



US007113917B2

(12) **United States Patent**  
**Jacobi et al.**

(10) **Patent No.:** **US 7,113,917 B2**  
(45) **Date of Patent:** **\*Sep. 26, 2006**

(54) **PERSONALIZED RECOMMENDATIONS OF  
ITEMS REPRESENTED WITHIN A  
DATABASE**

(75) Inventors: **Jennifer A. Jacobi**, Seattle, WA (US);  
**Eric A. Benson**, Seattle, WA (US);  
**Gregory D. Linden**, Seattle, WA (US)

(73) Assignee: **Amazon.com, Inc.**, Seattle, WA (US)

(\*) Notice: Subject to any disclaimer, the term of this  
patent is extended or adjusted under 35  
U.S.C. 154(b) by 499 days.

This patent is subject to a terminal dis-  
claimer.

(21) Appl. No.: **09/850,263**

(22) Filed: **May 7, 2001**

(65) **Prior Publication Data**

US 2001/0021914 A1 Sep. 13, 2001

**Related U.S. Application Data**

(63) Continuation of application No. 09/156,237, filed on  
Sep. 18, 1998, now Pat. No. 6,317,722.

(51) **Int. Cl.**  
**G06F 17/60** (2006.01)

(52) **U.S. Cl.** ..... **705/14; 705/10; 705/26**

(58) **Field of Classification Search** ..... **705/14,**  
**705/10, 26**

See application file for complete search history.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

4,870,579 A	9/1989	Hey	364/419
4,992,940 A	2/1991	Dworkin	705/26
4,996,642 A	2/1991	Hey	364/419
5,235,509 A	8/1993	Mueller et al.	364/405
5,459,306 A	10/1995	Stein et al.	235/383
5,583,763 A	12/1996	Atcheson et al.	364/551.01

5,704,017 A	12/1997	Heckerman et al.	706/12
5,732,398 A *	3/1998	Tagawa	705/5
5,745,681 A	4/1998	Levine et al.	395/200.3
5,749,081 A	5/1998	Whiteis	707/102
5,758,257 A	5/1998	Herz et al.	
5,774,670 A	6/1998	Montulli	395/200.57
5,790,426 A	8/1998	Robinson	702/179
5,790,935 A	8/1998	Payton	455/5.1

(Continued)

**FOREIGN PATENT DOCUMENTS**

EP 0 265 083 4/1988

(Continued)

**OTHER PUBLICATIONS**

"A. 'Obvious to Try' Motivation Insufficient", JTPOS, v. 86, n. 10,  
Oct. 2004: 838.\*

(Continued)

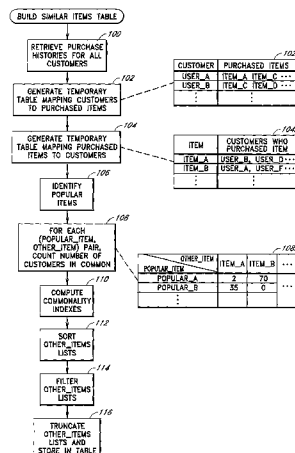
*Primary Examiner*—Donald L. Champagne

(74) *Attorney, Agent, or Firm*—Knobbe, Martens, Olson &  
Bear LLP

(57) **ABSTRACT**

A computer-implemented service recommends items to a user based on items previously selected by the user, such as items previously purchased, viewed, or placed in an electronic shopping cart by the user. The items may, for example, be products represented within a database of an online merchant. In one embodiment, the service generates the recommendations using a previously generated table that maps items to respective lists of "similar" items. To generate the table, historical data indicative of users' affinities for particular items is processed periodically to identify correlations between item interests of users (e.g., items A and B are similar because a large portion of those who selected A also selected B). Personal recommendations are generated by accessing the table to identify items similar to those selected by the user. In one embodiment, items are recommended based on the current contents of a user's shopping cart.

**52 Claims, 7 Drawing Sheets**



## U.S. PATENT DOCUMENTS

5,794,210	A	8/1998	Goldhaber et al.	705/14
5,867,799	A	2/1999	Lang et al.	707/1
5,884,282	A	3/1999	Robinson	705/27
5,905,973	A	5/1999	Yonezawa et al.	705/27
5,909,023	A	6/1999	Ono et al.	235/380
5,909,492	A	6/1999	Payne et al.	380/24
5,918,014	A	6/1999	Robinson	709/219
5,974,396	A	10/1999	Anderson et al.	
6,006,218	A	12/1999	Breese et al.	707/3
6,006,225	A *	12/1999	Bowman et al.	707/5
6,018,738	A	1/2000	Breese et al.	707/100
6,195,657	B1 *	2/2001	Rucker et al.	707/5
6,266,649	B1 *	7/2001	Linden et al.	705/26
6,317,722	B1 *	11/2001	Jacobi et al.	705/14
6,321,221	B1 *	11/2001	Bieganski	705/7
6,460,036	B1	10/2002	Herz	
6,484,149	B1	11/2002	Jammes et al.	
6,629,079	B1	9/2003	Spiegel et al.	
6,782,370	B1	8/2004	Stack	

## FOREIGN PATENT DOCUMENTS

EP	0 751 471	A	1/1997
EP	0 827 063	A	3/1998
GB	2 336 925	A	3/1999

## OTHER PUBLICATIONS

Upendra Shardanand and Pattie Maes with MIT Media-Lab, Social Information Filtering: Algorithms for Automating "Word of Mouth", 8 pgs (undated).

Combining Social Networks and Collaborative Filtering, Communications of the ACM, Mar. 1997/vol. 40, No. 3, pp. 63-65.

Pointing the Way: Active Collaborative Filtering, CHI '95 Proceedings Papers, 11 pgs.

Bradley N. Miller, John T. Riedl, Joseph A. Konstan with Department of Computer Science, University of Minnesota, Experiences with GroupLens: Making Usenet Useful Again, 13 pgs.

A System for Sharing Recommendations, Communications of the ACM, Mar. 1997/vol. 40, No. 3, pp. 59-62.

Recommender Systems for Evaluating Computer Messages, Communications of the ACM, Mar. 1997/vol. 40, No. 3, pp. 88 and 89.

Content-Based, Collaborative Recommendation, Communications of the ACM, Mar. 1997/vol. 40, No. 3, pp. 66-72.

Applying Collaborative Filtering to Usenet News, Communications of the ACM, Mar. 1997/vol. 40, No. 3, pp. 77-87.

Personalized Navigation for the Web, Communications of the ACM, Mar. 1997/vol. 40, No. 3, pp. 73-76.

Resnick, P., et al. GroupLens: An Open Architecture for Collaborative Filtering of Netnews, Proceedings of ACM 1994 Conf., Chapel Hill, NC, pp. 175-186.

Net Perceptions, Inc., White Paper, Building Customer Loyalty and High-Yield Relationships Through GroupLens Collaborative Filtering, 9 pgs., Nov. 22, 1996.

Christos Faloutsos and Douglas Oard with University of Maryland, A Survey of Information Retrieval and Filtering Methods, 22 pgs (undated).

Joaquin Delgado, "Intelligence Collaborative Information Retrieval", undated.

Joaquin Delgado, "Content-based Collaborative Information Filtering", undated.

Marko Balabanovic and Yoav Shoham, "Content-Based, Collaborative Recommendation," Communications of the ACM, vol. 40n3, pp. 66-72, Mar. 1997.

"COSMOCOM," Computer Telephony, p. 124, Jul. 1998.

Brier, S.E., "Smart Devices Peep Into Your Grocery Cart," New York Times Co., Section G, p. 3, col. 3, Circuits, Jul. 16, 1998.

Nash, E.L., "Direct Marketing; Strategy, Planning, Execution," 3<sup>rd</sup> Ed., McGraw-Hill, Inc., pp. 165, & 365-6. (1994).

"iCatElectronic Commerce Suite Takes 'Best of Show' Award At WebINNOVATION 97," PR Newswire, Jun. 1997.

"iCatCorporation: iCat's Commerce Suite Makes Setting Up Shop on Net Even Easier Than High Street," M2 Presswire, Feb. 1997.

Dragan et al., "Advice From the Web," PC Magazine, v. 16, n. 15, p. 133, Sep. 1997.

"Able Solutions Announces Able Commerce 2.6," PR Newswire, Sep. 1998.

"Internet World—IBM To Expand E-Comm Features," Newsbytes News Network, Dec. 1996.

McMains, A., "Weiss, Whitten, Staliano's," ADWEEK Eastern Edition, V. 39, N. 24, p. 82, Jun. 1998.

"Cdnw Rated Top Music Site by eMarketer, the Authority on Business Online," PR Newswire, Sep. 1998.

"Net Perceptions Closes Second Round of Financing: GroupLens secures No. 1 recommendation system spot with strong endorsement by investment community," Business Wire, p. 3020013, Dialog File 16, AN 05495619, Mar. 1998.

"LinkShare Launches Affiliates Profiling Software; First to Integrate Personalization Software Into Affiliates Program," PR Newswire, LinkShare Corp., Dialog File 813, AN 1232636, Feb. 1998.

"Fort Point Partners Teams With LikeMinds to Offer Breakthrough Personalization Technology for Increased Sales Online," Business Wire, p. 3110064, Dialog File 16, AN 05510541, Mar. 1998.

"Net Perceptions Debuts GroupLens Version 3.0 at Internet World Spring; 'Industrial Strength Tool Matures Into Essential Website Technology'," Business Wire, p. 3090007, Dialog File 16, AN 05505690, Mar. 1998.

"Home Box Office Selects Like Minds Personalization Software for Second Network Site," PR Newswire, p. 1117SFM023, Dialog File 148, AN 09869396, Nov. 1997.

"GroupLens Recommendation Engine to Standarize Internet Personalization For Singapore's Online Technologies Consortium," Business Wire, Dialog File 20, AN 01951318, Jun. 1998.

Borchers, A. et al., "Ganging up on Information Overload," Computer, pp. 106-108, Apr. 1998.

Konstan, J. et al., "GroupLens: Applying Collaborative Filtering to Usenet News," Communications of the ACM, vol. 30, No. 3, pp. 77-87, dated Mar. 1997.

Miller, B. et al., "Experiences with GroupLens: Making Usenet Useful Again." USENIX Association 1997 Annual Technical Conference, pp. 219-232.

Resnick, P. et al., "Recommender Systems," Communications of the ACM, vol. 40, No. 3, pp. 56-58, Mar. 1997.

Rucker, J. et al., "Seseeer: Personalized Navigation for the Web," Communications of the ACM, vol. 40, No. 3, pp. 73-76, Mar. 1997.

