CUSTOMER CLUSTERING USING INTEGER PROGRAMMING

Applicant: Sears Brands, LLC, Hoffman Estates, IL (US)

Inventors: Burcu Aydin, Mountain View, CA (US); Michael Tamir, San Jose, CA (US)

Assignee: Sears Brands, LLC, Hoffman Estates, IL (US)

Appl. No.: 14/084,903

Filed: Nov. 20, 2013

ABSTRACT

Methods and apparatus are disclosed regarding an e-commerce system that clusters customers based on demographic data and purchase history data for the customers. In some embodiments, the e-commerce system solves an Integer Program that accounts for the demographic data and purchase history data in order to identify a hyperplane that splits a selected cluster of customers.
Fig. 5

Begin

Pre-process purchase history data to obtain CI matrix

Pre-process demographic data to obtain feature space

Standardize CI matrix to obtain standardized CI matrix or transaction space

End
<table>
<thead>
<tr>
<th>Customer ID</th>
<th>100001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>1/1/2013</td>
</tr>
<tr>
<td>Transaction Location</td>
<td>San Jose Sears</td>
</tr>
<tr>
<td>Quantity</td>
<td>1</td>
</tr>
<tr>
<td>Price</td>
<td>10$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category ID</th>
<th>001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>1</td>
</tr>
<tr>
<td>Category ID</td>
<td>002</td>
</tr>
<tr>
<td>Quantity</td>
<td>1</td>
</tr>
</tbody>
</table>
Begin
Solve Integer Program To Obtain Hyperplane For Data Set 610

Select Partition As Data Set Further Partitioning 630

Continue Partitioning? 620

Yes

No

End

Fig. 12
CUSTOMER CLUSTERING USING INTEGER PROGRAMMING

FIELD OF THE INVENTION

Various embodiments relate to electronic commerce (e-commerce), and more particularly, to classifying customers in an e-commerce environment.

BACKGROUND OF THE INVENTION

Electronic commerce (e-commerce) websites are an increasingly popular venue for consumers to research and purchase products without physically visiting a conventional brick-and-mortar retail store. An e-commerce website may provide products and/or services to a vast number of customers. As a result of providing such products and/or services, the e-commerce website may obtain extensive amounts of data about their customer base. Such customer data may aid the e-commerce website to provide products and/or services that are relevant and/or otherwise desirable to a particular customer.

In particular, an e-commerce website may attempt to identify groups of customers with similar interests or similar lifestyles. The e-commerce website may analyze these identified groups to derive generalizations regarding members of the group. The e-commerce website may then tailor its services to members of each group based upon the derived generalizations.

Limitations and disadvantages of conventional and traditional approaches should become apparent to one of skill in the art, through comparison of such systems with aspects of the present invention as set forth in the remainder of the present application.

BRIEF SUMMARY OF THE INVENTION

Apparatus and methods of classifying or grouping customers are substantially shown in and/or described in connection with at least one of the figures, and are set forth more completely in the claims.

These and other advantages, aspects and novel features of the present invention, as well as details of an illustrated embodiment thereof, will be more fully understood from the following description and drawings.

BRIEF DESCRIPTION OF SEVERAL VIEWS OF THE DRAWINGS

FIG. 1 shows an e-commerce environment comprising a computing device and an e-commerce system in accordance with an embodiment of the present invention.

FIG. 2 shows an embodiment of a computing device for use in the e-commerce environment of FIG. 1.

FIG. 3 shows user profiles and product catalogs maintained by an e-commerce system of FIG. 1.

FIG. 4 shows an embodiment of a product listing provided by the e-commerce system of FIG. 1.

FIG. 5 shows a flowchart for an embodiment of a process that may be used by the e-commerce system of FIG. 1 to obtain a transaction space and a feature space from purchase history data and demographic data.

FIG. 6 shows an example entry of the purchase history data for the e-commerce system of FIG. 1.

FIG. 7 shows an example purchase history table for the e-commerce system of FIG. 1 after evaluating and retaining data of the purchase history data for a time window of interest.

