p_1 .. p_n}$ (e.g., national brand vs. private label, package size, low sales vs. high volume, unit of measures ("UOM"), etc.);
b. Store features $\mathbf{s_1 .. s_m}$ (e.g., convenience store vs. supermarket, population dense area vs. non population area, different store layouts, etc.);

c. Timing related features $\mathbf{t_1}$.. $\mathbf{t_k}$ (e.g., holiday season vs. non holiday season, hot day vs cold day, low store traffic vs. high store traffic, football season, etc.);
d. Promotions and prices $\mathbf{x_1}$.. $\mathbf{x_j}$.

[0073] At 306, embodiments generate "N" clusters of the data points from 302 using the product features ($\mathbf{p_1}$.. $\mathbf{p_n}$), location features($\mathbf{s_1}$.. $\mathbf{s_m}$) and timing features ($\mathbf{t_1}$.. $\mathbf{t_k}$) from 304 to generate N clusters ($\mathbf{c_1}$.. $\mathbf{c_N}$). The number of clusters N can be a user defined number. Clustering is a form of data mining in which a set of objects (i.e., the features) are grouped in such a way that objects in the same group (the cluster) are more similar in some sense to each other than to those in other groups (clusters). In one embodiment, the k-means clustering algorithm is used. k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. However, other clustering algorithms such as hierarchical clustering, etc., can also be used.

[0074] In one embodiment, the following k-means clustering algorithm is used:

Given an initial set of *k* means $m_1^{(1)},\ldots,m_k^{(1)}$, the algorithm proceeds by alternating between two steps:

(1) Assignment step: Assign each observation to the cluster with the nearest mean: that with the least squared Euclidean distance. (Mathematically, this means partitioning the observations according to the Voronoi diagram generated by the means.)

$$S_i^{(t)} = \left\{ x_p : \left\| x_p - m_i^{(t)} \right\|^2 \leq \left\| x_p - m_j^{(t)} \right\|^2 \forall j, 1 \leq j \leq k \right\},$$

where each x_p is assigned to exactly one $S^{(t)}$, even if it could be assigned to two or more of them.

(2) Update step: Recalculate means (centroids) for observations assigned to each cluster.

$$m_i^{(t+1)} = \frac{1}{\left|S_i^{(t)}\right|} \sum_{x_j \in S_i^{(t)}} x_j$$

The algorithm has converged when the assignments no longer change.

[0075] At 308, each of the clusters generated at 306 are trained (i.e., estimating the effects) to generate trained cluster models (i.e., promotion effects per cluster generated at 306). In one embodiments, the training is done by running a regression for each cluster *i* of the N clusters and generate N sets of promotion/price effects $(\mathbf{e_1}..\mathbf{e_k})_1...(\mathbf{e_1}..\mathbf{e_k})_N$. In one embodiment, linear regression is used for the training, or estimating the effects, but other regression algorithms can be used. Each set of promotion/price effects is an optimized set of effects for the features of the corresponding cluster. The features in each cluster can be decomposed into product/location/time period. Every product/location/time period maps to one of the N sets of effects.

[0076] At 310, the clustering models are saved into a binary file that can easily be loaded on demand. For example, in embodiments, the models are trained with features in a machine learning specific language, such as Python or C++. These models can be saved into a binary file with a specific serialization format. For example, a model trained in Python can be saved to a python pickle binary file.

[0077] At 312, when a demand forecast for a retail product is desired, for each product/location/week, the saved cluster trained model is used to map it to one of the N clusters by identifying the features for the particular week. For example, for cluster *i*, promotion/price effect from effects set $(\mathbf{e_1}..\mathbf{e_k})_i$ is fetched.

[0078] At 314, the promotion/price effects from 312 are used to predict a final demand. In one embodiment, the following demand forecast algorithm is used to predict demand: demand = base demand * promotion lift * price lift. This is a relatively simple demand algorithm and much more complex demand forecast algorithms can be used in other embodiments as long as they incorporate the promotion effects in the demand forecast in any way. Embodiments can forecast demand on a per SKU/store basis, or forecast demand for all SKUs in all stores at the same time using a large matrix (i.e., each row is one SKU/store).

[0079] The functionality of Fig. 3 can be used to generate multiple sales curves for an SKU at different intersections by, for each given intersection, such as product class/region, clustering the product/location based on the product attributes such as color, size, store format and sales length into different groups. Then the short lifecycle sales curves disclosed below as used in connection with Fig. 6 are generated (i.e., the demand forecast for each intersection).

[0080] Figs. 4A-B and 5 illustrate an example of an implementation of embodiments of the invention, and how each example corresponds to the functionality of Fig. 3. In the relatively simplified example, it is assumed that the retailers use the following formula to model demand:

demand = base demand * promotion lift * price lift;

and, for simplicity, it is assumed that the base demand is 5 for all product/store combinations.

[0081] At 302, historical sales data points are received. In the simplified example, it is assumed that there are two stores and information regarding 6 products from week1_2018 to week52_2019 (i.e., every week in 2018 and 2019).

[0082] At 304, embodiments associate every item/store/week sales data received at 302 with product, location, and time period related extracted features, including the following:

product brand, package size, store type, weather during the corresponding week, what sporting events are on during the corresponding week, price, promotions 1, 2, 3, 4, etc.

[0083] At 306, the clustering algorithm is implemented to cluster the sales into N clusters (i.e., three clusters in this example). In Fig. 4A, each cluster (402, 403, 404) is denoted by a different cross-hatching scheme. The tables shown in Fig. 4A assign one of the three clusters to each product, at each store, during each of the weeks as a result of the clustering.

[0084] At 308, for each cluster 402-404, the regression models are trained with the sales data and its features by using data in each cluster to generate promotion effects per cluster 402-404. Fig. 4B shows for each of the clusters, the trained promotion effects for price and promotions 1-4. The generated effects at 308 are based on the effects that form the demand model (in this simplified example,

only price and promotions). In more complicated demand models, effects generated can also include holidays, back-to-school events, weather, etc. Each effect number in Fig. 4B functions as a multiplier.

[0085] At 310, the trained clustering models and three trained regression models are saved in a binary file. The binary file can include both clustering models and regression models. Fig 4A illustrates clustering models while Fig 4B illustrates the regression models. In embodiments, clustering and regression could be done separately and then saved to different binary files or a single file.

[0086] At 312, to create the forecast for week1_2020 to week4_2020 (i.e., the first 4 weeks of 2020), the features of those weeks are first identified. Then the saved cluster model is used to predict every item/store/week to one of the three clusters. Fig. 5 illustrates how embodiments map each product/location/week to a cluster. So, for example, Store1, Product B, week 2 is mapped to cluster 404, while Store2, Product C, Week 3 is mapped to cluster 403.

[0087] At 314, for each item/store/week, the effects (price, promotions) in the corresponding cluster are applied on top of the base demand, using the demand formula, to create the forecast. Specifically, for this example, where demand = base demand * promotion lift * price lift, assume base demand of product A in store 1 is 2.0, and promotion 1 & 3 is on during week2_2020, then the demand of week2_2020 of product A at store 1 is: Demand = 2 * 1.05*3.2=6.72.

[0088] Fig. 6 is a flow diagram of the functionality of short life cycle product demand forecast tool 110 of Fig. 1 when generating a sales/demand curve for the short life cycle item (i.e., a given item/location or SKU) in accordance with one embodiment.

[0089] At 602, multiple sales curves related to an item or items that are similar to the short life cycle item are generated. In general, these multiple sales curves are then “blended” together using the remaining Fig. 6 functionality.

[0090] In one embodiment, for a similar item to the short life cycle item, a sales curve is generated at different (“n”) intersection levels by using historical sales data for similar item (since there is no or sparse sales history available for the short life cycle item). For example, the intersection levels may be SKU/store, subclass/store, department/region, etc. The historical sales from the SKU/store are

aggregated to the higher intersection levels, and the sales curve for that intersection level is determined. Any known functionality can be used for determining the sales curves (i.e., forecasted demand) in some embodiments.