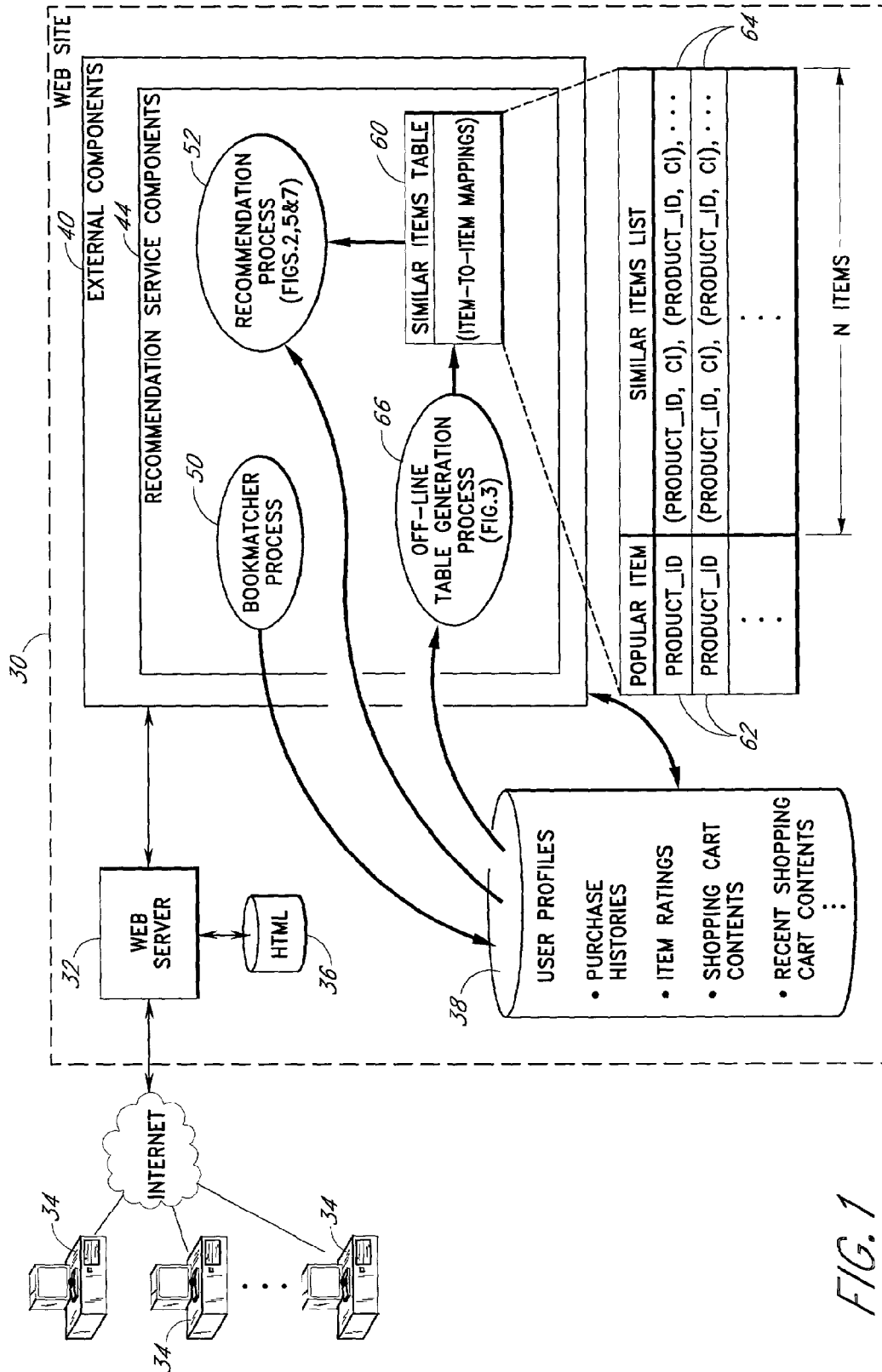
Description of Amazon.com's "book information page" feature, Paper No. 2, U.S. Appl. No. 09/850,263.

Watt, P., "Varied Collaborative Approaches Entice Intranet Users," Network World, vol. 14, Issue 16, p. S14, dated Apr. 21, 1997 (4 pages printed from ProQuest).

Hibbard, J., "Just Add People, Collaborative filtering brings human input to information retrieval in the enterprise," Dec. 22, 1997 issue of Information Week, p. 65.

R. Agrawal et al., "Mining association rules between sets of items in large databases," Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, pp. 207-216, 1993.

\* cited by examiner



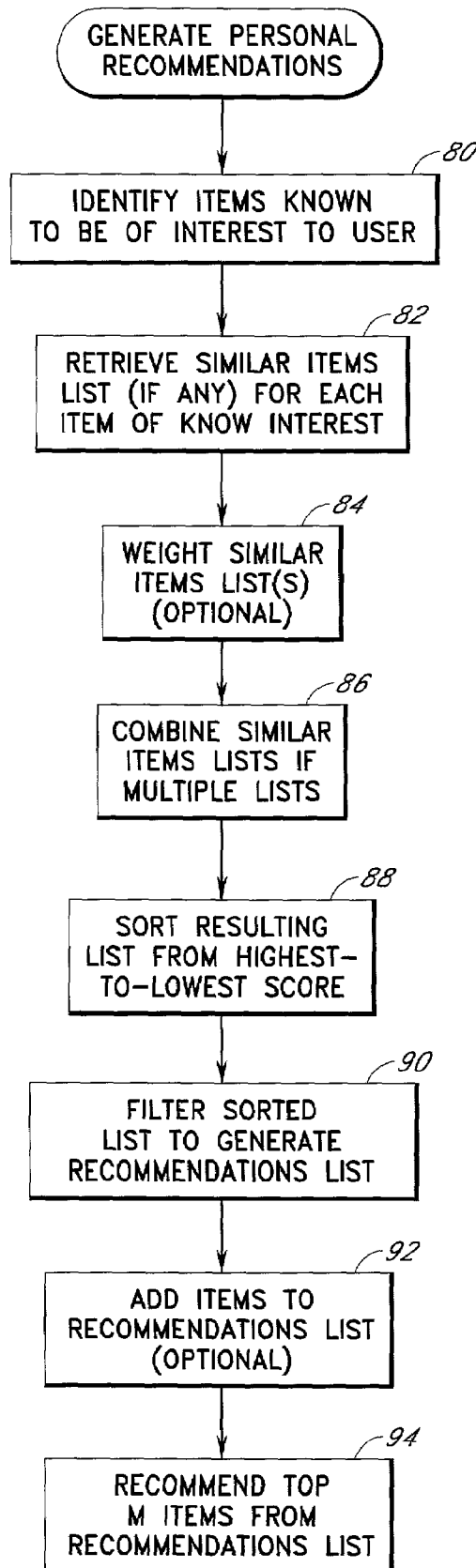


FIG. 2

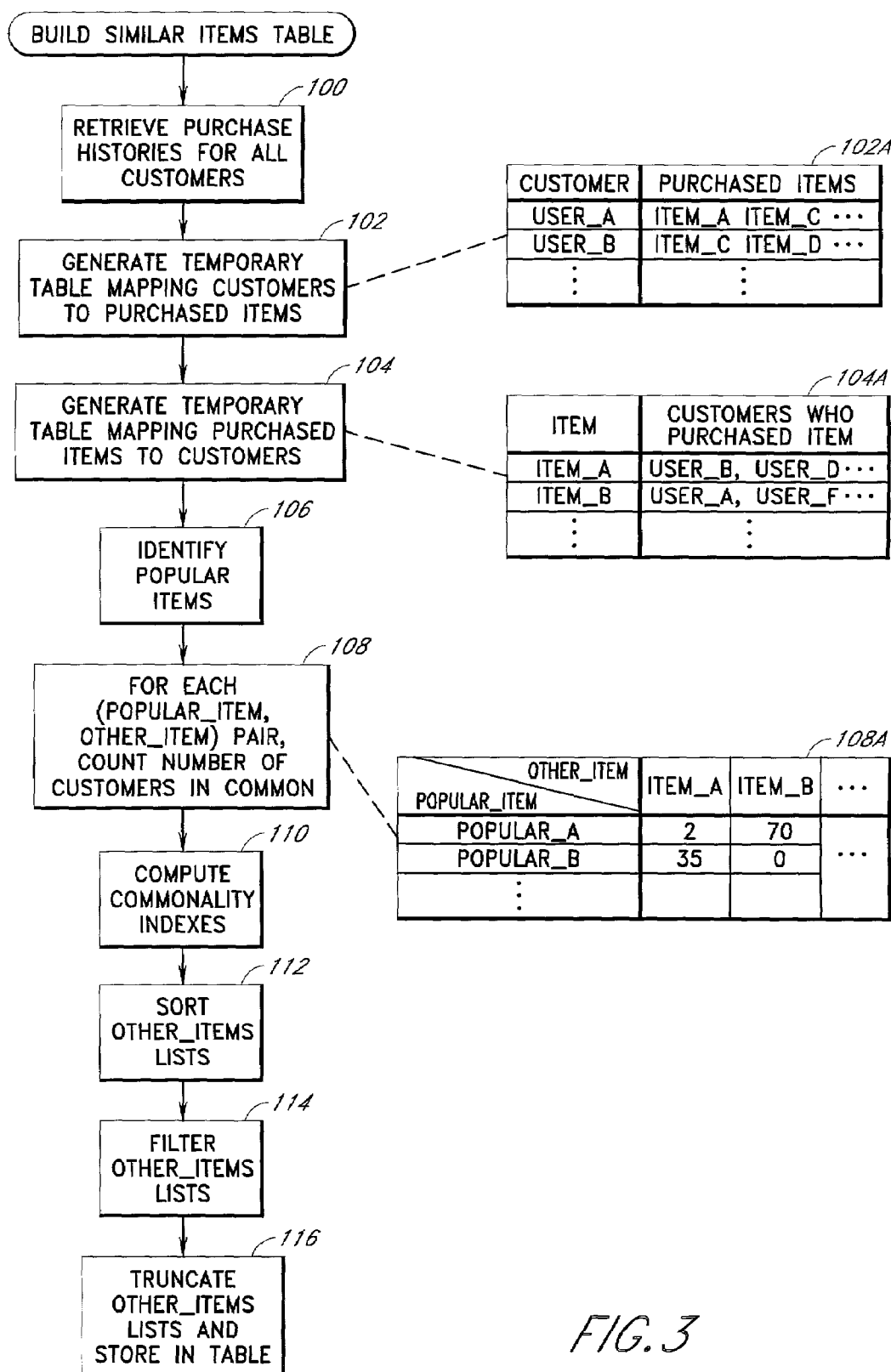
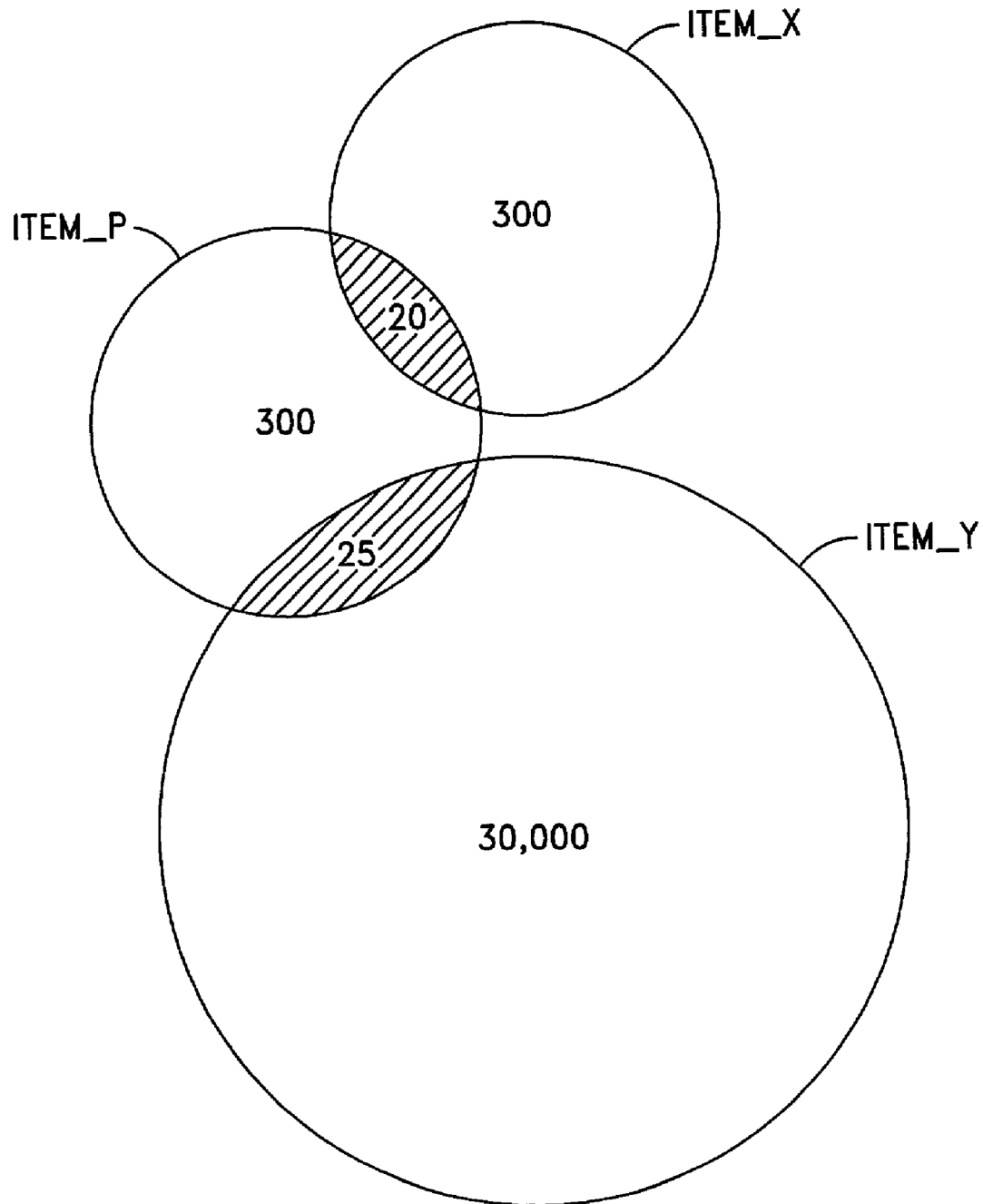


