



Improving the Customer Grouping Algorithm for Spare-Parts Distribution Using k-Means Tools

M.M. Sepehri* & M. Kargari

Mohammad Mehdi Sepehri, Associate Professor, Industrial Engineering Department, Tarbiat Modares University
Mehrdad Kargari, Ph.D. Student, Industrial Engineering Department, Tarbiat Modares University

Keywords

Customer grouping algorithm,
Customer Relationship
Management, K-means,
Similarity function,
Spare-parts distribution

ABSTRACT

Customer classification using k-means algorithm for optimizing the transportation plans is one of the most interesting subjects in the Customer Relationship Management context. In this paper, the real-world data and information for a spare-parts distribution company (ISACO) during the past 36 months has been investigated and these figures have been evaluated using k-means tool developed for spare-part demand similarity function in different regions of the country. Similarity function for customer behavior in different regions has been defined. Based on this function and with help of k-means algorithm, customers have been grouped and similar customers were put in the same groups. Customer similarity function has been developed through 5 steps and has been defined individually based on each factors of Euclidean distance, customer's order time and bulk value of the order. Then, these three factors have been combined and DCB function has been defined. In the final step, different weights have been allocated to different years and seasons and BCD function has been improved. The grouping process has been improved for three functions of Euclidean distance, DCB and BCD. This process was executed using the R software and the improved BCD function was recognized as the optimum grouping function. Then, using DMT model, customer behavior has been analyzed at each part and the proper distribution policies have been defined. Results indicate a significant cost reduction (32%) in spare-parts distribution costs for ISACO.

© 2012 IUST Publication, IJIEPM. Vol. 23, No. 2, All Rights Reserved

*
Corresponding author. Mohammad Mehdi Sepehri
Email: mehdi.sepehri@modares.ac.ir



بهبود الگوریتم خوش بندی مشتریان برای توزیع قطعات یدکی با رویکرد داده کاوی (k-means)

محمد مهدی سپهری* و مهرداد کارگری

چکیده:

خوش بندی مشتریان با رویکرد داده کاوی برای بهینه سازی برنامه حمل و نقل یکی از مباحث مطرح در حوزه مدیریت ارتباط با مشتریان است. در این مقاله داده ها و اطلاعات واقعی توزیع قطعات یدکی شرکت ایساکو در طی ۳۶ ماه گذسته مورد بررسی قرار گرفته است و به کمک ابزار داده کاوی شباهت رفتار تقاضای قطعات یدکی توسط مشتریان در مناطق مختلف براساس ترکیب قواعد سنجیده شده است.تابع سنجش شباهت رفتار مشتریان در مناطق مختلف براساس ترکیب قواعد تعریف گردیده است. براساس این تابع و به کمک الگوریتم k-means عملیات خوش بندی انجام گرفته و مشتریان مشابه در یک خوش قرار گرفته اند. تابع سنجش شباهت رفتار مشتریان در پنج مرحله تشکیل شده است. تابع شباهت مشتریان براساس فواصل اقلیدسی (مکان استقرار آنها)، زمان تقاضای مشتریان و مقدار ارزش حجمی تقاضای مشتریان به صورت جداگانه تعریف شده است. سپس این سه عامل با هم ترکیب شده و تابع DCB به وجود آمده است. در مرحله نهایی با درنظر گرفتن وزن های مختلف برای سال ها و فصول گوناگون تابع BCD بهبود داده شده است. عملیات خوش بندی به وسیله سه تابع فاصله اقلیدسی، تابع DCB و تابع BCD بهبود یافته در نرم افزار R انجام و تابع BCD بهبود یافته به عنوان بهترین تابع برای خوش بندی انتخاب شده است. سپس با استفاده از مدل DTM رفتار هر بخش تحلیل شده و سیاست های توزیع مناسب برای آن بخش تبیین شده است. نتایج حاصل بیانگر کاهش ۳۲ درصد هزینه های توزیع در شرکت ایساکو می باشد.