[0091] In one embodiment, the functionality disclosed above in conjunction with Fig. 3 is used to generate multiple sales curves for similar items at 602. Specifically, the end result of the functionality of Fig. 3 may be a first SKU and corresponding cluster of similar items, a second SKU and a different corresponding cluster of similar items, etc. The sales curve for each SKU and the sales curve for each corresponding cluster can be generated using known methods. Assuming the first and second SKUs are also similar to the short life cycle item, as disclosed above, these sales curves can be used as input to the functionality of Fig. 6 below and are then "blended" together to form the short life cycle item sales curve. In one embodiment, only two sales curves may be generated – a single similar SKU and corresponding cluster. In other embodiments, when multiple similar SKUs are available, many sales curves may be generated for each similar SKU and for each corresponding cluster.

[0092] When the functionality of Fig. 3 is used, at 304 the features extracted in one embodiment can include extracting a plurality of different types of features related to sales of each of the products, including merchandise properties such as color and size, location properties such as price zone, store format, calendar related properties such as total weeks on the shelf, first week on the shelf, etc.

[0093] The time frame of the sales curves at 602 will approximately match the time frame of the desired sales curve for the short life cycle item in embodiments. In other words, if the short life cycle item is expected to have a 10-week selling season during the Summer of 2021, then the sales curves for the similar item should also be directed to those same 10 weeks, or closely match (e.g., 12 weeks).

[0094] At 604, the sales curve at each intersection "*i*" is parameterized using the "Bass Diffusion" model. The Bass Diffusion model includes a differential equation that describes the process of how new products get adopted in a population. The model presents a rationale of how current adopters and potential adopters of a new product interact. The basic premise of the model is that adopters can be classified as innovators or as imitators and the speed and timing of adoption

depends on their degree of innovativeness and the degree of imitation among adopters. The Bass Diffusion model has been widely used in forecasting, especially new products' sales forecasting and technology forecasting. Mathematically, the basic Bass Diffusion model is a Riccati equation with constant coefficients.

[0095] In embodiments, the Bass Diffusion model is used to parameterize each sales curve from 602 as follows:

For each intersection (or each sales curve in other embodiments) *i*,:

estimate the m_i, p_i, q_i parameters/coefficients based on the following equation by using the least square optimizer to minimize the objection MSE(*i*)

$$F_i(t) = m_i \frac{(p_i + q_i)^2}{p_i} * \frac{e^{-(p_i+q_i)t}}{(1 + \frac{q_i}{p_i} e^{-(p_i+q_i)t})^2}$$

$$E_i = \sum_{t=1}^{\square} (F_i(t) - S_i(t))^2$$

Where $F_i(t)$ is the estimate sales at week *t* of intersection *i* and $S_i(t)$ is the actual sales history at week *t* of intersection *i*. E_i is the error for current intersection *i*.

[0096] At 606, the p_i and q_i parameters from 604 are combined based on a weighted error. The p_i is also referred to as the coefficient of innovation, external influence or advertising effect. The q_i parameter is also referred to as the coefficient of imitation, internal influence or word-of-mouth effect. The m_i parameter is not needed or used. The p_i and q_i parameters from 604 are combined as follows, where W_i is a determined weight for each intersection *i*:

$$W_i = \frac{E_i}{\sum_{i=1}^{n} E_i}$$

$$p = \sum_{i=1}^{n} W_i * p_i$$

$$q = \sum_{i=1}^{n} W_i * q_i$$

[0097] At 608, for the given length of short life cycle sales length l,

embodiments calculate the curve value at τ'th week as follows:

$$c(\tau) = l\frac{(p+q)^2}{p} * \frac{e^{-(p+q)\tau}}{(1+\frac{q}{p}e^{-(p+q)\tau})^2}$$

The final curve of the given product/location at τ'th will be calculated at 606 by normalizing to l as:

$$C(\tau) = \frac{l * c(\tau)}{\sum_{\tau=1}^{l} c(\tau)}$$

which generates a final sales/demand curve that provides a forecast of sales volume for the short life cycle item for each week (or other predefined time period) over the selling time period (e.g., weeks 20-30 of year 2021).

[0098] At 610, the sales curve or demand forecast from 608 is used for logistics, including being the basis for determining how many of the items to manufacture, inventory levels, and levels of shipments to specific stores. The output at 608 may be in the form of a specialized data structure that can be used with fully automated manufacturing, inventory, and logistic systems, as described more fully below.

[0099] Figs. 7 and 8 illustrate an example of an implementation of embodiments of the invention, and how each example corresponds to the functionality of Fig. 6. In the relatively simplified example, at 602, sales curves for the similar item are generated for only two intersection levels: SKU/store and SKU/region. Fig. 7 illustrates the SKU/store sales curve at 701, and the SKU/region sales curve at 702. In Fig. 7, the X-axis is the calendar (e.g., week 1, week 2, etc.) and the Y-axis is the shape/value of the curve. The sales curve is usually normalized to the number of weeks. For example, for a ten week sales curve, the curve value may be: 0.76, 0.88, 1.02, 1.15, 1.2, 1.28, 1.11, 1.05, 0.8, 0.75. As noted, the values add up to ten, which is equal to the length of the curve in weeks.

[00100] At 604 and 606 the m_i, p_i, q_i E_i and W_i parameters are calculated (i.e., the sales curves are "parameterized"). 801 in Fig. 8 illustrate the calculations in the example.

[00101] At 606, the *p* and *q* parameters are combined based on the weighted error. 802 in Fig. 8 illustrates the calculations in the example. After the *p* and *q* parameters are combined, the combined sales curve is calculated at 608. Sales curve 703 of Fig. 3 is the combined sales curve, and represents the demand forecast for the short life cycle item. The length of the curves of Fig. 7 do not necessarily have to be equal. The equations include a factor 'l' - which is the length of the curve. This allows combining curves of different lengths. The final sales curve is not showing number of items forecast to be sold. Instead, it is showing the shape of the sale curve during its life cycle.

[00102] Fig. 9 illustrates an integrated manufacturing, inventory and logistics system 900 that includes demand forecasting as disclosed herein in accordance with one embodiment. Embodiments can use the specialized demand forecast data structure of the short life cycle item generate at 608 to fully automate the processes. As shown in Fig. 9, system 900 can include a product demand forecasting system 970 that forecasts future product demand and in some instances forecasts and/or considers future demand for hundreds of thousands of products, or in some applications tens of millions or more products at one or more retail stores 901-904. Forecasting system 970 is in communication through a cloud network 950 or other type of communications network with one or more inventory systems 920 and one or more manufacturing systems 980.

[00103] Forecasting system 970 generates demand forecasting by implementing the functionality disclosed in conjunction with Fig. 3 and or Fig. 6 above. Inventory system 920 stores inventory and provides transportation logistics to deliver items to stores 901-904 using trucks 910-913 or some other transportation mechanisms. Inventory system 920 in one embodiment implements an Enterprise Resource Planning ("ERP") specialized computer system or a specialized inventory control system that uses input from demand forecasting system 970 to determine levels of inventories and the amount and timing of the delivery of items to stores 901-904. The functionality of Fig. 9 can be completely automated in some embodiments using automated loading mechanisms and self-driving transportation and the specialized data structures.

[00104] Manufacturing system 980 manufactures items to be sent to inventory

system 920 and provides transportation logistics to deliver the items to inventory system 920 using a truck 981 or some other transportation mechanisms. Manufacturing system 980 in one embodiment implements an ERP specialized computer system or a specialized manufacturing system that uses input from forecasting system 970 to determine an amount of items to manufacture, inventory of resources that are used for the manufacturing, and the amount and timing of the delivery of items to inventory system 920.

[00105] Forecasting system 970 can utilize information from inventory system 920, a sales tracking system (not shown) and/or databases in forecasting demand for products. In forecasting demand, forecasting system 970 attempts to predict uncharacteristic demand of one or more products that results from events, weather, social demand, economic factors and other factors. Tens, to hundreds to thousands of different variables may be tracked that can have an effect on the demand of one or more products. Changes in these variables can result in uncharacteristic demands. For example, changes in forecasted weather can be tracked, and one or more variables associated with the forecasted weather can be used in determining whether such a change is weather may have an effect on demand, and may further forecast a change in demand.