FIG. 3

*FIG. 4*

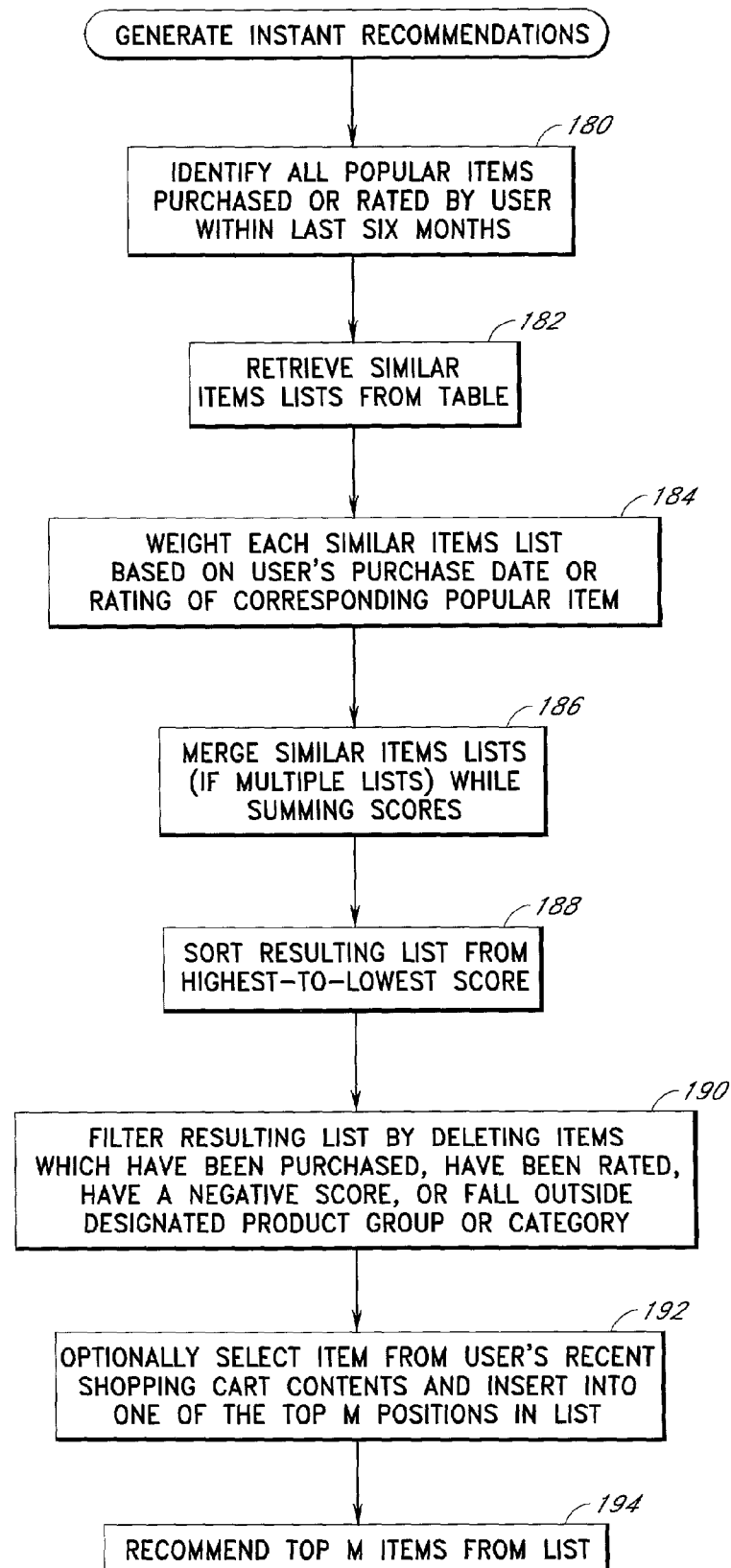


FIG. 5

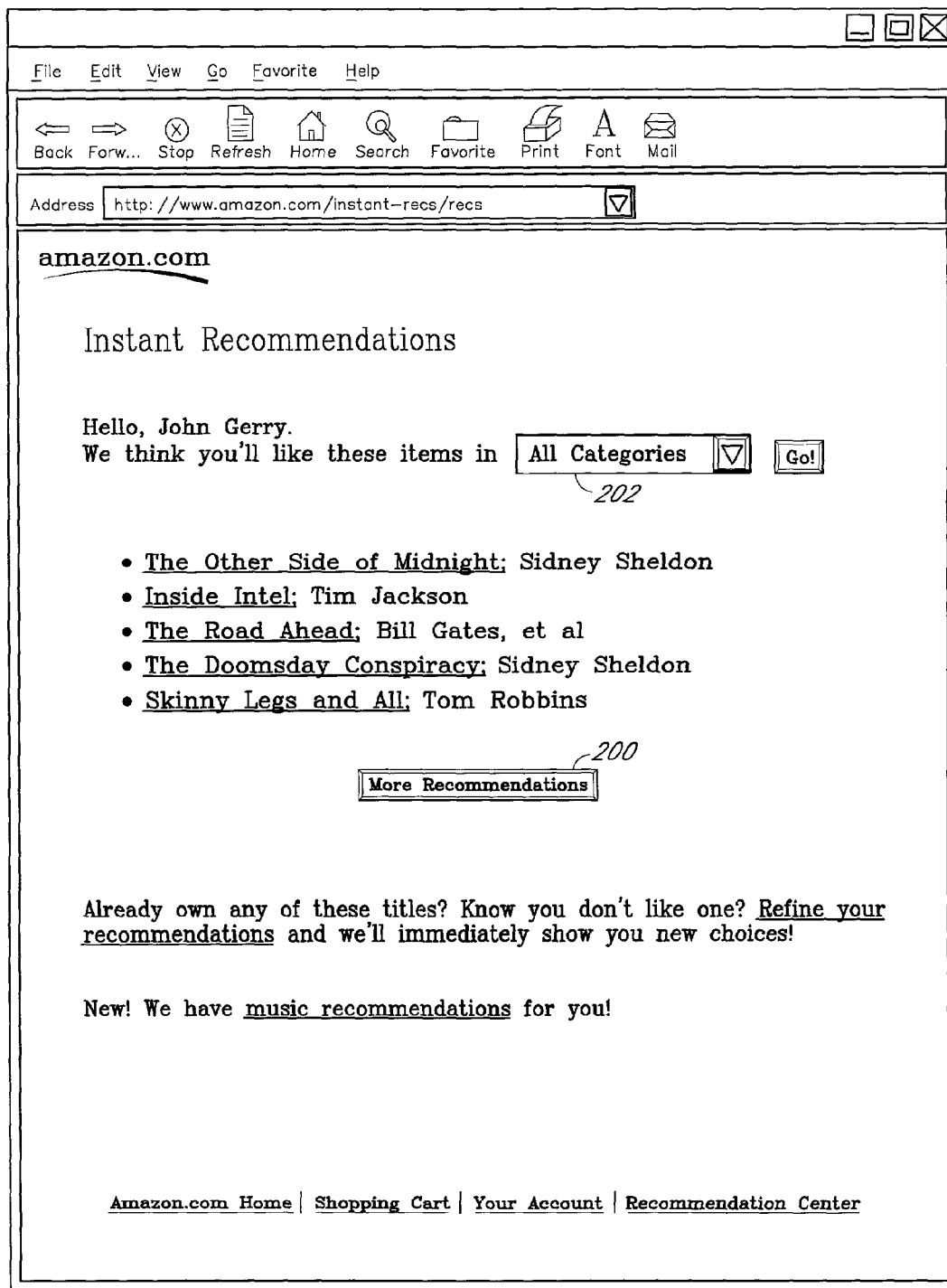
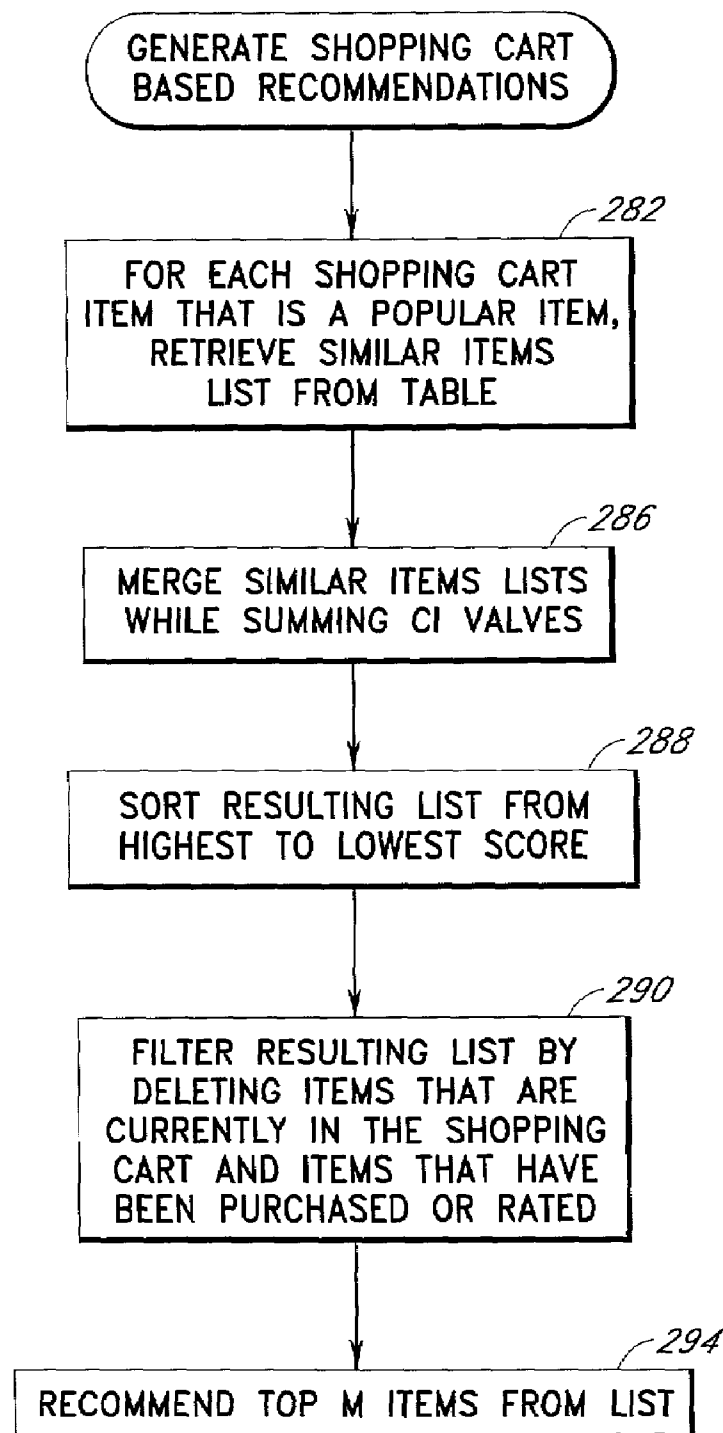