FIG. 8 shows an example purchase history table for the e-commerce system of FIG. 1 after combining rows that correspond to the same customer and product category.

FIG. 9 shows an entry from an example customer-item (CI) matrix for the e-commerce system of FIG. 1.

FIG. 10 shows an example quantile table for the e-commerce system of FIG. 1.

FIG. 11 shows a standardized entry from the example quantile table of FIG. 10.

FIG. 12 shows a flowchart of a process that may be used by the e-commerce system of FIG. 1 to cluster customers based on the transaction space and feature space.

FIGS. 13-16 depict an example partitioning of a customer base.

DETAILED DESCRIPTION OF THE INVENTION

Aspects of the present invention are related to classifying and/or grouping customers together that exhibit similar interests, lifestyles, and/or purchase behavior. More specifically, certain embodiments of the present invention relate to apparatus, hardware and/or software systems, and associated methods that cluster customers based on solving an integer program that accounts for purchase history data and demographic data of the customers.

Referring now to FIG. 1, an e-commerce environment 10 is depicted. As shown, the e-commerce environment 10 may include a computing device 20 connected to an e-commerce system 30 via a network 40. The network 40 may include a number of private and/or public networks such as, for example, wireless and/or wired LAN networks, cellular networks, and the Internet that collectively provide a communication path and/or paths between the computing device 20 and the e-commerce system 30. The computing device 20 may include a desktop, a laptop, a tablet, a smart phone, and/or some other type of computing device which enables a user to communicate with the e-commerce system 30 via the network 40. The e-commerce system 30 may include one or more web servers, database servers, routers, load balancers, and/or other computing and/or networking devices that operate to provide an e-commerce experience for users that connect to the e-commerce system 30 via the computing device 20 and the network 40.

The e-commerce system 30 may further include a customer classifier 33, one or more tailored services 35, and one or more electronic databases 37 upon which are stored purchase history data 38 and demographic data 39 for customers of the e-commerce system 30. The classifier 33 may include one or more firmware and/or software instructions, routines, modules, etc. that the e-commerce system 30 may execute in order to classify, group, or cluster customers of the e-commerce system 30 into classes, groups, or clusters of customers that exhibit similar purchasing habits. The classifier 33 may analyze purchase history data and demographic data for the customers to identify clusters of customers with similar purchasing preferences.

The tailored services 35 may comprise one or more firmware and/or software instructions, routines, modules, etc. that the e-commerce system 30 may execute in order to tailor one or more aspects of the e-commerce system 30 for a particular customer. The tailored services 35 may include
advertisements, promotions, product recommendations, email campaigns, etc. that are tailored based upon the cluster to which the customer has been placed.

The classifier 33 and tailored services 35 may be executed concurrently by a single computing device of the e-commerce system 30. However, in some embodiments, a computing device may execute the classifier 33 offline in order to obtain appropriate clusters and other input data for the tailored services 35. Moreover, the classifier 33 may periodically (e.g., once an hour, once a day, once a week, etc.) provide one or more of the tailored services 35 with updated cluster and other input data. In this manner, the e-commerce system 30 may continue to provide tailored services 35 without the constant overhead of the classifier 33 and/or without the overhead of constant updates. For example, the e-commerce system 30 may execute the classifier 33 only during generally idle periods (e.g., after normal business hours).

Further details regarding the classifier 33 and the tailored services 35 are presented below in regard to FIGS. 5-11.

FIG. 1 depicts a simplified embodiment of the e-commerce environment 10 which may be implemented in numerous different manners using a wide range of different computing devices, platforms, networks, etc. Moreover, while aspects of the e-commerce environment 10 may be implemented using a client/server architecture, aspects of the e-commerce may be implemented using a peer-to-peer architecture or another networking architecture.