[00106] In general, the elements of Fig. 9 perform sales, manufacturing, or consumption of inventory. Retail locations/stores 901-904 for direct consumer sales exhibit the most volatile inventory patterns, due to the random nature and external factors affecting sales. However, manufacturing facilities and sites that consume inventory (such as product integrators, internet shippers, etc. products used in the local facility) also benefit from demand forecasting as disclosed herein. As disclosed, each retail location 901-904 sends sales data and historic forecast data to forecasting system 970. The sales data includes inventory depletion statistics for each item, or SKU/UPC for each sales period, typically days, in the previous sales cycles (i.e. weeks), typically 4-7 weeks of inventory cycles.

[00107] Forecasting system 970 stores the sales data in a repository 972, and employs the sales data for generating orders to replenish inventory. The orders include a set of items and a quantity for each item for maintaining the inventory level at a store 901-904.

[00108] Many retail ordering schemes rely on days of the week for sales periods and sales cycles. In one configuration, in an inventory management environment having inventory statistics, in which the inventory statistics are specific to each day of the week, inventory system 920 determines target inventory levels by gathering, for each day of the week, inventory level statistics from previous sales. Embodiments compute, based on the inventory level statistics, an inventory level for each day of the week, such that the safety stock accommodates variations in inventory between the different days of the week. Embodiments render, for each of a plurality of items, a stocking level indicative of the target inventory level including the safety stock for each day of the week. Embodiments compute an ordering quantity based on a lead time such that the ordered quantity arrives to satisfy the rendered stocking level on the determined day of the week. Identifying the actual stock levels includes identifying stock levels on the day of the week from previous weeks from the history data, thus focusing on the same day of the week over time, rather than an average of all days in the week.

[00109] In particular configurations, the disclosed embodiments may be employed in conjunction with specialized and/or particularly high volume retail sales environments. In large logistics and distribution operations, it is beneficial to load trucks as full as possible, and in the event deferral of items to a successive trip is needed, to select those items which will have a least likely chance of interrupting sales activity. Accordingly, embodiments are operable in conjunction with POS system 99 to identify high velocity or high turnover items that tend to be sold and replenished faster than other items. A UPC bar code symbol or radio-frequency identification ("RFID") on an item includes a field, designation or value, that alone or in conjunction with a database lookup, designates an item as a high velocity item appropriate for safety stock treatment as defined herein.

[00110] A high velocity item may be accommodated by identifying, for each of a plurality of items represented in an inventory database, a field for a product identifier and a field denoting a safety stock for the item, and determining, for each of the product identifiers, a product segmentation field based on product velocity indicative of increased product replenishment demands resulting from a sales volume. The disclosed embodiments determine based on the velocity field, whether

to compute a safety stock, i.e. whether the overhead and burden to resupply according to the safety stock is worthwhile given the product throughput.

[00111] In other embodiments, supply logistics may invoke a delivery frequency higher than one truck a day, hence triggering a resupply window with a higher granularity. In such a case, the safety stock may be more specific than an individual day, such as a Monday AM and Monday PM, or to designate multiple delivery or time windows within a particular day of the week, such as 7:00 AM, 11:00 AM and 4:00 PM.

[00112] Embodiments, including the generated demand forecast, may be employed in implementing supply logistics and designating deliveries (i.e., trucks) and manifest (i.e., contained items) in accordance with demand and profit margins of the transported items. High velocity items might be deemed to have priority space on a particular delivery, but could further be selected based on a profit margin or markup on the included items, and items with the greatest revenue generation potential selected for inclusion.

[00113] In such a product inventory shipping environment that uses the demand forecast disclosed herein and has a plurality of transport vehicles, each vehicle (e.g., truck) is configured for receiving a fixed payload of items for delivery to a sales location for inventory replenishment. Embodiments can provide guidance in loading a delivery vehicle, by, for each item of a plurality of items including a first item and a second item, computing a safety stock and determining, based on the computed safety stock of the first item and the second item, a quantity of each of the first item and the second item to be loaded into the delivery vehicle. Embodiments recompute a truck loading quantity based on the safety stock if insufficient space is available in the delivery vehicle for the determined quantity of the first item and the second item, meaning that certain items would need to be omitted and deferred to a successive delivery.

[00114] As disclosed, embodiments combine the sales pattern of an item similar to the short life cycle item at different intersections, such as SKU/store, subclass/store, department/region, etc. This makes the curves of the short life merchandise more robust. Embodiments then parametrize the curve using two parameters, and then regenerates the curve by applying the planned number of

weeks the product will be on shelf. This methodology makes the recalculation of the curve very convenient, even if the sales pattern changes year over year. Given the parametrization of the curves, combining the sales pattern is not dependent on the curves' lengths.

[00115] Several embodiments are specifically illustrated and/or described herein. However, it will be appreciated that modifications and variations of the disclosed embodiments are covered by the above teachings and within the purview of the appended claims without departing from the spirit and intended scope of the invention.

WHAT IS CLAIMED IS:

1. A method of generating a short life cycle sales curve for a short life cycle item comprising:

generating a plurality of similar sales curves corresponding to at least one similar item that is similar to the short life cycle item;

parametrizing each of the similar sales curves comprising estimating, for each similar sales curve, a coefficient of innovation parameter, a coefficient of imitation parameter, and an error parameter;

determining a weight for each error parameter;

combining the coefficient of innovation parameters and the coefficient of imitation parameters using the weights; and

generating the short life cycle sales curve using the combined coefficient of innovation parameters and the combined coefficient of imitation parameters.

2. The method of claim 1, further comprising receiving historical sales data for the similar item at different intersection levels, the historical sales data used as input for the generating the plurality of similar sales curves, wherein each similar sales curve is generated for a different intersection level.

3. The method of claim 1, further comprising generating a demand forecast for the short life cycle item using the short life cycle sales curve.

4. The method of claim 1, wherein the plurality of similar sales curves extend for a first duration of time, and the short life cycle sales curve extends for a second duration of time that is different than the first duration of time.

5. The method of claim 3, further comprising:

generating a specialized data structure that represents the demand forecast;

transmitting the specialized data structure to an automated logistics system;

wherein, in response to receiving the specialized data structure, the automated logistics system automatically transports one or more of the short life cycle item to at least one retailer.

6. The method of claim 1, the generating the plurality of similar sales curves comprising

receiving historical sales data for an aggregate products/locations level, the historical sales data comprising a plurality of sales data points, including sales data points for a first item at each of a plurality of locations;

extracting a plurality of different types of features related to sales of each of the similar items;

generating a plurality of clusters of sales data points based on the plurality of different types of features;

training each of the clusters to generate a plurality of trained cluster models comprising promotion effects per cluster;

for a particular time period, a particular location and the first item, identifying the features for the time period and mapping to one of the trained cluster models to fetch the promotion effects for the time period; and

generating a first sales curve corresponding to the first item and generating a second sales curve corresponding to the mapped trained cluster model.

7. The method of claim 2, the combining, at each intersection level, the coefficient of innovation parameter (p), a coefficient of imitation parameter (q), and an weight at each intersection i (W_i) based on the error at the intersection (E_i) comprising:

$$W_i = \frac{E_i}{\sum_{i=1}^{n} E_i}$$

$$p = \sum_{i=1}^{n} W_i * p_i$$

$$q = \sum_{i=1}^{n} W_i * q_i$$

.