FIG. 6



*FIG. 7*

1

# PERSONALIZED RECOMMENDATIONS OF ITEMS REPRESENTED WITHIN A DATABASE

## RELATED APPLICATIONS

This application is a continuation of U.S. application Ser. No. 09/156,237, filed Sep. 18, 1998, now U.S. No. 6,317, 722.

## FIELD OF THE INVENTION

The present invention relates to information filtering and recommendation systems. More specifically, the invention relates to methods for predicting the interests of individual users based on the known interests of a community of users.

## BACKGROUND OF THE INVENTION

A recommendation service is a computer-implemented service that recommends items from a database of items. The recommendations are customized to particular users based on information known about the users. One common application for recommendation services involves recommending products to online customers. For example, online merchants commonly provide services for recommending products (books, compact discs, videos, etc.) to customers based on profiles that have been developed for such customers. Recommendation services are also common for recommending Web sites, articles, and other types of informational content to users.

One technique commonly used by recommendation services is known as content-based filtering. Pure content-based systems operate by attempting to identify items which, based on an analysis of item content, are similar to items that are known to be of interest to the user. For example, a content-based Web site recommendation service may operate by parsing the user's favorite Web pages to generate a profile of commonly-occurring terms, and then use this profile to search for other Web pages that include some or all of these terms.

Content-based systems have several significant limitations. For example, content-based methods generally do not provide any mechanism for evaluating the quality or popularity of an item. In addition, content-based methods generally require that the items include some form of content that is amenable to feature extraction algorithms; as a result, content-based systems tend to be poorly suited for recommending movies, music titles, authors, restaurants, and other types of items that have little or no useful, parsable content.

Another common recommendation technique is known as collaborative filtering. In a pure collaborative system, items are recommended to users based on the interests of a community of users, without any analysis of item content. Collaborative systems commonly operate by having the users rate individual items from a list of popular items. Through this process, each user builds a personal profile of ratings data. To generate recommendations for a particular user, the user's profile is initially compared to the profiles of other users to identify one or more "similar users." Items that were rated highly by these similar users (but which have not yet been rated by the user) are then recommended to the user. An important benefit of collaborative filtering is that it overcomes the above-noted deficiencies of content-based filtering.

As with content-based filtering methods, however, existing collaborative filtering techniques have several problems.

2

One problem is that the user is commonly faced with the onerous task of having to rate items in the database to build up a personal ratings profile. This task can be frustrating, particularly if the user is not familiar with many of the items that are presented for rating purposes. Further, because collaborative filtering relies on the existence of other, similar users, collaborative systems tend to be poorly suited for providing recommendations to users that have unusual tastes.

Another problem with collaborative filtering techniques is that an item in the database normally cannot be recommended until the item has been rated. As a result, the operator of a new collaborative recommendation system is commonly faced with a "cold start" problem in which the service cannot be brought online in a useful form until a threshold quantity of ratings data has been collected. In addition, even after the service has been brought online, it may take months or years before a significant quantity of the database items can be recommended.

Another problem with collaborative filtering methods is that the task of comparing user profiles tends to be time consuming—particularly if the number of users is large (e.g., tens or hundreds of thousands). As a result, a tradeoff tends to exist between response time and breadth of analysis. For example, in a recommendation system that generates real-time recommendations in response to requests from users, it may not be feasible to compare the user's ratings profile to those of all other users. A relatively shallow analysis of the available data (leading to poor recommendations) may therefore be performed.

Another problem with both collaborative and content-based systems is that they generally do not reflect the current preferences of the community of users. In the context of a system that recommends products to customers, for example, there is typically no mechanism for favoring items that are currently "hot sellers." In addition, existing systems do not provide a mechanism for recognizing that the user may be searching for a particular type or category of item.

## SUMMARY OF THE DISCLOSURE

One aspect of the invention is a method of recommending items to users from a database of items. The method comprises maintaining item selection histories of each of a plurality of users of a server system that provides functionality for browsing and selecting items from an electronic catalog of items, each item selection history corresponding to, and identifying items selected by, a particular user. Each item selection history may, for example, be a history of items selected for purchase by a user, or a history of items selected for viewing by a user. The method additionally comprises collectively analyzing at least the item selection histories of the plurality of users, as collected over a period of time, in an off-line processing mode to generate a plurality of data values that represent degrees to which specific items in the electronic catalog are related; and storing a selected subset of the plurality of data values in a mapping structure that maps items to related items. For each of a plurality of users of the electronic catalog, the mapping structure, including the data values stored therein, is used to generate personalized recommendations of items within the catalog.

Another aspect of the invention is a computer-implemented method of generating a mapping of items to related items. The method comprises maintaining item selection histories of each of a plurality of users of a server system that provides functionality for browsing and selecting items from an electronic catalog of items, each item selection history

3

corresponding to, and identifying items selected by, a particular user. The item selection histories of the plurality of users are collectively and programmatically analyzed to generate a data value that represents a degree to which a first item and a second item in the catalog are related. The data value is dependent upon at least (a) a number of users that selected both the first item and the second item, (b) a total number of users that selected the first item, and (c) a total number of users that selected the second item. The data value is used to determine whether the first item should be mapped to the second item in a mapping structure that maps items to related items.

Another aspect of the invention is a computer-implemented method of recommending items to users of a server system that provides functionality for selecting items represented in an electronic catalog. The method comprises identifying a plurality of items that are currently selected by a target user to obtain from an online business entity, including multiple items currently represented in a user-specific data repository in which item selections are maintained persistently over multiple sessions. The method additionally comprises identifying a set of additional items to recommend to the target user according to a selection algorithm in which a candidate item is considered for inclusion in the set based at least in-part upon a degree to which the candidate item is related to each of said plurality of items currently selected by the target user to obtain. The method further comprises recommending at least some of the additional items in the set to the target user while the plurality of items are selected by the target user to obtain.

Another aspect of the invention is a data mining method. The method comprises programmatically analyzing user activity data associated with a plurality of users of an electronic catalog, in an off-line processing mode, to generate data values representing degrees to which specific catalog items are related to one another. The user activity data reflects user interest in specific catalog items. The data values are used to select, for each of a plurality of the catalog items, a corresponding set of related catalog items. Mappings of the plurality of catalog items to the corresponding sets of related catalog items are generated, and are used to programmatically generate personalized item recommendations for each of a plurality of users of the electronic catalog.

### BRIEF DESCRIPTION OF THE DRAWINGS

These and other features of the invention will now be described with reference to the drawings summarized below. These drawings and the associated description are provided to illustrate a preferred embodiment of the invention, and not to limit the scope of the invention.

FIG. 1 illustrates a Web site which implements a recommendation service which operates in accordance with the invention, and illustrates the flow of information between components.

FIG. 2 illustrates a sequence of steps that are performed by the recommendation process of FIG. 1 to generate personalized recommendations.

FIG. 3 illustrates a sequence of steps that are performed by the table generation process of FIG. 1 to generate a similar items table, and illustrates temporary data structures generated during the process.

FIG. 4 is a Venn diagram illustrating a hypothetical purchase history profile of three items.

FIG. 5 illustrates one specific implementation of the sequence of steps of FIG. 2.

4

FIG. 6 illustrates the general form of a Web pages used to present the recommendations of the FIG. 5 process to the user.

FIG. 7 illustrates another specific implementation of the sequence of steps of FIG. 2.

### DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

The various features and methods of the invention will now be described in the context of a recommendation service, including two specific implementations thereof, that is used to recommend book titles, music titles, video titles, and other types of items to individual users of the Amazon.com Web site. As will be recognized to those skilled in the art, the disclosed methods can also be used to recommend other types of items, including non-physical items. By way of example and not limitation, the disclosed methods can also be used to recommend authors, artists, categories or groups of titles, Web sites, chat groups, movies, television shows, downloadable content, restaurants, and other users.

Throughout the description, reference will be made to various implementation-specific details of the recommendation service, the Amazon.com Web site, and other recommendation services of the Web site. These details are provided in order to fully illustrate preferred embodiments of the invention, and not to limit the scope of the invention. The scope of the invention is set forth in the appended claims.