As noted above, the e-commerce system 30 may include one or more computing devices. FIG. 2 depicts an embodiment of a computing device 50 suitable for the computing device 20 and/or the e-commerce system 30. As shown, the computing device 50 may include a processor 51, a memory 53, a mass storage device 55, a network interface 57, and various input/output (I/O) devices 59. The processor 51 may be configured to execute instructions, manipulate data and generally control operation of other components of the computing device 50 as a result of its execution. To this end, the processor 51 may include a general purpose processor such as an x86 processor or an ARM processor which are available from various vendors. However, the processor 51 may also be implemented using an application specific processor and/or other logic circuitry.

The memory 53 may store instructions and/or data to be executed and/or otherwise accessed by the processor 51. In some embodiments, the memory 53 may be completely and/or partially integrated with the processor 51.

In general, the mass storage device 55 may store software and/or firmware instructions which may be loaded in memory 53 and executed by processor 51. The mass storage device 55 may further store various types of data which the processor 51 may access, modify, and/or otherwise manipulate in response to executing instructions from memory 53. To this end, the mass storage device 55 may comprise one or more redundant array of independent disks (RAID) devices, traditional hard disk drives (HDD), solid-state device (SSD) drives, flash memory devices, read only memory (ROM) devices, etc.

The network interface 57 may enable the computing device 50 to communicate with other computing devices directly and/or via network 40. To this end, the networking interface 57 may include a wired networking interface such as an Ethernet (IEEE 802.3) interface, a wireless networking interface such as a WiFi (IEEE 802.11) interface, a radio or mobile interface such as a cellular interface (GSM, CDMA, LTE, etc.), and/or some other type of networking interface capable of providing a communications link between the computing device 50 and network 40 and/or another computing device.

Finally, the I/O devices 59 may generally provide devices which enable a user to interact with the computing device 50 by either receiving information from the computing device 50 and/or providing information to the computing device 50. For example, the I/O devices 59 may include display screens, keyboards, mice, touch screens, microphones, audio speakers, etc.

While the above provides general aspects of a computing device 50, those skilled in the art readily appreciate that there may be significant variation in actual implementations of a computing device. For example, a smart phone implementation of a computing device may use vastly different components and may have a vastly different architecture than a database server implementation of a computing device. However, despite such differences, computing devices generally include processors that execute software and/or firmware instructions in order to implement various functionality.

As such, aspects of the present application may find utility across a vast array of different computing devices and the intention is not to limit the scope of the present application to a specific computing device and/or computing platform beyond any such limits that may be found in the appended claims.

As part of the provided e-commerce experience, the e-commerce system 30 may enable customers, which may be guests or members of the e-commerce system 30, to browse and/or otherwise locate products. The e-commerce system 30 may further enable such customers to purchase products offered for sale. To this end, the e-commerce system 30 may maintain an electronic product database or product catalog 300 which may be stored on an associated mass storage device 55. As shown in FIG. 3, the product catalog 300 includes product listings 310 for each product available for purchase. Each product listing 310 may include various information or attributes regarding the respective product, such as a unique product identifier (e.g., stock-keeping unit “SKU”), a product description, product image(s), manufacture information, available quantity, price, product features, etc. Moreover, while the e-commerce system 30 may enable guests to purchase products without registering and/or otherwise signing-up for a membership, the e-commerce system 30 may provide additional and/or enhanced functionality to those users that become a member.

To this end, the e-commerce system 30 may enable members to create a customer profile 330. As shown, a customer profile 330 may include personal information 331, purchase history data 335, and other customer activity data 337. The personal information 331 may include such items as name, mailing address, email address, phone number, billing information, clothing sizes, birthdates of friends and family, etc. The purchase history data 335 may include information regarding products previously purchased by the customer from the e-commerce system 30. The customer history data 335 may further include products previously purchased from affiliated online and brick-and-mortar vendors.

The other customer activity data 337 may include information regarding prior customer activities such as products for which the customer has previously searched, products for which the customer has previously viewed, products for which the customer has provide comments, products for
which the customer has rated, products for which the customer has written reviews, etc. and/or purchased from the e-commerce system. The other customer activity data may further include similar activities associated with affiliated online and brick-and-mortar vendors.