8. The method of claim 7, the generating the short life cycle sales curve comprising, at τ'th week:

$$c(\tau) = l\frac{(p+q)^2}{p} * \frac{e^{-(p+q)\tau}}{(1+\frac{q}{p}e^{-(p+q)\tau})^2}$$

and normalizing to l as:

$$C(\tau) = \frac{l*c(\tau)}{\sum_{\tau=1}^{l} c(\tau)} .$$

9. A computer-readable medium having instructions stored thereon that, when executed by one or more processors, cause the processors to generate a short life cycle sales curve for a short life cycle item, the generating the sales curve comprising:

generating a plurality of similar sales curves corresponding to at least one similar item that is similar to the short life cycle item;

parametrizing each of the similar sales curves comprising estimating, for each similar sales curve, a coefficient of innovation parameter, a coefficient of imitation parameter, and an error parameter;

determining a weight for each error parameter;

combining the coefficient of innovation parameters and the coefficient of imitation parameters using the weights; and

generating the short life cycle sales curve using the combined coefficient of innovation parameters and the combined coefficient of imitation parameters.

10. The computer-readable medium of claim 9, further comprising receiving historical sales data for the similar item at different intersection levels, the historical sales data used as input for the generating the plurality of similar sales curves, wherein each similar sales curve is generated for a different intersection level.

11. The computer-readable medium of claim 9, further comprising generating a demand forecast for the short life cycle item using the short life cycle sales curve.

12. The computer-readable medium of claim 9, wherein the plurality of similar

sales curves extend for a first duration of time, and the short life cycle sales curve extends for a second duration of time that is different than the first duration of time.

13. The computer-readable medium of claim 11, further comprising:

generating a specialized data structure that represents the demand forecast;

transmitting the specialized data structure to an automated logistics system;

wherein, in response to receiving the specialized data structure, the automated logistics system automatically transports one or more of the short life cycle item to at least one retailer.

14. The computer-readable medium of claim 9, the generating the plurality of similar sales curves comprising

receiving historical sales data for an aggregate products/locations level, the historical sales data comprising a plurality of sales data points, including sales data points for a first item at each of a plurality of locations;

extracting a plurality of different types of features related to sales of each of the similar items;

generating a plurality of clusters of sales data points based on the plurality of different types of features;

training each of the clusters to generate a plurality of trained cluster models comprising promotion effects per cluster;

for a particular time period, a particular location and the first item, identifying the features for the time period and mapping to one of the trained cluster models to fetch the promotion effects for the time period; and

generating a first sales curve corresponding to the first item and generating a second sales curve corresponding to the mapped trained cluster model.

15. The computer-readable medium of claim 10, the combining, at each intersection level, the coefficient of innovation parameter (*p*), a coefficient of imitation parameter (*q*), and an weight at each intersection *i* (W_i) based on the error at the intersection (E_i) comprising:

$$W_i = \frac{E_i}{\sum_{i=1}^{n} E_i}$$

$$p = \sum_{i=1}^{n} W_i * p_i$$

$$q = \sum_{i=1}^{n} W_i * q_i$$

.

16. The computer-readable medium of claim 15, the generating the short life cycle sales curve comprising, at τ'th week:

$$c(\tau) = l\frac{(p+q)^2}{p} * \frac{e^{-(p+q)\tau}}{(1+\frac{q}{p}e^{-(p+q)\tau})^2}$$

and normalizing to l as:

$C(\tau) = \frac{l*c(\tau)}{\sum_{\tau=1}^{l} c(\tau)}$.

17. A retail item demand forecasting system comprising:

one or more processors coupled to one or more point of sale systems, the processors:

generating a plurality of similar sales curves corresponding to at least one similar item that is similar to the short life cycle item;

parametrizing each of the similar sales curves comprising estimating, for each similar sales curve, a coefficient of innovation parameter, a coefficient of imitation parameter, and an error parameter;

determining a weight for each error parameter;

combining the coefficient of innovation parameters and the coefficient of imitation parameters using the weights; and

generating the short life cycle sales curve using the combined coefficient of innovation parameters and the combined coefficient of imitation parameters.

18. The system of claim 17, the processors further receiving historical sales data for the similar item at different intersection levels, the historical sales data used

as input for the generating the plurality of similar sales curves, wherein each similar sales curve is generated for a different intersection level.

19. The system of claim 17, the processors further generating a demand forecast for the short life cycle item using the short life cycle sales curve.

20. The system of claim 19, the processors further:

generating a specialized data structure that represents the demand forecast;

transmitting the specialized data structure to an automated logistics system;

wherein, in response to receiving the specialized data structure, the automated logistics system automatically transports one or more of the short life cycle item to at least one retailer.

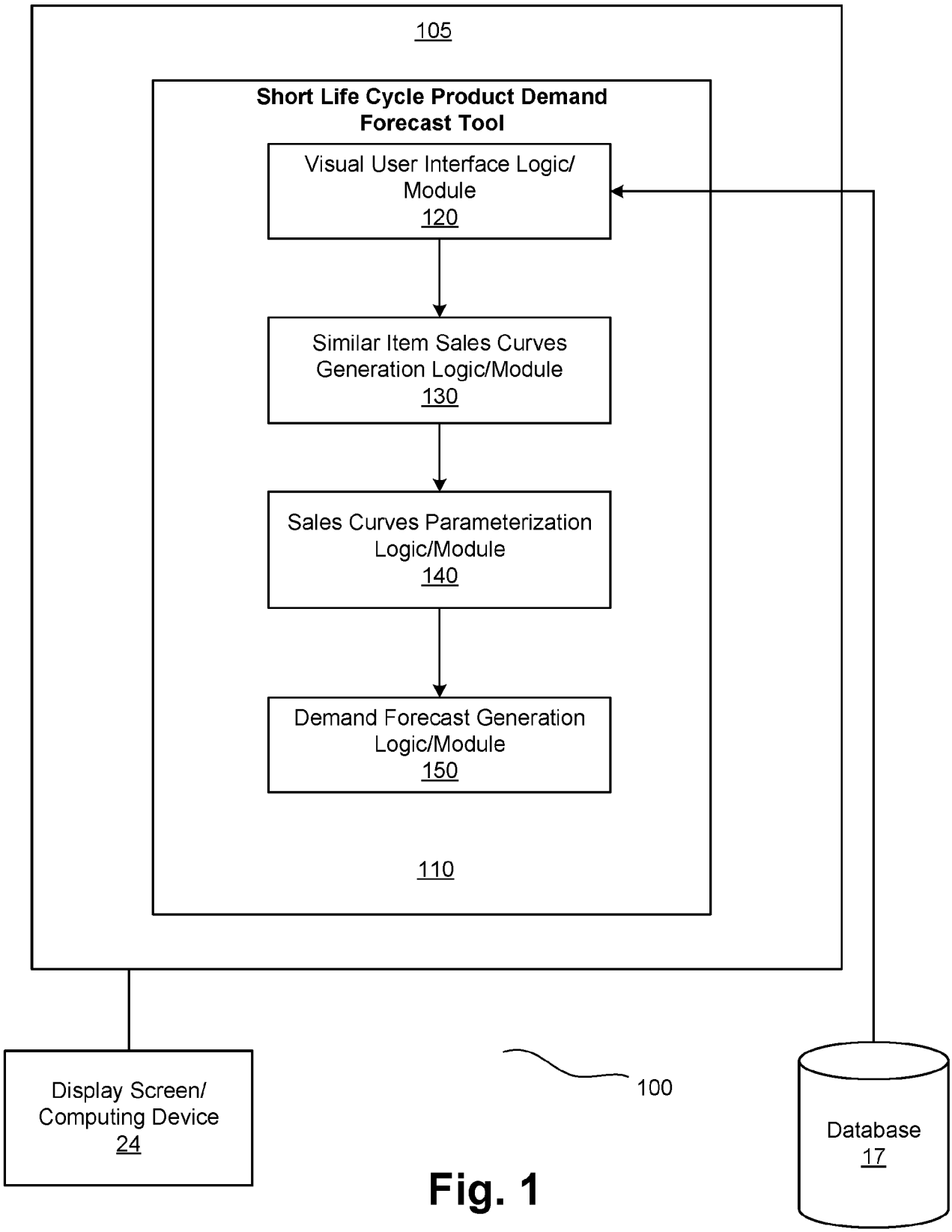

105
Short Life Cycle Product Demand Forecast Tool
Visual User Interface Logic/ Module 120
Similar Item Sales Curves Generation Logic/Module 130
Sales Curves Parameterization Logic/Module 140
Demand Forecast Generation Logic/Module 150
110
Display Screen/ Computing Device 24
100
Database 17