#### I. Overview of Web Site and Recommendation Services

The Amazon.com Web site includes functionality for allowing users to search, browse, and make purchases from an online catalog of several million book titles, music titles, video titles, and other types of items. Using a shopping cart feature of the site, users can add and remove items to/from a personal shopping cart which is persistent over multiple sessions. (As used herein, a "shopping cart" is a data structure and associated code which keeps track of items that have been selected by a user for possible purchase.) For example, a user can modify the contents of the shopping cart over a period of time, such as one week, and then proceed to a check out area of the site to purchase the shopping cart contents.

The user can also create multiple shopping carts within a single account. For example, a user can set up separate shopping carts for work and home, or can set up separate shopping carts for each member of the user's family. A preferred shopping cart scheme for allowing users to set up and use multiple shopping carts is disclosed in U.S. application Ser. No. 09/104,942, filed Jun. 25, 1998, titled METHOD AND SYSTEM FOR ELECTRONIC COMMERCE USING MULTIPLE ROLES, the disclosure of which is hereby incorporated by reference.

The site also implements a variety of different recommendation services for recommending book titles, music titles, and/or video titles to users. One such service, known as BookMatcher™, allows users to interactively rate individual books on a scale of 1-5 to create personal item ratings profiles, and applies collaborative filtering techniques to these profiles to generate personal recommendations. The BookMatcher service is described in detail in U.S. application Ser. No. 09/040,171 filed Mar. 17, 1998, the disclosure of which is hereby incorporated by reference. The site may also include associated services that allow users to rate other types of items, such as CDs and videos. As described below, the ratings data collected by the BookMatcher service and

5

similar services is optionally incorporated into the recommendation processes of the present invention.

Another type of service is a recommendation service which operates in accordance with the invention. The service ("Recommendation Service") is preferably used to recommend book titles, music titles and/or videos titles to users, but could also be used in the context of the same Web site to recommend other types of items, including authors, artists, and groups or categories of titles. Briefly, given a unary listing of items that are "known" to be of interest to a user (e.g., a list of items purchased, rated, and/or viewed by the user), the Recommendation Service generates a list of additional items ("recommendations") that are predicted to be of interest to the user. (As used herein, the term "interest" refers generally to a user's liking of or affinity for an item; the term "known" is used to distinguish items for which the user has implicitly or explicitly indicated some level of interest from items predicted by the Recommendation Service to be of interest.)

The recommendations are generated using a table which maps items to lists of "similar" items ("similar items lists"), without the need for users to rate any items (although ratings data may optionally be used). For example, if there are three items that are known to be of interest to a particular user (such as three items the user recently purchased), the service may retrieve the similar items lists for these three items from the table, and appropriately combine these lists (as described below) to generate the recommendations.

In accordance with one aspect of the invention, the mappings of items to similar items ("item-to-item mappings") are generated periodically, such as once per week, from data which reflects the collective interests of the community of users. More specifically, the item-to-item mappings are generated by an off-line process which identifies correlations between known interests of users in particular items. For example, in the embodiment described in detail below, the mappings are generated by analyzing user purchase histories to identify correlations between purchases of particular items (e.g., items A and B are similar because a relatively large portion of the users that purchased item A also bought item B). The item-to-item mappings could also reflect other types of similarities, including content-based similarities extracted by analyzing item descriptions or content.

An important aspect of the Recommendation Service is that the relatively computation-intensive task of correlating item interests is performed off-line, and the results of this task (item-to-item mappings) are stored in a mapping structure for subsequent look-up. This enables the personal recommendations to be generated rapidly and efficiently (such as in real-time in response to a request by the user), without sacrificing breadth of analysis.

In accordance with another aspect of the invention, the similar items lists read from the table are appropriately weighted (prior to being combined) based on indicia of the user's affinity for or current interest in the corresponding items of known interest. For example, in one embodiment described below, if the item of known interest was previously rated by the user (such as through use of the Book-Matcher service), the rating is used to weight the corresponding similar items list. Similarly, the similar items list for a book that was purchased in the last week may be weighted more heavily than the similar items list for a book that was purchased four months ago.

Another feature of the invention involves using the current and/or recent contents of the user's shopping cart as inputs to the Recommendation Service. For example, if the

6

user currently has three items in his or her shopping cart, these three items can be treated as the items of known interest for purposes of generating recommendations, in which case the recommendations may be generated and displayed automatically when the user views the shopping cart contents. If the user has multiple shopping carts, the recommendations are preferably generated based on the contents of the shopping cart implicitly or explicitly designated by the user, such as the shopping cart currently being viewed. This method of generating recommendations can also be used within other types of recommendation systems, including content-based systems and systems that do not use item-to-item mappings.

Using the current and/or recent shopping cart contents as inputs tends to produce recommendations that are highly correlated to the current short-term interests of the user—even if these short term interests are not reflected by the user's purchase history. For example, if the user is currently searching for a father's day gift and has selected several books for prospective purchase, this method will have a tendency to identify other books that are well suited for the gift recipient.

Another feature of the invention involves generating recommendations that are specific to a particular shopping cart. This allows a user who has created multiple shopping carts to conveniently obtain recommendations that are specific to the role or purpose to the particular cart. For example, a user who has created a personal shopping cart for buying books for her children can designate this shopping cart to obtain recommendations of children's books. In one embodiment of this feature, the recommendations are generated based solely upon the current contents of the shopping cart selected for display. In another embodiment, the user may designate one or more shopping carts to be used to generate the recommendations, and the service then uses the items that were purchased from these shopping carts as the items of known interest.

As will be recognized by those skilled in the art, the above-described techniques for using shopping cart contents to generate recommendations can also be incorporated into other types of recommendation systems, including pure content-based systems.

FIG. 1 illustrates the basic components of the Amazon.com Web site 30, including the components used to implement the Recommendation Service. The arrows in FIG. 1 show the general flow of information that is used by the Recommendation Service. As illustrated by FIG. 1, the Web site 30 includes a Web server application 32 ("Web server") which processes HTTP (Hypertext Transfer Protocol) requests received over the Internet from user computers 34. The Web server 34 accesses a database 36 of HTML (Hypertext Markup Language) content which includes product information pages and other browsable information about the various products of the catalog. The "items" that are the subject of the Recommendation Service are the titles (regardless of media format such as hardcover or paperback) that are represented within this database 36.

The Web site 30 also includes a "user profiles" database 38 which stores account-specific information about users of the site. Because a group of individuals can share an account, a given "user" from the perspective of the Web site may include multiple actual users. As illustrated by FIG. 1, the data stored for each user may include one or more of the following types of information (among other things) that can be used to generate recommendations in accordance with the invention: (a) the user's purchase history, including dates of purchase, (b) the user's item ratings profile (if any), (c) the

current contents of the user's personal shopping cart(s), and (d) a listing of items that were recently (e.g., within the last six months) removed from the shopping cart(s) without being purchased ("recent shopping cart contents"). If a given user has multiple shopping carts, the purchase history for that user may include information about the particular shopping cart used to make each purchase; preserving such information allows the Recommendation Service to be configured to generate recommendations that are specific to a particular shopping cart.

As depicted by FIG. 1, the Web server 32 communicates with various external components 40 of the site. These external components 40 include, for example, a search engine and associated database (not shown) for enabling users to interactively search the catalog for particular items. Also included within the external components 40 are various order processing modules (not shown) for accepting and processing orders, and for updating the purchase histories of the users.

The external components 40 also include a shopping cart process (not shown) which adds and removes items from the users' personal shopping carts based on the actions of the respective users. (The term "process" is used herein to refer generally to one or more code modules that are executed by a computer system to perform a particular task or set of related tasks.) In one embodiment, the shopping cart process periodically "prunes" the personal shopping cart listings of items that are deemed to be dormant, such as items that have not been purchased or viewed by the particular user for a predetermined period of time (e.g. two weeks). The shopping cart process also preferably generates and maintains the user-specific listings of recent shopping cart contents.

The external components 40 also include recommendation service components 44 that are used to implement the site's various recommendation services. Recommendations generated by the recommendation services are returned to the Web server 32, which incorporates the recommendations into personalized Web pages transmitted to users.

The recommendation service components 44 include a BookMatcher application 50 which implements the above-described BookMatcher service. Users of the BookMatcher service are provided the opportunity to rate individual book titles from a list of popular titles. The book titles are rated according to the following scale:

---

1 = Bad!  
 2 = Not for me  
 3 = OK  
 4 = Liked it  
 5 = Loved it!

---

Users can also rate book titles during ordinary browsing of the site. As depicted in FIG. 1, the BookMatcher application 50 records the ratings within the user's items rating profile. For example, if a user of the BookMatcher service gives the book *Into Thin Air* a score of "5," the BookMatcher application 50 would record the item (by ISBN or other identifier) and the score within the user's item ratings profile. The BookMatcher application 50 uses the users' item ratings profiles to generate personal recommendations, which can be requested by the user by selecting an appropriate hyperlink. As described in detail below, the item ratings profiles are also used by an "Instant Recommendations" implementation of the Recommendation Service.

The recommendation services components 44 also include a recommendation process 52, a similar items table 60, and an off-line table generation process 66, which collectively implement the Recommendation Service. As depicted by the arrows in FIG. 1, the recommendation process 52 generates personal recommendations based on information stored within the similar items table 60, and based on the items that are known to be of interest ("items of known interest") to the particular user.

In the embodiments described in detail below, the items of known interest are identified based on information stored in the user's profile, such as by selecting all items purchased by the user or all items in the user's shopping cart. In other embodiments of the invention, other types of methods or sources of information could be used to identify the items of known interest. For example, in a service used to recommend Web sites, the items (Web sites) known to be of interest to a user could be identified by parsing a Web server access log and/or by extracting URLs from the "favorite places" list of the user's Web browser. In a service used to recommend restaurants, the items (restaurants) of known interest could be identified by parsing the user's credit card records to identify restaurants that were visited more than once.

The various processes 50, 52, 66 of the recommendation services may run, for example, on one or more Unix or NT based workstations or physical servers (not shown) of the Web site 30. The similar items table 60 is preferably stored as a B-tree data structure to permit efficient look-up, and may be replicated across multiple machines (together with the associated code of the recommendation process 52) to accommodate heavy loads.

## II. Similar Items Table (FIG. 1)

The general form and content of the similar items table 60 will now be described with reference to FIG. 1. As this table can take on many alternative forms, the details of the table are intended to illustrate, and not limit, the scope of the invention.

As indicated above, the similar items table 60 maps items to lists of similar items based at least upon the collective interests of the community of users. The similar items table 60 is preferably generated periodically (e.g., once per week) by the off-line table generation process 66. The table generation process 66 generates the table 60 from data that reflects the collective interests of the community of users. In the embodiment described in detail herein, the similar items table is generated exclusively from the purchase histories of the community of users (as depicted in FIG. 1). In other embodiments, the table 60 may additionally or alternatively be generated from other indicia of user-item interests, including indicia based on users viewing activities, shopping cart activities, and item rating profiles. For example, the table 60 could be built exclusively from the present and/or recent shopping cart contents of users. The similar items table 60 could also reflect non-collaborative type item similarities, including content-based similarities derived by comparing item contents or descriptions.