As part of the e-commerce experience, the e-commerce system may cause a computing device to display a product listing as shown in FIG. 4. In particular, the e-commerce system may provide such a product listing in response to a member browsing products by type, price, kind, etc., viewing a list of products obtained from a product search, and/or other techniques supported by the e-commerce system for locating products of interest. As shown, the product listing may include one or more representative images of the product as well as a product description. The product listing may further include one or more products recommended by a recommendation engine of the tailored services. In particular, the recommendation engine may provide product recommendations based on the personal information provided in the customers' profile. The e-commerce system may collect and maintain purchase history data for the customer over a period of time. The purchase history data, in its raw form, may include information recorded for each purchase. An example entry is shown in FIG. 6. As shown, the e-commerce system may maintain the purchase history data in one or more relational database tables. Each row of the purchase history table may include a row for each transaction, and each row may include a customer identifier (ID) that uniquely identifies the customer associated with the corresponding transaction.

The classifier may preprocess the raw purchase history information found in the purchase history table. To this end, the classifier may select a time window (e.g., the most recent 24 months). The classifier may extract entries from the purchase history table that have a transaction date that falls within the selected time window. The classifier may then discard all fields other than the Customer ID, Item ID and Quantity of that particular item purchased in that transaction.

Many e-commerce sites maintain a product hierarchy of product identifiers where the Item ID corresponds to the lowest level of such hierarchy and various Category IDs lie higher up in the product hierarchy. Moreover, in many environments, the Item IDs are at such a fine granularity that correlations between purchases may be lost. In such situations, the classifier may be configured to coalesce purchased items of multiple Item IDs under a single Category ID that lies at a high level in the product hierarchy.

FIG. 7 shows an example table after evaluating the time window as described above. As may be seen from FIG. 7, the resulting table may still include multiple entries or rows for each Customer ID and Category ID pair. The classifier may apply a pivoting step to the resulting table in order to combine rows having the same Customer ID and Category ID pair into a single row. As shown in FIG. 8, the resulting table includes a single row for each Customer ID and Category ID pair and includes Quantity data that contains the sum of all purchased quantities for this ID pair.

From the table shown in FIG. 8, the classifier may create a Customer-Item (CI) matrix. In the CI matrix, each row corresponds to a unique Customer ID, and each column corresponds to a unique Category ID, and the entry CI corresponds to the quantity of this Customer ID and Category ID pair from the table shown in FIG. 7. If a particular customer did not purchase from a product in a category of CI matrix, then corresponding entry is zero.

At 515, the classifier may further preprocess the demographic data of its customers to obtain a feature space. The e-commerce system may collect demographic data from customers such as personal information provided in the customers' profile. The e-commerce system may further obtain demographic data for customers from various providers of demographic data. Based on such collected demographic data, the classifier may maintain and/or create a demographic table. The demographic table may include a row for each Customer ID. Moreover, each column of the table may represent a different feature such as, for example, age, gender, occupation, number of children, etc. During preprocessing, the classifier may preprocess each demographic entry into a numerical value. For example, the "Gender" column may contain only two kinds of entries, male and female. The classifier may preprocess the demographic table such that that Gender column includes a 1 for each female customer and a 0 otherwise. The preprocessing demographic table may form the feature space for later classification.

After preprocessing the purchase history and demographic data, the classifier may standardize the CI matrix to obtain a standardized CI matrix which refers to transaction space. Standardizing the CI matrix may ensure that the columns of the standardized CI matrix are scale-wise comparable with each other. In one embodiment, the classifier applies standardization to each column separately using a bin quantiles standardization (BQS) technique. However, other standardization techniques may be utilized.