Fig. 1

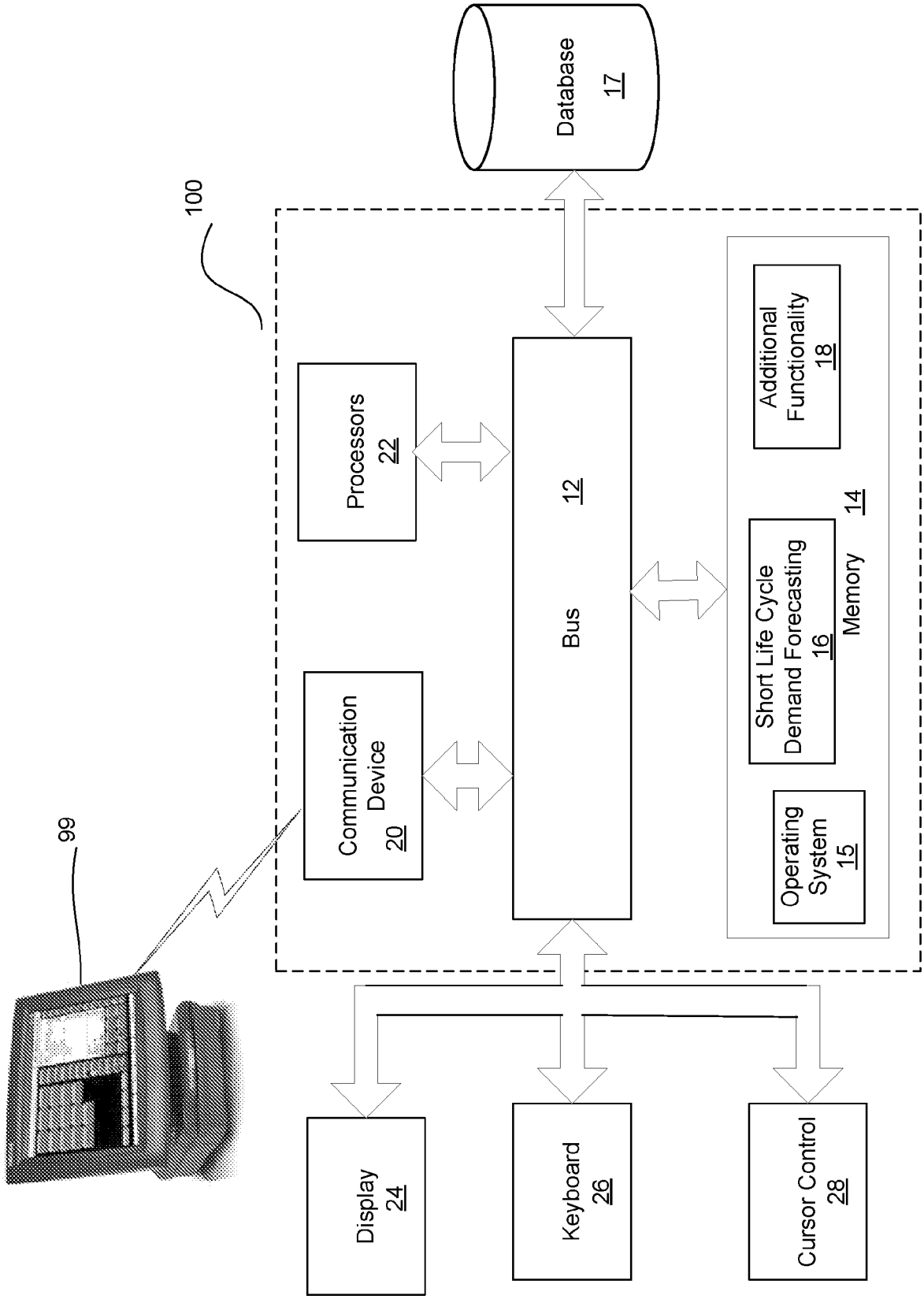

Database
17
100
Processors
22
12
Bus
Additional
Functionality
18
Short Life Cycle
Demand Forecasting
16
Memory
14
Operating
System
15
Communication
Device
20
99
Display
24
Keyboard
26
Cursor Control
28

Fig. 2

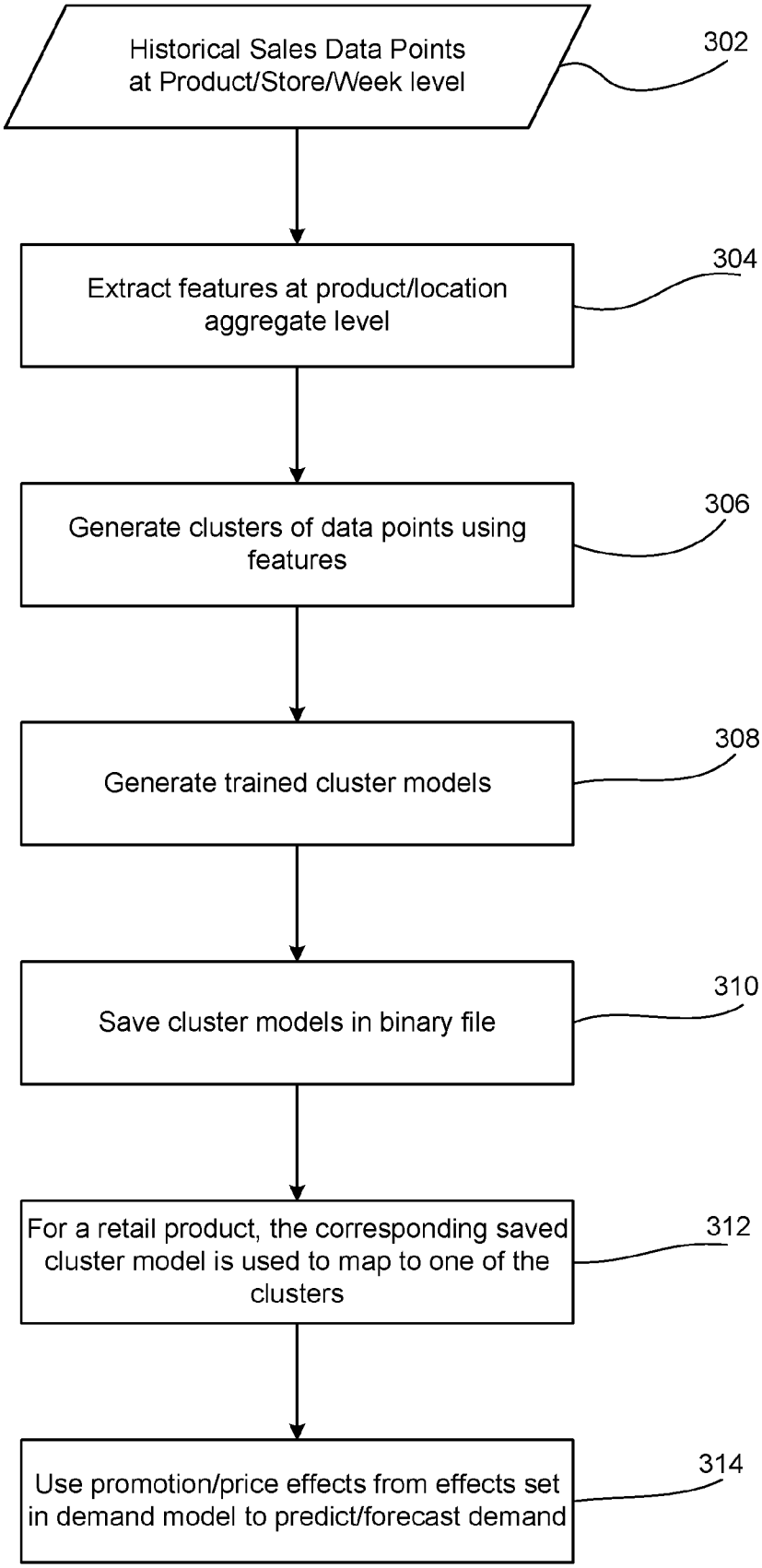

Historical Sales Data Points at Product/Store/Week level
302
Extract features at product/location aggregate level
304
Generate clusters of data points using features
306
Generate trained cluster models
308
Save cluster models in binary file
310
For a retail product, the corresponding saved cluster model is used to map to one of the clusters
312
Use promotion/price effects from effects set in demand model to predict/forecast demand
314