Each entry in the similar items table 60 is preferably in the form of a mapping of a popular item 62 to a corresponding list 64 of similar items ("similar items lists"). As used herein, a "popular" item is an item which satisfies some pre-specified popularity criteria. For example, in the embodiment described herein, an item is treated as popular if it has been purchased by more than 30 customers during the life of the Web site. Using this criteria produces a set of popular items (and thus a recommendation service) which grows

over time. The similar items list **64** for a given popular item **62** may include other popular items.

In other embodiments involving sales of products, the table **60** may include entries for most or all of the products of the online merchant, rather than just the popular items. In the embodiment described herein, several different types of items (books, CDs, videos, etc.) are reflected within the same table **60**, although separate tables could alternatively be generated for each type of item.

Each similar items list **64** consists of the N (e.g., 20) items which, based on correlations between purchases of items, are deemed to be the most closely related to the respective popular item **62**. Each item in the similar items list **64** is stored together with a commonality index ("CI") value which indicates the relatedness of that item to the popular item **62**, based on sales of the respective items. A relatively high commonality index for a pair of items ITEM A and ITEM B indicates that a relatively large percentage of users who bought ITEM A also bought ITEM B (and vice versa). A relatively low commonality index for ITEM A and ITEM B indicates that a relatively small percentage of the users who bought ITEM A also bought ITEM B (and vice versa). As described below, the similar items lists are generated, for each popular item, by selecting the N other items that have the highest commonality index values. Using this method, ITEM A may be included in ITEM B's similar items list even though ITEM B is not present in ITEM A's similar items list.

In the embodiment depicted by FIG. 1, the items are represented within the similar items table **60** using product IDs, such as ISBNs or other identifiers. Alternatively, the items could be represented within the table by title ID, where each title ID corresponds to a given "work" regardless of its media format. In either case, different items which correspond to the same work, such as the hardcover and paperback versions of a given book or the VCR cassette and DVD versions of a given video, are preferably treated as a unit for purposes of generating recommendations.

Although the recommendable items in the described system are in the form of book titles, music titles and videos titles, it will be appreciated that the underlying methods and data structures can be used to recommend a wide range of other types of items. For example, in the system depicted by FIG. 1, the Recommendation Service could also be used to recommend authors, artists, and categorizations or groups of works.

### III. General Process for Generating Recommendations (FIG. 2)

The general sequence of steps that are performed by the recommendation process **52** to generate a set of personal recommendations will now be described with reference to FIG. 2. This process, and the more specific implementations of the process depicted by FIGS. 5 and 7 (described below), are intended to illustrate, and not limit, the scope of the invention.

The FIG. 2 process is preferably invoked in real-time in response to an online action of the user. For example, in an Instant Recommendations implementation (FIGS. 5 and 6) of the service, the recommendations are generated and displayed in real-time (based on the user's purchase history and/or item ratings profile) in response to selection by the user of a corresponding hyperlink, such as a hyperlink which reads "Instant Book Recommendations" or "Instant Music Recommendations." In a shopping cart based implementation (FIG. 7), the recommendations are generated (based on the user's current and/or recent shopping cart contents) in

real-time when the user initiates a display of a shopping cart, and are displayed on the same Web page as the shopping cart contents. The Instant Recommendations and shopping cart based embodiments are described separately below under corresponding headings.

Any of a variety of other methods can be used to initiate the recommendations generation process and to display the recommendations to the user. For example, the recommendations can automatically be generated periodically and sent to the user by e-mail, in which case the e-mail listing may contain hyperlinks to the product information pages of the recommended items. Further, the personal recommendations could be generated in advance of any request or action by the user, and cached by the Web site **30** until requested.

As illustrated by FIG. 2, the first step (step **80**) of the recommendations-generation process involves identifying a set of items that are of known interest to the user. The "knowledge" of the user's interest can be based on explicit indications of interest (e.g., the user rated the item highly) or implicit indications of interest (e.g., the user added the item to a shopping cart). Items that are not "popular items" within the similar items table **60** can optionally be ignored during this step.

In the embodiment depicted in FIG. 1, the items of known interest are selected from one or more of the following groups: (a) items in the user's purchase history (optionally limited to those items purchased from a particular shopping cart); (b) items in the user's shopping cart (or a particular shopping cart designated by the user), (c) items rated by the user (optionally with a score that exceeds a certain threshold, such as two), and (d) items in the "recent shopping cart contents" list associated with a given user or shopping cart. In other embodiments, the items of known interest may additionally or alternatively be selected based on the viewing activities of the user. For example, the recommendations process **52** could select items that were viewed by the user for an extended period of time and/or viewed more than once. Further, the user could be prompted to select items of interest from a list of popular items.

For each item of known interest, the service retrieves the corresponding similar items list **64** from the similar items table **60** (step **82**), if such a list exists. If no entries exist in the table **60** for any of the items of known interest, the process **52** may be terminated; alternatively, the process could attempt to identify additional items of interest, such as by accessing other sources of interest information.

In step **84**, the similar items lists **64** are optionally weighted based on information about the user's affinity for the corresponding items of known interest. For example, a similar items list **64** may be weighted heavily if the user gave the corresponding popular item a rating of "5" on a scale of 1-5, or if the user purchased multiple copies of the item. Weighting a similar items list **64** heavily has the effect of increasing the likelihood that the items in that list will be included in the recommendations that are ultimately presented to the user. In one implementation described below, the user is presumed to have a greater affinity for recently purchased items over earlier purchased items.

The similar items lists **64** are preferably weighted by multiplying the commonality index values of the list by a weighting value. The commonality index values as weighted by any applicable weighting value are referred to herein as "scores." In other embodiments, the recommendations may be generated without weighting the similar items lists **64**.

If multiple similar items lists **64** are retrieved in step **82**, the lists are appropriately combined (step **86**), such as by merging the lists while summing the scores of like items.

11

The resulting list is then sorted (step 88) in order of highest-to-lowest score. In step 90, the sorted list is filtered to remove unwanted items. The items removed during the filtering process may include, for example, items that have already been purchased or rated by the user, and items that fall outside any product group (such as music or books), product category (such as non-fiction), or content rating (such as PG or adult) designated by the user. The filtering step could alternatively be performed at a different stage of the process, such as during the retrieval of the similar items lists from the table 60. The result of step 90 is a list ("recommendations list") of other items to be recommended to the user.

In step 92, one or more additional items are optionally added to the recommendations list. In one embodiment, the items added in step 92 are selected from the set of items (if any) in the user's "recent shopping cart contents" list. As an important benefit of this step, the recommendations include one or more items that the user previously considered purchasing but did not purchase. The items added in step 92 may additionally or alternatively be selected using another recommendations method, such as a content-based method.

Finally, in step 94, a list of the top M (e.g., 15) items of the recommendations list are returned to the Web server 32 (FIG. 1). The Web server incorporates this list into one or more Web pages that are returned to the user, with each recommended item being presented as a hypertextual link to the item's product information page. The recommendations may alternatively be conveyed to the user by email, facsimile, or other transmission method. Further, the recommendations could be presented as advertisements for the recommended items.

#### IV. Generation of Similar Items Table (FIGS. 3 and 4)

The table-generation process 66 is preferably executed periodically (e.g., once a week) to generate a similar items table 60 that reflects the most recent purchase history data. The recommendation process 52 uses the most recently generated version of the table 60 to generate recommendations.

FIG. 3 illustrates the sequence of steps that are performed by the table generation process 66 to build the similar items table 60. The general form of temporary data structures that are generated during the process are shown at the right of the drawing. As will be appreciated by those skilled in the art, any of a variety of alternative methods could be used to generate the table 60.

As depicted by FIG. 3, the process initially retrieves the purchase histories for all customers (step 100). Each purchase history is in the general form of the user ID of a customer together with a list of the product IDs (ISBNs, etc.) of the items (books, CDs, videos, etc.) purchased by that customer. In embodiments which support multiple shopping carts within a given account, each shopping cart could be treated as a separate customer for purposes of generating the table. For example, if a given user (or group of users that share an account) purchased items from two different shopping carts within the same account, these purchases could be treated as the purchases of separate users.

The product IDs may be converted to title IDs during this process, or when the table 60 is later used to generate recommendations, so that different versions of an item (e.g., hardcover and paperback) are represented as a single item. This may be accomplished, for example, by using a separate database which maps product IDs to title IDs. To generate a similar items table that strongly reflects the current tastes of

12

the community, the purchase histories retrieved in step 100 can be limited to a specific time period, such as the last six months.

In steps 102 and 104, the process generates two temporary tables 102A and 104A. The first table 102A maps individual customers to the items they purchased. The second table 104A maps items to the customers that purchased such items. To avoid the effects of "ballot stuffing," multiple copies of the same item purchased by a single customer are represented with a single table entry. For example, even if a single customer purchased 4000 copies of one book, the customer will be treated as having purchased only a single copy. In addition, items that were sold to an insignificant number (e.g., <15) of customers are preferably omitted or deleted from the tables 102A, 104B.

In step 106, the process identifies the items that constitute "popular" items. This may be accomplished, for example, by selecting from the item-to-customers table 104A those items that were purchased by more than a threshold number (e.g., 30) of customers. In the context of the Amazon.com Web site, to resulting set of popular items may contain hundreds of thousands or millions of items.

In step 108, the process counts, for each (popular\_item, other\_item) pair, the number of customers that are in common. A pseudocode sequence for performing this step is listed in Table 1. The result of step 108 is a table that indicates, for each (popular\_item, other\_item) pair, the number of customers the two have in common. For example, in the hypothetical table 108A of FIG. 3, POPULAR\_A and ITEM\_B have seventy customers in common, indicating that seventy customers bought both items.

TABLE 1

```

for each popular_item
  for each customer in customers of item
    for each other_item in items of customer
      increment common-customer-count(popular_item, other_item)

```

In step 110, the process generates the commonality indexes for each (popular\_item, other\_item) pair in the table 108A. As indicated above, the commonality index (CI) values are measures of the similarity between two items, with larger CI values indicating greater degrees of similarity. The commonality indexes are preferably generated such that, for a given popular<sub>13</sub> item, the respective commonality indexes of the corresponding other\_items take into consideration both (a) the number of customers that are common to both items, and (b) the total number of customers of the other\_item. A preferred method for generating the commonality index values is set forth in the equation below, in which  $N_A$  represents the number of customers of item\_A,  $N_B$  represents the number of customers of item\_B, and  $N_{common}$  represents the number of customers of item\_A and item\_B.

$$CI(\text{item}_A, \text{item}_B) = \frac{N_{common}}{\sqrt{N_A \times N_B}}$$

FIG. 4 illustrates this method in example form. In the FIG. 4 example, item\_P (a popular item) has two "other items," item\_X and item\_Y. Item\_P has been purchased by 300 customers, item\_X by 300 customers, and item\_Y by 30,000 customers. In addition, item\_P and item\_X have 20 customers in common, and item\_P and item\_Y have 25 customers

13

in common. Applying the equation above to the values shown in FIG. 4 produces the following results:

$$CI(\text{item\_P}, \text{item\_X}) = 20 / \sqrt{(300 \times 300)} = 0.0667$$

$$CI(\text{item\_P}, \text{item\_Y}) = 25 / \sqrt{(300 \times 30,000)} = 0.0083$$

Thus, even though items P and Y have more customers in common than items P and X, items P and X are treated as being more similar than items P and Y. This result desirably reflects the fact that the percentage of item\_X customers that bought item\_P (6.7%) is much greater than the percentage of item\_Y customers that bought item\_P (0.08%).