To illustrate the BQS technique, one example column of the CI matrix is shown in FIG. 9. If depicted column corresponds to a category ID CID in the CI matrix, then the information in column suggests that customer 1 bought 1 unit of an item corresponding to category ID 4, customer 4 bought 2 items, customer 6 bought 1 item, and customer 7 bought 8 items. The classifier in accordance with the BQS technique may traverse the column, record every unique quantity except zero that appears along with how many times each unique quantity appears in the column. The classifier may sort the results based on occurrence of each unique quantity. See, e.g., the Occurrences column of FIG. 10. The classifier may traverse the occurrences to obtain a cumulative sum of the number of occurrences. See, e.g., Cumulative Occurrences column of FIG. 10. Furthermore, the classifier for each row may divide the respective cumulative occurrence value by the last number in the cumulative occurrence column (i.e., the total number of occurrences) to obtain the quantile value for that row. See, e.g., Quantile column of FIG. 10.

The BQS result shown in FIG. 10 suggests that the customers who bought 1 item associated with the category ID constitute the first 50% quantile, customers who bought 2 or
less such items are the 75% quantile, and customers who bought 8 or less such items are the 100% quantile. The classifier 33 may then update the quantity values of the original column with their corresponding quantile values as shown in FIG. 11 to obtain the standardized column.

[0045] The BQS technique may provide two advantages. One, all the numbers in the columns of CI matrix are guaranteed to be between 0 and 1, therefore the purchase patterns of high-frequency items such as grocery items and a low-frequency items such as expensive electronics items are comparable. Second, because the quantity values are thought in terms of frequencies of each number appearing and their relative order rather than their nominal values, the occasional very large number observed in the columns do not skew the analysis.

[0046] After obtaining feature space the standardized transaction space, the classifier 33 may classify or cluster the customers. In particular, the classifier 33 may attempt to find linear partitions in the feature space that divides the data points (customers) into groups or clusters with the smallest sum of distances within themselves. The distances are defined using the standardized transaction space.

[0047] The distance between customer A and customer B is a measure of the dissimilarity between their purchase history data 335. While many distance functions may be used, the classifier 33 in one embodiment uses the Minkowski distance in Euclidean space. The Minkowski distance for an integer p may be represented by the following expression:

\[ (\sum_{i=1}^{n} |x_{i} - y_{i}|^{p})^{1/p} \]

where \( C_{IA} \) represents the row in the standardized CI matrix for the customer A; \( C_{IB} \) represents the row in the standardized CI matrix for the customer B; \( B_{ij} \) represents the \( i^{th} \) element of the row \( C_{IA} \); \( B_{ij} \) represents the \( j^{th} \) element of the row \( C_{IB} \). The cases where \( p=1 \) and \( p=2 \) correspond to the Manhattan distance and Euclidean distance, respectively.

[0048] The classifier 33 may alternatively utilize a distance function that provides a metric of the similarity between customers. In such an embodiment, the classifier 33 may attempt to minimize the sum of inner-similarities per cluster. For example, the classifier 33 may use Mahalanobis distance functions, correlation functions, and/or some other similarity function in such an embodiment.

[0049] After obtaining the feature space and transaction space, the classifier 33 may proceed to analyze the feature space and transaction space in order to identify clusters of customers with similar purchasing behaviors. To this end, the classifier 33 may iteratively divide customer sets into two partitions until a suitable number of partitions for the customer base is obtained. In particular, the classifier 33 may divide the feature space into two partitions that minimizes the inner-distance between members of the cluster in the transaction space by solving an Integer Program that takes into account both the feature space and transaction space of the customer base.

[0050] In one embodiment, the following parameters, data, variables, and formulation define a Integer Program which may be solved to obtain a hyperplane that suitably divides the customer base into two clusters.

[0051] Parameters and Data:

[0052] \( n \) = number of customers;

[0053] \( m \) = number of dimensions in feature space;

[0054] \( x_{i} \) = length-m coordinate vector of customer i in feature space for \( i=1 \ldots n \);

[0055] \( d_{ij} \) = distance between customers i and j in transaction space according to a pre-selected distance metric;

[0056] \( C \) = a large constant; and

[0057] \( \epsilon \) = a small constant (epsilon).

[0058] Variables:

[0059] \( I_{i} \) = indicator variable of customer i, which is one if customer i is in cluster 1 (one side of the optimum hyperplane), and zero if the customer is in cluster 2 (the other side of the hyperplane) in the feature space.