Fig. 3

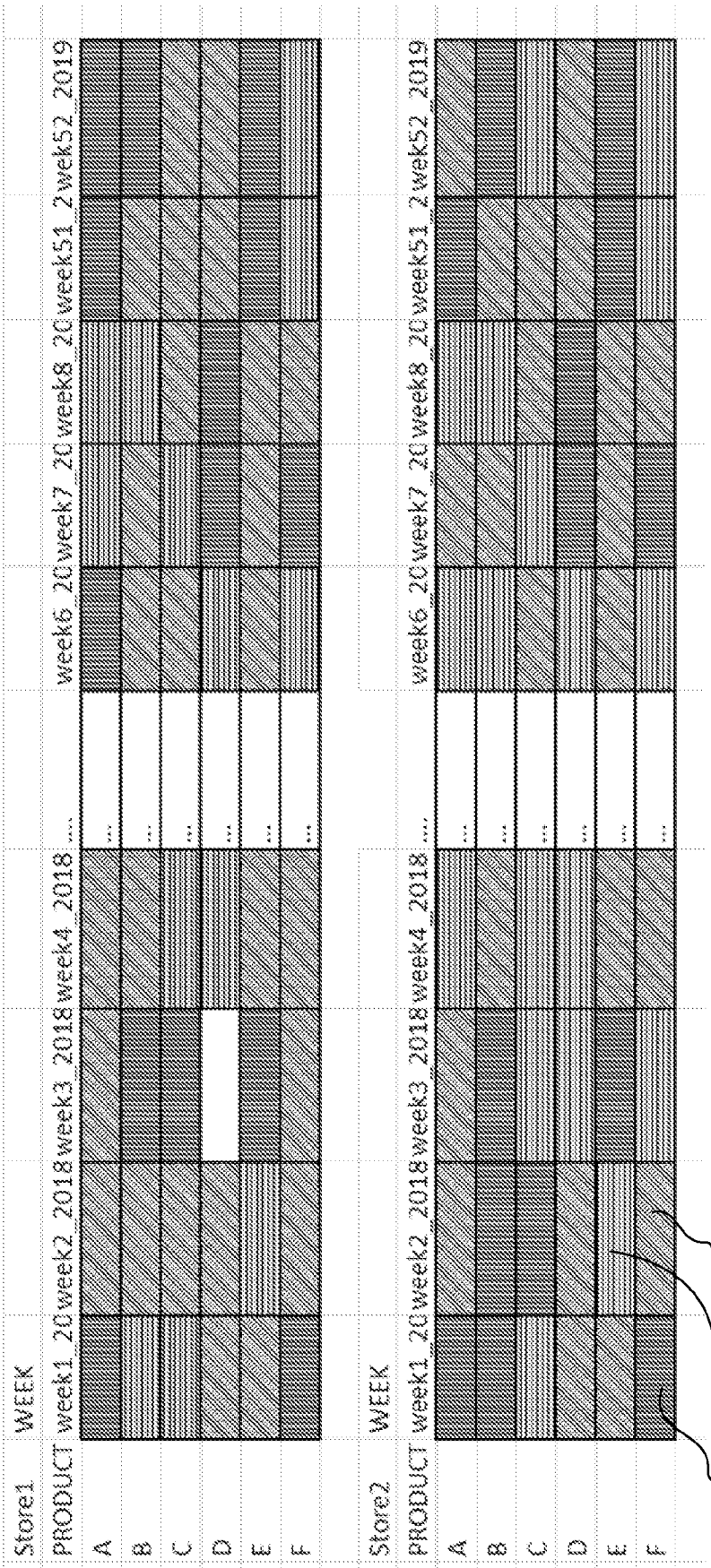

Store1 WEEK
PRODUCT week1_20 week2_2018 week3_2018 week4_2018 .... week6_20 week7_20 week8_20 week51_2 week52_2019
A
B
C
D
E
F
Store2 WEEK
PRODUCT week1_20 week2_2018 week3_2018 week4_2018 .... week6_20 week7_20 week8_20 week51_2 week52_2019
A
B
C
D
E
F
402
403
404

Fig. 4A

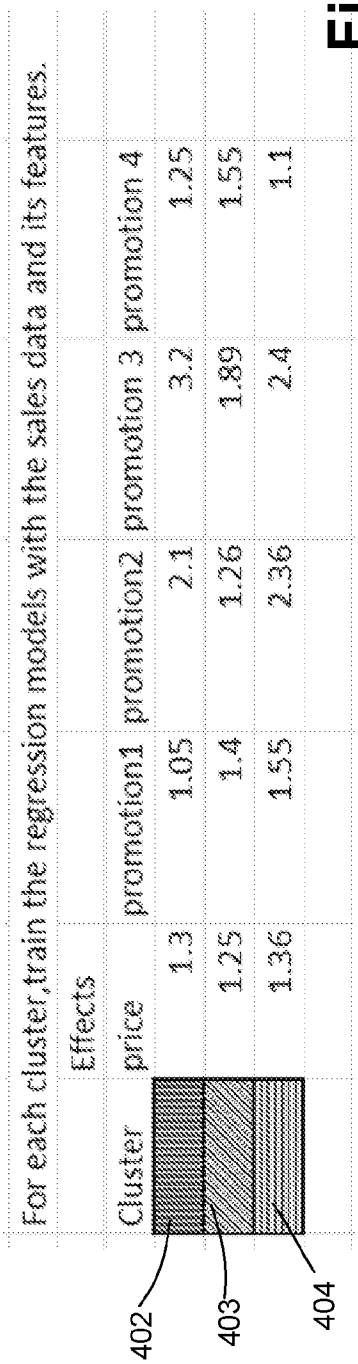

For each cluster,train the regression models with the sales data and its features.
Effects
Cluster price promotion1 promotion2 promotion 3 promotion 4
1.3 1.05 2.1 3.2 1.25
1.25 1.4 1.26 1.89 1.55
1.36 1.55 2.36 2.4 1.1
402
403
404