Because this equation is symmetrical (i.e.,  $CI(\text{item\_A}, \text{item\_B}) = CI(\text{item\_B}, \text{item\_A})$ ), it is not necessary to separately calculate the CI value for every location in the table 108A. In other embodiments, an asymmetrical method may be used to generate the CI values. For example, the CI value for a (popular\_item, other\_item) pair could be generated as (customers of popular item and other\_item)/(customers of other\_item).

Following step 110 of FIG. 3, each popular item has a respective "other\_items" list which includes all of the other\_items from the table 108A and their associated CI values. In step 112, each other\_items list is sorted from highest-to-lowest commonality index. Using the FIG. 4 values as an example, item\_X would be positioned closer to the top of the item\_B's list than item\_Y, since  $0.014907 > 0.001643$ .

In step 114, the sorted other\_items lists are filtered by deleting all list entries that have fewer than 3 customers in common. For example, in the other\_items list for POPULAR\_A in table 108A, ITEM\_A would be deleted since POPULAR\_A and ITEM\_A have only two customers in common. Deleting such entries tends to reduce statistically poor correlations between item sales.

In step 116, the sorted other\_items lists are truncated to length N to generate the similar items lists, and the similar items lists are stored in a B-tree table structure for efficient look-up.

As indicated above, any of a variety of other methods for evaluating similarities between items could be incorporated into the table generation process 66. For example, the table generation process could compare item contents and/or use previously-assigned product categorizations as additional indicators of item similarities. An important benefit of the FIG. 3 method, however, is that the items need not contain any content that is amenable to feature extraction techniques, and need not be pre-assigned to any categories. For example, the method can be used to generate a similar items table given nothing more than the product IDs of a set of products and user purchase histories with respect to these products.

Another important benefit of the Recommendation Service is that the bulk of the processing (the generation of the similar items table 60) is performed by an off-line process. Once this table has been generated, personalized recommendations can be generated rapidly and efficiently, without sacrificing breadth of analysis.

#### V. Instant Recommendations Service (FIGS. 5 and 6)

A specific implementation of the Recommendation Service, referred to herein as the Instant Recommendations service, will now be described with reference to FIGS. 5 and 6.

As indicated above, the Instant Recommendations service is invoked by the user by selecting a corresponding hyperlink from a Web page. For example, the user may select an "Instant Book Recommendations" or similar hyperlink to

14

obtain a listing of recommended book titles, or may select a "Instant Music Recommendations" or "Instant Video Recommendations" hyperlink to obtain a listing of recommended music or video titles. As described below, the user can also request that the recommendations be limited to a particular item category, such as "non-fiction," "jazz" or "comedies." The Instant Recommendations service generates the recommendations based exclusively on the purchase history and any item ratings profile of the particular user. The service becomes available to the user (i.e., the appropriate hyperlink is presented to the user) once the user has purchased and/or rated a threshold number (e.g. three) of popular items within the corresponding product group. If the user has established multiple shopping carts, the user may also be presented the option of designating a particular shopping cart to be used in generating the recommendations.

FIG. 5 illustrates the sequence of steps that are performed by the Instant Recommendations service to generate personal recommendations. Steps 180–194 in FIG. 5 correspond, respectively, to steps 80–94 in FIG. 2. In step 180, the process 52 identifies all popular items that have been purchased by the user (from a particular shopping cart, if designated) or rated by the user, within the last six months. In step 182, the process retrieves the similar items lists 64 for these popular items from the similar items table 60.

In step 184, the process 52 weights each similar items list based on the duration since the associated popular item was purchased by the user (with recently-purchased items weighted more heavily), or if the popular item was not purchased, the rating given to the popular item by the user. The formula used to generate the weight values to apply to each similar items list is listed in C in Table 2. In this formula, "is<sub>13</sub> purchased" is a boolean variable which indicates whether the popular item was purchased, "rating" is the rating value (1–5), if any, assigned to the popular item by the user, "order\_date" is the date/time (measured in seconds since 1970) the popular item was purchased, "now" is the current date/time (measured in seconds since 1970), and "6 months" is six months in seconds.

TABLE 2

1	Weight = ((is <sub>13</sub> purchased ? 5 : rating) * 2 - 5) *
2	(1 + (max((is <sub>13</sub> purchased ? order_date : 0) - (now - 6 months), 0))
3	/(6 months))

In line 1 of the formula, if the popular item was purchased, the value "5" (the maximum possible rating value) is selected; otherwise, the user's rating of the item is selected. The selected value (which may range from 1–5) is then multiplied by 2, and 5 is subtracted from the result. The value calculated in line 1 thus ranges from a minimum of –3 (if the item was rated a "1") to a maximum of 5 (if the item was purchased or was rated a "5").

The value calculated in line 1 is multiplied by the value calculated in lines 2 and 3, which can range from a minimum of 1 (if the item was either not purchased or was purchased at least six months ago) to a maximum of 2 (if order\_date=now). Thus, the weight can range from a minimum of –6 to a maximum of 10. Weights of zero and below indicate that the user rated the item a "2" or below. Weights higher than 5 indicate that the user actually purchased the item (although a weight of 5 or less is possible even if the item was purchased), with higher values indicating more recent purchases.

The similar items lists 64 are weighted in step 184 by multiplying the CI values of the list by the corresponding



15

weight value. For example, if the weight value for a given popular item is ten, and the similar items list 64 for the popular item is

```
(productid_A, 0.10), (productid_B, 0.09), (productid_C, 0.08), . . .
```

the weighted similar items list would be:

```
(productid_A, 1.0), (productid_B, 0.9), (productid_C, 0.8), . . .
```

The numerical values in the weighted similar items lists are referred to as “scores.”

In step 186, the weighted similar items lists are merged (if multiple lists exist) to form a single list. During this step, the scores of like items are summed. For example, if a given other\_item appears in three different similar items lists 64, the three scores (including any negative scores) are summed to produce a composite score.

In step 188, the resulting list is sorted from highest-to-lowest score. The effect of the sorting operation is to place the most relevant items at the top of the list. In step 190, the list is filtered by deleting any items that (1) have already been purchased or rated by the user, (2) have a negative score, or (3) do not fall within the designated product group (e.g., books) or category (e.g., “science fiction,” or “jazz”).

In step 192 one or more items are optionally selected from the recent shopping cart contents list (if such a list exists) for the user, excluding items that have been rated by the user or which fall outside the designated product group or category. The selected items, if any, are inserted at randomly-selected locations within the top M (e.g., 15) positions in the recommendations list. Finally, in step 194, the top M items from the recommendations list are returned to the Web server 32, which incorporates these recommendations into one or more Web pages.

The general form of such a Web page is shown in FIG. 6, which lists five recommended items. From this page, the user can select a link associated with one of the recommended items to view the product information page for that item. In addition, the user can select a “more recommendations” button 200 to view additional items from the list of M items. Further, the user can select a “refine your recommendations” link to rate or indicate ownership of the recommended items. Indicating ownership of an item causes the item to be added to the user’s purchase history listing.

The user can also select a specific category such as “non-fiction” or “romance” from a drop-down menu 202 to request category-specific recommendations. Designating a specific category causes items in all other categories to be filtered out in step 190 (FIG. 5).

#### VI. Shopping Cart Based Recommendations (FIG. 7)

Another specific implementation of the Recommendation Service, referred to herein as shopping cart recommendations, will now be described with reference to FIG. 7.

The shopping cart recommendations service is preferably invoked automatically when the user displays the contents of a shopping cart that contains more than a threshold number (e.g., 1) of popular items. The service generates the recommendations based exclusively on the current contents of the shopping cart. As a result, the recommendations tend to be highly correlated to the user’s current shopping interests. In other implementations, the recommendations may also be based on other items that are deemed to be of current interest to the user, such as items in the recent shopping cart contents of the user and/or items recently viewed by the user. Further, other indications of the user’s current shopping interests

16

could be incorporated into the process. For example, any search terms typed into the site’s search engine during the user’s browsing session could be captured and used to perform content-based filtering of the recommended items list.

FIG. 7 illustrates the sequence of steps that are performed by the shopping cart recommendations service to generate a set of shopping-cart-based recommendations. In step 282, the similar items list for each popular item in the shopping cart is retrieved from the similar items table 60. The similar items list for one or more additional items that are deemed to be of current interest could also be retrieved during this step, such as the list for an item recently deleted from the shopping cart or recently viewed for an extended period of time.

In step 286, these similar items lists are merged while summing the commonality index (CI) values of like items. In step 288, the resulting list is sorted from highest-to-lowest score. In step 290, the list is filtered to remove any items that exist in the shopping cart or have been purchased or rated by the user. Finally, in step 294, the top M (e.g., 5) items of the list are returned as recommendations. The recommendations are preferably presented to the user on the same Web page (not shown) as the shopping cart contents.

If the user has defined multiple shopping carts, the recommendations generated by the FIG. 7 process may be based solely on the contents of the shopping cart currently selected for display. As described above, this allows the user to obtain recommendations that correspond to the role or purpose of a particular shopping cart (e.g., work versus home).

The various uses of shopping cart contents to generate recommendations as described above can be applied to other types of recommendation systems, including content-based systems. For example, the current and/or past contents of a shopping cart can be used to generate recommendations in a system in which mappings of items to lists of similar items are generated from a computer-based comparison of item contents. Methods for performing content-based similarity analyses of items are well known in the art, and are therefore not described herein.

Although this invention has been described in terms of certain preferred embodiments, other embodiments that are apparent to those of ordinary skill in the art are also within the scope of this invention. For example, although the embodiments described herein employ item lists, other programming methods for keeping track of and combining sets of similar items can be used. Accordingly, the scope of the present invention is intended to be defined only by reference to the appended claims.

In the claims which follow, reference characters used to denote process steps are provided for convenience of description only, and not to imply a particular order for performing the steps.