[0060] \( I_{ij} \) = indicator variable for customer pair (i,j), which is equal to one if i and j are in the same cluster, and zero if they are in different clusters.

[0061] \( \beta \) = the length-m direction vector in feature space that defines the direction of the dividing hyperplane.

[0062] \( \beta_{0} \) = scalar intercept of the dividing hyperplane.

[0063] Formulation:

[0064] Minimize:

\[ \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} I_{ij} \]

[0065] Subject to:

\[ \beta_{0} \sum_{i=1}^{n} I_{i} - \sum_{i=1}^{n} C \epsilon \sum_{i=1}^{n} I_{i} \]

\[ \sum_{i=1}^{n} I_{ij} = 2 \forall j \]

\[ I_{i} \in \{0,1\} \forall i \]

[0066] The above Integer Program, when solved by an Integer Programming solver of the classifier 33, returns the clustering of customers in the feature space together with hyperplane variables \( \beta \) and \( \beta_{0} \) that define the division rule for the clusters. The classifier 33 may use the division rule to place new customers into one of the defined clusters based on known demographic features. By doing so, the classifier 33 may obtain some insight into the likely purchasing behavior for a new customer despite not having much or any purchase history data for the new customer.

[0067] The above Integer Program, however, divides the customer base into only two clusters or partitions, which is most likely not enough number of clusters to provide meaningful insight into the purchasing behaviors of the customer base. Accordingly, the classifier 33 may iteratively apply the above Integer Program in order to further divide the clusters until a suitable number of clusters are obtained. Such an iterative clustering method 600 is shown in FIG. 12.

[0068] At 610, the classifier 33 at 610 may solve the above Integer Program to obtain a hyperplane that divides or partitions the customer base or data set into two partitions or clusters. After dividing the data set into two clusters, the classifier 33 at 620 may determine whether further partitioning of the data set is warranted. To this end, the classifier 33 may make such a determination based upon a stopping rule. A stopping rule may define conditions for stopping further partitioning of the data set and for identifying which cluster or clusters to further divide. A first example stopping rule may be to pre-define the desired number of clusters, and iteratively
keep dividing the cluster with the largest population until the
desired number of clusters is reached. A second example
stopping rule may be to define the largest population to be
allowed in a single cluster, and keep dividing the clusters that
are more populated than this limit until no cluster exceeds this
limit. It should be appreciated that the above two stopping
rules are merely examples and that other stopping rules and/or
a combination of rules may be used by the classifier 33 to
ascertain whether to cease partitioning and/or selecting which
clusters to further partition.

[0069]  If the classifier 33 determines that no further parti-
tioning is warranted, then the classifier 33 may cease further
partitioning of the data set. However, if the classifier 33 deter-
mines that the stopping rules indicates further partitioning is
warranted, then the classifier 33 at 630 may select a cluster for
further partitioning based on the stopping rule. For example,
the classifier 33 per the first example stopping rule may select
the cluster having the largest population for further partition-
ing. If the second example stopping rule is being used, then
the classifier 33 may select a cluster having a population
greater than the predefined limit.

[0070]  After selecting an appropriate cluster for further
partitioning, the classifier 33 may return to 610 in order to
solve the Integer Program and obtain a hyperplane that par-
titions the selected cluster into two smaller clusters. In this
manner, the classifier 33 may continue to obtain further par-
titions until a suitable number of partitions is achieved per the
stopping rule in effect.

[0071]  Referring now to FIGS. 13-16, an example of par-
titioning a data set of customers per the method 600 is shown.
In particular, the example illustrates partitioning based on a
stopping rule of the largest allowable cluster having a popula-
tion of 3. Starting with FIG. 13, an unclustered data set of 9
customers in a two dimensional feature space is shown. FIG.
14 shows a hyperplane H1 obtained by the classifier 33 as a
result of solving the Integer Program in order to partition the
9 customers of FIG. 13. After such partitioning of FIG. 14, the
lower partition has a data set of 3 customers and is thus not
divided further per the stopping rule. The upper partition,
however, defines a data set of 6 customers and thus exceeds
the population limit of 3 for the stopping rule. As such, the
classifier solves the Integer Program for the upper data set to
obtain the hyperplane H2 shown in FIG. 15.