Fig. 4B

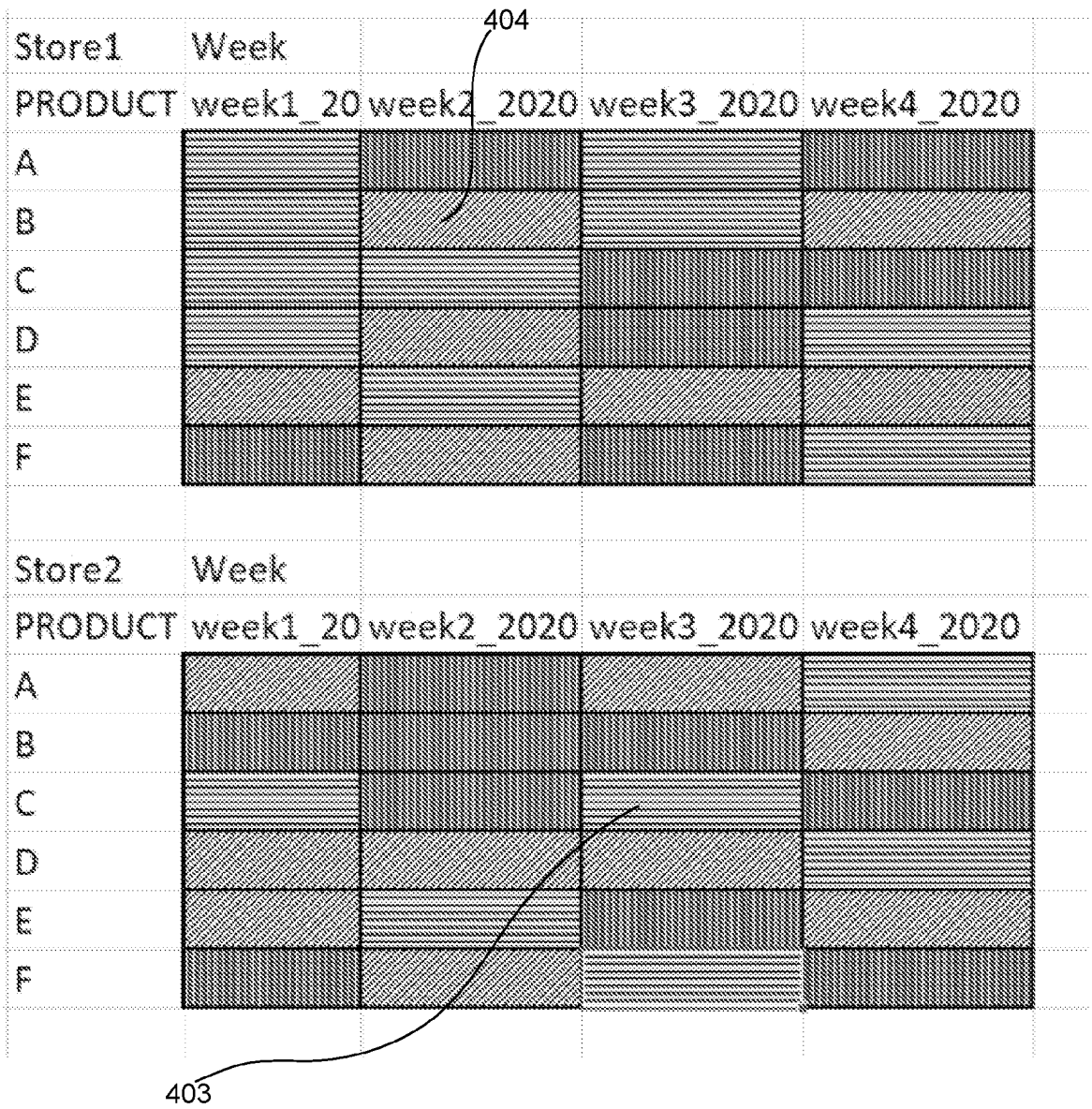

404
Store1
Week
PRODUCT
week1_20
week2_2020
week3_2020
week4_2020
A
B
C
D
E
F
Store2
Week
PRODUCT
week1_20
week2_2020
week3_2020
week4_2020
A
B
C
D
E
F
403

Fig. 5

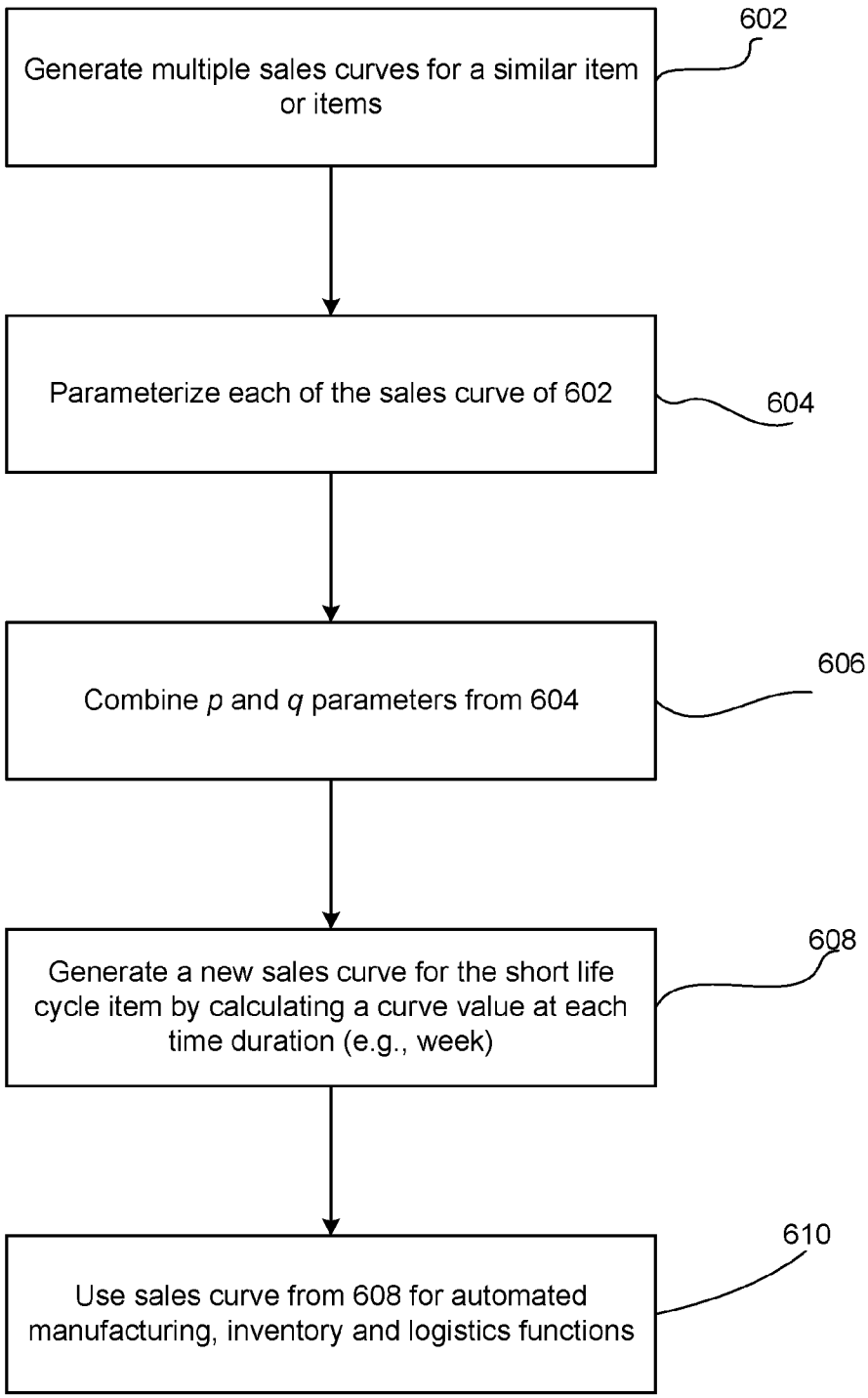

Generate multiple sales curves for a similar item or items
602
Parameterize each of the sales curve of 602
604
Combine p and q parameters from 604
606
Generate a new sales curve for the short life cycle item by calculating a curve value at each time duration (e.g., week)
608
Use sales curve from 608 for automated manufacturing, inventory and logistics functions
610