What is claimed is:

1. A method of recommending items to users from a database of items, the method comprising:

maintaining item selection histories of each of a plurality of users of a server system that provides functionality for browsing and selecting items from an electronic catalog of items, each item selection history corresponding to, and identifying items selected by, a particular user;

collectively analyzing at least the item selection histories of the plurality of users, as collected over a period of time, in an off-line processing mode to generate a

17

plurality of data values that represent degrees to which specific items in the electronic catalog are related;  
storing a selected subset of the plurality of data values in a mapping structure that maps items to related items; and

for each of a plurality of users of the electronic catalog, using the mapping structure, including the data values stored therein, to generate personalized recommendations of items within the catalog.

2. The method of claim 1, wherein the personalized recommendations are generated substantially in real time.

3. The method of claim 1, wherein the mapping structure is a B-tree data structure.

4. The method of claim 3, wherein the B-tree data structure includes the data values of selected pairs of related items.

5. The method of claim 1, further comprising replicating the mapping structure, together with associated executable code used to generate personalized recommendations, across multiple machines of said server system to improve real time performance.

6. The method of claim 1, wherein storing a selected subset of the plurality of data values comprises using the data values to identify items that are sufficiently related to be mapped within the mapping structure.

7. The method of claim 1, wherein collectively analyzing the item selection histories comprises generating a data value that indicates a degree to which a first item and a second item are related, wherein the data value is dependent upon at least (a) a number of users that selected both the first item and the second item, (b) a total number of users that selected the first item, and (c) a total number of users that selected the second item.

8. The method of claim 7, wherein the data value is generated substantially according to the equation  $X/(\text{SQRT}(Y \times Z))$ , where X is the number of users that selected both the first item and the second item, Y is the total number of users that selected the first item, and Z is the total number of users that selected the second item.

9. The method of claim 1, wherein collectively analyzing the item selection histories comprises generating a data value that indicates a degree to which a first item and a second item are related, wherein the data value is dependent upon at least (a) a number of users that purchased both the first item and the second item, (b) a total number of users that purchased the first item, and (c) a total number of users that purchased the second item.

10. The method of claim 1, further comprising repeating the step of collectively analyzing the item selection histories periodically using a most recent set of item selection history data.

11. The method of claim 1, wherein each item selection history is a history of items selected for purchase by a user.

12. The method of claim 1, wherein each item selection history is a history of items selected for viewing by a user.

13. The method of claim 1, wherein generating personalized recommendations comprises identifying a plurality of items selected by a target user to which personalized recommendations are to be provided, and using the mapping structure and data values stored therein to select, in real time, a set of additional items that are collectively related to the plurality of items.

14. The method of claim 13, wherein identifying a plurality of items selected by a target user comprises identifying items currently in an electronic shopping cart of the target user.

18

15. The method of claim 14, further comprising displaying at least some of the additional items on a shopping cart page that displays current contents of the electronic shopping cart, such that the shopping cart page suggests additional items that may be of interest to the target user.

16. The method of claim 1, wherein generating personalized recommendations comprises identifying a plurality of items purchased by a target user, and using the mapping structure to select, in real time, a set of additional items that are collectively related to the plurality of items.

17. The method of claim 1, wherein generating personalized recommendations comprises identifying a set of items that have been viewed by a target user during browsing of the electronic catalog, and using the mapping structure to select, in real time, a set of additional items that are collectively related to the set of viewed items.

18. The method of claim 1, wherein the personalized recommendations are generated and provided to a target user without requiring the target user to explicitly rate items or to create an input list of items.

19. The method of claim 1, wherein the personalized recommendations are generated and provided to a target user without requiring the target user to explicitly request recommendations.

20. The method of claim 1, wherein the personalized recommendations are generated and provided to a target user without requesting preference information from the target user.

21. The method of claim 1, wherein the personalized recommendations are generated for and provided to each of the plurality of users without updating the mapping structure.

22. The method of claim 1, wherein maintaining item selection histories comprises maintaining an item selection history reflective of items selected by a user over multiple days.

23. The method of claim 1, wherein maintaining item selection histories comprises maintaining an item selection history reflective of actions performed by a user over multiple days.

24. The method of claim 1, wherein each item selection history comprises data indicative of a date that an item was selected by a corresponding user.

25. A computer-implemented method of generating a mapping of items to related items, the method comprising: maintaining item selection histories of each of a plurality of users of a server system that provides functionality for browsing and selecting items from an electronic catalog of items, each item selection history corresponding to, and identifying items selected by, a particular user;

collectively and programmatically analyzing the item selection histories of the plurality of users to generate a data value that represents a degree to which a first item and a second item in the catalog are related, wherein the data value is dependent upon at least (a) a number of users that selected both the first item and the second item, (b) a total number of users that selected the first item, and (c) a total number of users that selected the second item; and

using the data value to determine whether the first item should be mapped to the second item in a mapping structure that maps items to related items.

26. The method of claim 25, wherein the data value is generated according to the equation  $X/(\text{SQRT}(Y \times Z))$ , where X is the number of users that selected both the first

19

item and the second item, Y is the total number of users that selected the first item, and Z is the total number of users that selected the second item.

27. The method of claim 25, wherein the item selection histories identify items selected for purchase by the users, and the data value is dependent upon at least (a) a number of users that purchased both the first item and the second item, (b) a total number of users that purchased the first item, and (c) a total number of users that purchased the second item.

28. The method of claim 25, further comprising determining whether a given item should be recommended to a user based at least in part on a degree to which the given item is related to each of multiple items known to be of interest to the user, as indicated by the mapping structure.

29. A computer-implemented method of recommending items to users of a server system that provides functionality for selecting items represented in an electronic catalog, the method comprising:

identifying a plurality of items that are currently selected by a target user to obtain from an online business entity, including multiple items currently represented in a user-specific data repository in which item selections are maintained persistently over multiple sessions;

identifying a set of additional items to recommend to the target user according to a selection algorithm in which a candidate item is considered for inclusion in the set based at least in part upon a degree to which the candidate item is related to each of said plurality of items currently selected by the target user to obtain; and recommending at least some of the additional items in the set to the target user while the plurality of items are selected by the target user to obtain.

30. The method of claim 29, wherein the method is performed substantially in real time when the target user accesses an electronic shopping cart in which said plurality of items are represented.

31. The method of claim 29, wherein the degree to which the candidate item is related to each of the plurality of items is based at least in part on results of a programmatic correlation analysis of purchase histories of a plurality of users.

32. The method of claim 29, wherein the degree to which the candidate item is related to each of the plurality of items is based at least in part on results of a programmatic correlation analysis of item viewing histories of a plurality of users.

33. The method of claim 29, wherein the set of additional items is identified at least in part by retrieving, from a B-tree data structure, a related items list for each of the plurality of items.

34. The method of claim 29, wherein recommending at least some of the additional items to the target user comprises displaying representations of the additional items on a web page that displays a representation of said plurality of items.

35. The method of claim 29, wherein recommending at least some of the additional items to the target user comprises filtering out additional items that are included within a purchase history of the target user.

36. The method of claim 29, wherein identifying a plurality of items that are currently selected comprises identifying a plurality of items represented in an electronic shopping cart associated with the target user.

37. The method of claim 29, wherein identifying a plurality of items that are currently selected comprises identifying items currently selected by the target user to purchase.

20

38. The method of claim 29, wherein identifying a plurality of items that are currently selected comprises identifying items currently selected for delivery to the target user.

39. The method of claim 29, wherein the items are physical products.

40. The method of claim 29, wherein the items include at least one of (a) an audio-visual work and (b) a video title.

41. The method of claim 29, wherein the degree to which the candidate item is related to each of the plurality of items is based, at least in part, on results of an off-line correlation analysis of item selection histories of a plurality of users.

42. A data mining method, comprising:

programmatically analyzing user activity data associated with a plurality of users of an electronic catalog, in an off-line processing mode, to generate data values representing degrees to which specific catalog items are related to one another, said user activity data reflecting user interest in specific catalog items;

using the data values to select, for each of a plurality of the catalog items, a corresponding set of related catalog items;

generating mappings of said plurality of catalog items to the corresponding sets of related catalog items, wherein generating the mappings comprises storing selected ones of said data values, together with associated item identifiers, in a mapping data structure; and

programmatically generating personalized item recommendations for each of a plurality of target users of the electronic catalog using said data values stored in the mapping data structure and using said mappings, wherein the personalized item recommendations are generated for each of the plurality of target users without updating the mapping data structure.

43. The method of claim 42, wherein programmatically analyzing the user activity data comprises determining, for a selected pair of catalog items, how many users that purchased the first item of the pair also purchased the second item of the pair.

44. The method of claim 42, wherein programmatically analyzing the user activity data comprises analyzing item viewing activities of users.

45. The method of claim 42, wherein programmatically analyzing the user activity data comprises analyzing shopping cart activities of users.

46. The method of claim 42, wherein programmatically generating personalized item recommendations comprises using the mappings to select catalog items that are related to a plurality of catalog items previously purchased by a target user, as determined from a purchase history of the target user.

47. The method of claim 42, wherein programmatically generating personalized item recommendations comprises using the mappings to select catalog items that are related to a plurality of catalog items previously viewed by a target user.

48. The method of claim 42, wherein programmatically generating personalized item recommendations comprises using the mappings to select catalog items that are related to a plurality of catalog items currently selected by a target user to obtain from an online business entity.

49. The method of claim 42, wherein the items include video titles.

50. The method of claim 42, wherein the data values are based on at least one type of item selection activity performed by users for purposes other than to obtain recommendations.

**21**

**51.** A data mining method, comprising:  
 programmatically analyzing user activity data associated  
 with a plurality of users of an electronic catalog, in an  
 off-line processing mode, to generate data values rep-  
 resenting degrees to which specific catalog items are  
 related to one another, said user activity data reflecting  
 user interest in specific catalog items; 5  
 using the data values to select, for each of a plurality of  
 the catalog items, a corresponding set of related catalog  
 items; 10  
 generating mappings of said plurality of catalog items to  
 the corresponding sets of related catalog items; and  
 programmatically generating personalized item recom-  
 mendations for each of a plurality of users of the  
 electronic catalog using said mappings; 15  
 wherein programmatically generating personalized item  
 recommendations comprises using the mappings, in  
 combination with an item ratings profile of a target  
 user, to select catalog items to recommend to the target  
 user, wherein the item ratings profile comprises scores  
 explicitly assigned to specific items by the target user. 20

**22**

**52.** A data mining method, comprising:  
 programmatically analyzing user activity data associated  
 with a plurality of users of an electronic catalog, in an  
 off-line processing mode, to generate data values rep-  
 resenting degrees to which specific catalog items are  
 related to one another, said user activity data reflecting  
 user interest in specific catalog items;  
 using the data values to select, for each of a plurality of  
 the catalog items, a corresponding set of related catalog  
 items;  
 generating mappings of said plurality of catalog items to  
 the corresponding sets of related catalog items; and  
 programmatically generating personalized item recom-  
 mendations for each of a plurality of users of the  
 electronic catalog using said mappings;  
 wherein programmatically generating personalized item  
 recommendations comprises using the mappings to  
 select catalog items that are related to a plurality of  
 catalog items currently represented in a user-specific  
 data repository in which item selections are maintained  
 persistently over multiple sessions.

\* \* \* \* \*