[0072]  After such partitioning of FIG. 15, the upper left
partition has a data set of 2 customers and is thus divided
further per the stopping rule. The upper right partition, how-
ever, defines a data set of 4 customers and thus still exceeds
the population limit of 3 for the stopping rule. As such, the
classifier solves the Integer Program for the upper right data
set to obtain the hyperplane H3 shown in FIG. 16. After such
partitioning of FIG. 16, all partitions have less than the popula-
tion limit of 3. As such, the classifier 33 ceases further parti-
tioning of the customer base per the stopping rule.

[0073]  Various embodiments of the invention have been
described herein by way of example and not by way of limi-
tation in the accompanying figures. For clarity of illustra-
tion, exemplary elements illustrated in the figures may not nec-
ecessarily be drawn to scale. In this regard, for example, the
dimensions of some of the elements may be exaggerated rela-
tive to other elements to provide clarity. Furthermore,
where considered appropriate, reference labels have been
repeated among the figures to indicate corresponding or anal-
ogous elements.

[0074]  Moreover, certain embodiments may be imple-
mented as a plurality of instructions on a non-transitory, com-
puter readable storage medium such as, for example, flash
memory devices, hard disk devices, compact disc media,
DVD media, EEPROMs, etc. Such instructions, when
executed by one or more computing devices, may result in the
one or more computing devices identifying customer clusters
based on purchase history data and demographic data for the
customer.

[0075]  While the present invention has been described with
reference to certain embodiments, it will be understood by
those skilled in the art that various changes may be made and
equivalents may be substituted without departing from the
scope of the present invention. For example, the above
embodiments were described primarily from the standpoint
of an e-commerce environment. However, it should be ap-
preciated that clustering of customers may be useful in other
environments as well. For example, a brick-and-mortar store
may cluster customers in order to provide targeted mailing,
coupons, and/or other types of promotions to its customers. In
addition, many modifications may be made to adapt a par-
ticular situation or material to the teachings of the present
invention without departing from its scope. Therefore, it is
intended that the present invention not be limited to the par-
ticular embodiment or embodiments disclosed, but that the
present invention encompasses all embodiments falling
within the scope of the appended claims.

What is claimed is:

1. A computer-implemented method, comprising:
   iteratively splitting a plurality of customers into a plurality
   of clusters based on purchase history data and demo-
graphic data for the plurality of customers, wherein each
   iteration comprises selecting a cluster of customers and
   splitting the selected cluster until a stopping rule is sat-
   isfied; and
tailing services provided to a customer based on a cluster
   from the plurality of clusters in which the customer resi-

2. The computer-implemented method of claim 1, wherein
   the tailoring comprises providing product recomenda-
   tions based on the cluster in which the customer resides.

3. The computer-implemented method of claim 1, wherein
   the tailoring comprises providing product promotions based
   on the cluster in which the customer resides.

4. The computer-implemented method of claim 1, wherein
   the tailoring comprises providing coupons based on the cluster
   in which the customer resides.

5. The computer-implemented method of claim 1, wherein
   the tailoring comprises solving an Integer Program that
   accounts for the purchase history data and demographic data
   for the selected cluster.

6. The computer-implemented method of claim 1, wherein
   the splitting comprises solving an Integer Program that
   accounts for the purchase history data and demographic data
   for the selected cluster.

7. The computer-implemented method of claim 1, wherein
   the selecting comprises selecting a cluster that has a popula-
   tion greater than a limit specified by the stopping rule.

8. The computer-implemented method of claim 1, wherein
   the selecting comprises selecting a cluster that has the largest
   population of the plurality of clusters.

9. The computer-implemented method of claim 1, further
   comprising ceasing the iteratively splitting in response to
determining the plurality of clusters comprises a quantity of
clusters desired by the stopping rule.
10. The computer-implemented method of claim 1, further comprising ceasing the iteratively splitting in response to determining that no cluster of the plurality of clusters comprises exceeds a population limit specified by the stopping rule.