Fig. 6

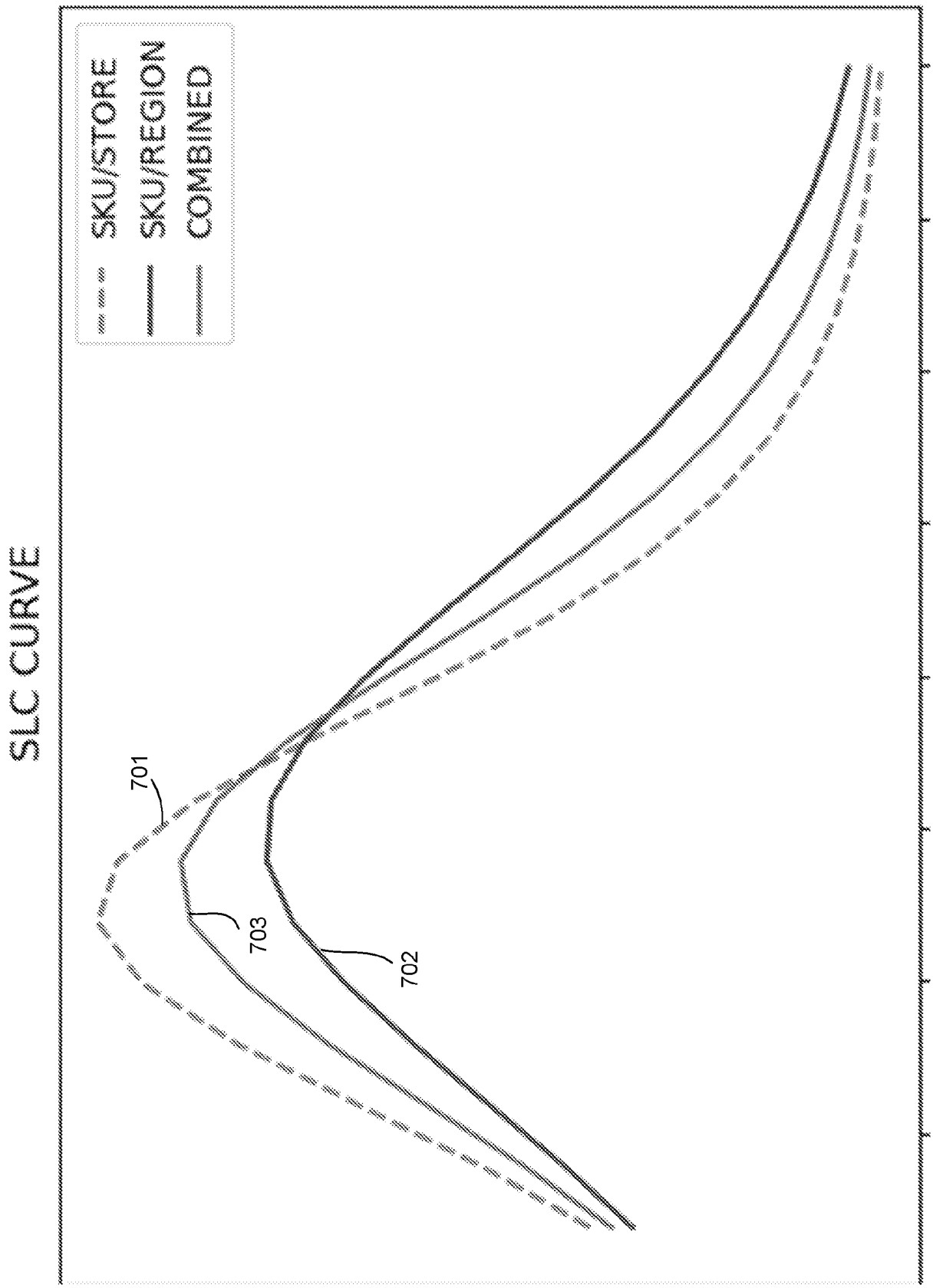


Fig. 7

801

Intersection	Pi	Qi	Mi	Ei	Wi
sku/store	0.03	0.38	100	12.5	0.51867
sku/region	0.028	0.29	110	11.6	0.48133

802

	Pi	Qi
Combined	0.02904	0.33668

Fig. 8

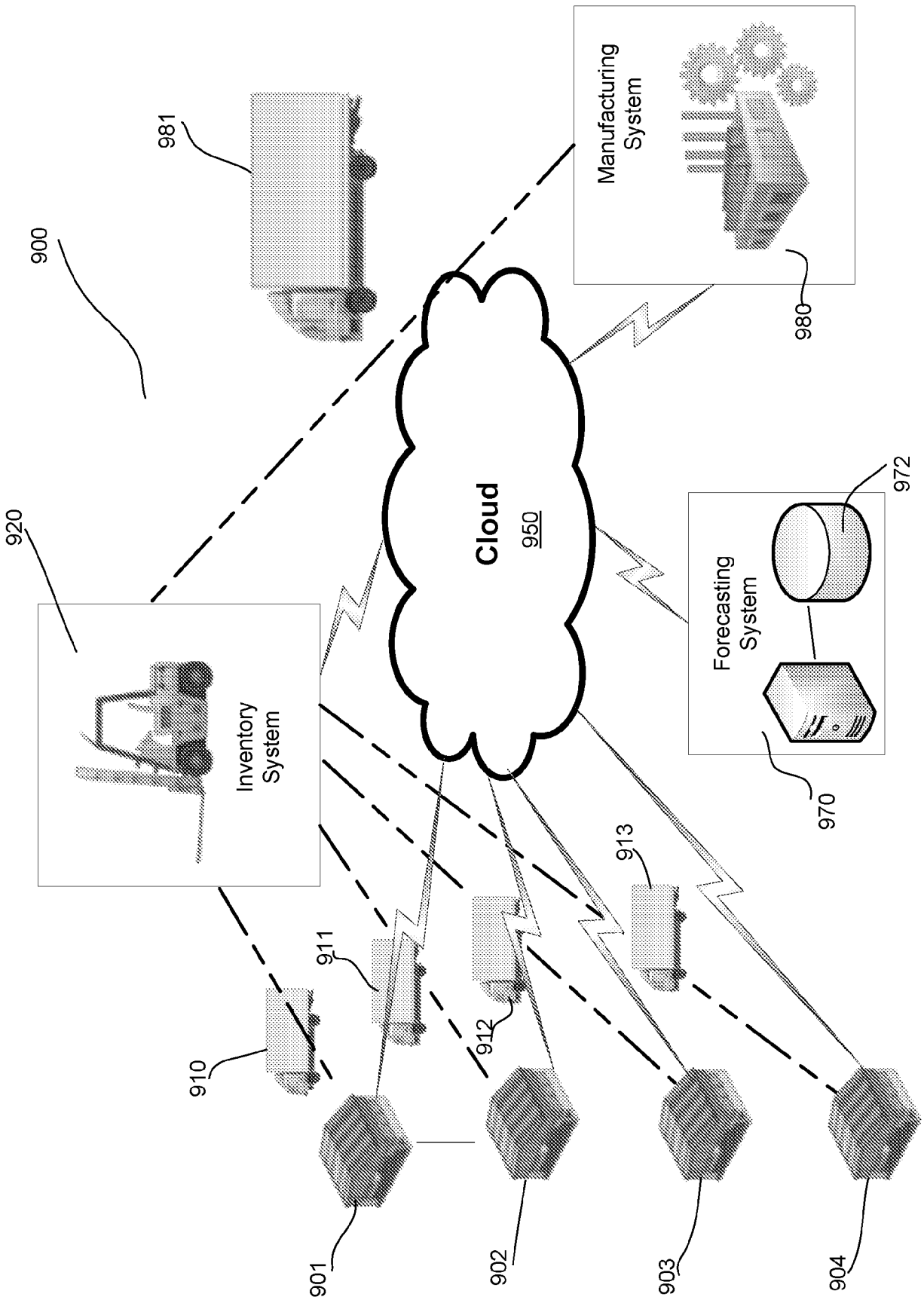

900
981
Manufacturing System
980
Cloud
950
920
Inventory System
Forecasting System
972
970
913
911
912
910
901
902
903
904

Fig. 9

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2021/041773

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06Q30/02
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06Q

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2020/074485 A1 (POPESCU CATALIN [US] ET AL) 5 March 2020 (2020-03-05) abstract paragraph [0001] - paragraph [0010] paragraph [0020] - paragraph [0025] paragraph [0033] - paragraph [0125] claims 1-20; figures 1-8	1-20
X	US 2019/130425 A1 (LEI MING [US] ET AL) 2 May 2019 (2019-05-02) abstract paragraph [0001] - paragraph [0006] paragraph [0011] - paragraph [0063] claims 1-20; figures 1-5	1-20

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 November 2021	22/11/2021

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Bassanini, Anna

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2021/041773

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2020074485 A1	05-03-2020	CN 111133460 A	08-05-2020
		US 2020074485 A1	05-03-2020
		WO 2020046497 A1	05-03-2020
US 2019130425 A1	02-05-2019	CN 111295681 A	16-06-2020
		JP 2021501421 A	14-01-2021
		US 2019130425 A1	02-05-2019
		WO 2019089146 A1	09-05-2019