11. The computer-implemented method of claim 1, wherein the splitting comprises solving the following Integer Program:

Minimize:

\[ \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} \]

Subject to:

- \[ \beta_{x} x + \beta_{y} y (1-I) \leq C \forall i \]
- \[ -\beta_{x} x - \beta_{y} y (1-I) \leq C - e \forall i \]
- \[ I_{i} - I_{i} - I_{j} \leq 0 \forall i, j \]
- \[ I_{i} - I_{j} \leq 0 \forall i, j \]
- \[ x_{i} = 0, 1 \forall i \]

where \( n \) is a number of customers; \( m \) is a number of dimensions in a feature space defined by demographic data for the number of customers; \( x_{i} \) is a length-\( m \) coordinate vector of customer \( i \) in the feature space for \( i = 1 \ldots n \); \( d_{ij} \) is a distance between customers \( i \) and \( j \) in transaction space according to a pre-selected distance metric; \( C \) is a large constant; \( e \) is a small constant; \( I_{i} \) is an indicator variable of customer \( i \), which is one if customer \( i \) is in cluster \( 1 \) (one side of the optimum hyperplane), and zero if the customer is in cluster \( 2 \) (the other side of the hyperplane) in the feature space; \( I_{i} \) is an indicator variable for customer pair \((i,j)\), which is equal to one if \( i \) and \( j \) are in the same cluster, and zero if they are in different clusters; \( \beta_{x} \) is a length-\( m \) direction vector in the feature space that defines the direction of the dividing hyperplane; and \( \beta_{y} \) is a scalar intercept of the dividing hyperplane.

12. A non-transitory computer-readable medium, comprising a plurality of instructions, that in response to being executed, result in a computing device:

- iteratively splitting a plurality of customers into a plurality of clusters based on purchase history data and demographic data for the plurality of customers, wherein each iteration comprises selecting a cluster of customers and splitting the selected cluster until a stopping rule is satisfied; and
- tailoring services provided to a customer based on a cluster from the plurality of clusters in which the customer resides.

13. The non-transitory computer-readable medium of claim 12, further comprising instructions that result in the computing device splitting the selected cluster by solving an Integer Program that accounts for the purchase history data and demographic data for the selected cluster.

14. The non-transitory computer-readable medium of claim 12, further comprising instructions that result in the computing device:

- selecting a cluster that has a population greater than a limit specified by the stopping rule; and
- ceasing the iteratively splitting in response to determining that no cluster of the plurality of clusters exceeds the limit specified by the stopping rule.

15. The non-transitory computer-readable medium of claim 12, further comprising instructions that result in the computing device:

- selecting a cluster that has the largest population of the plurality of clusters; and
- ceasing the iteratively splitting in response to determining the plurality of clusters comprises a quantity of clusters desired by the stopping rule.

16. A computing device, comprising an electronic database comprising demographic data and purchase history data for a plurality of customers; and a processor configured to:

- iteratively split a plurality of customers into a plurality of clusters based on purchase history data and demographic data for the plurality of customers, wherein each iteration comprises selecting a cluster of customers and splitting the selected cluster until a stopping rule is satisfied; and
- tailor services provided to a customer based on a cluster from the plurality of clusters in which the customer resides.

17. The computing device of claim 16, wherein the processor is further configured to split the selected cluster by solving an Integer Program that accounts for the purchase history data and demographic data for the selected cluster.

18. The computing device of claim 16, wherein the processor is further configured to:

- select a cluster that has a population greater than a limit specified by the stopping rule; and
- cease further splitting in response to determining that no cluster of the plurality of clusters exceeds the limit specified by the stopping rule.

19. The computing device of claim 16, wherein the processor is further configured to:

- select a cluster that has the largest population of the plurality of clusters; and
- cease further iteratively splitting in response to determining the plurality of clusters comprises a quantity of clusters desired by the stopping rule.