کلمات کلیدی

الگوریتم خوش بندی مشتریان، مدیریت ارتباط با مشتریان، داده کاوی (k-means)، تابع شباهت، توزیع قطعات یدکی

دیدگاهی کلی و سطح بالا از تمام بانک اطلاعاتی مشتریان ارائه داده و به صاحبان شرکت های توزیع امکان اعمال رفتار و سیاست های متفاوت و مناسب به مشتریان هربخشش را می دهد. در حالت ایده آل هر سازمان باید هر یک از مشتریان را بطور کامل بشناسد، ولی این کار امکان پذیر نیست و در واقع خوش بندی این امکان را فراهم می آورد تا مشتریان که شبیه به هم هستند در یک بخش قرار گیرند. معیار شباهت براساس نوع صنعت و کسب و کار متفاوت است. در این صورت مدیریت و شناخت این بخش ها بسیار ساده تراز شناخت تک تک مشتریان است. خوش بندی به شیوه های مختلفی صورت می گیرد. در برخی تحقیقات صورت گرفته در این حوزه، با در نظر گرفتن سهم سود ایجاد شده، سود بالقوه و تعریف سود آوری مشتری یک مدل LTV پیشنهاد شده

۱. مقدمه

خوش بندی مشتریان براساس مناطق جغرافیایی یکی از مباحث مطرح در حوزه مدیریت ارتباط با مشتریان است. خوش بندی در واقع شکستن جمعیت زیاد مشتریان به بخش های مختلف است. بطوریکه مشتریان موجود در هر بخش به یکدیگر شبیه و مشتریان بخش های مختلف با یکدیگر متفاوت هستند. خوش بندی

تاریخ وصول: ۸۹/۶/۹

تاریخ تصویب: ۹۰/۴/۷

*تویینده مسئول مقاله: دکتر محمد مهدی سپهری، عضو هیئت علمی دانشگاه تربیت مدرس، [mehdi.sepehri@modares.ac.ir](mailto:mehdii.sepehri@modares.ac.ir)
مهرداد کارگری، دانشجوی دکتری دانشگاه تربیت مدرس، M_kargari@modares.ac.ir

در قسمت دوم روش خوش بندی مشتریان بیان شده است . این قسمت شامل آماده داده ها، معنی توابع شباهت ، الگوریتم خوش بندی k-means براساس توابع شباهت ، ارزیابی کیفیت خوش بندی و بهبود توابع خوش بندی بصورت ترکیبی میباشد . قسمت سوم مدل DTM برای ارزیابی خوش بندیها را بیان میکند . مطالعه موردی در قسمت چهارم بیان شده است . این قسمت شامل مقایسه میزان تراکم در حالات مختلف ، کیفیت خوش بندیها در حالات مختلف و تحلیل خوش بندیها در مدل DTM میباشد . نتایج و تحقیقات آینده در قسمت پنجم بیان شده است .

۲. متدولوژی بخش بندی مشتریان

در این مقاله ابتدا مشتریان بخش بندی میشوند و پس از ارزیابی و بهبود تابع فاصله ، سیاست های رفتاری مناسب برای هر بخش تبیین می گردد . اولین مرحله مانند هر پروژه داده کاوی دیگر آماده سازی داده هاست . سپس تابع DCB برای سنجش تفاوت رفتار مشتریان براساس مفاهیم قواعد انجمان تعریف شده است و با جایگزینی این تابع در k-means خوش بندی صورت گرفته است . بهبود خوش بندی صورت گرفته نسبت به عملکرد تابع فاصله اقلیدسی در الگوریتم k-means با استفاده از تابع سنجش تراکم خوش بندی نشان داده است . تعداد خوش بندیها با استفاده از تابع سنجش کیفیت خوش بندی تعیین گردیده است . مراحل ذکر شده در ادامه به طور کامل تشریح شده اند .

۲-۱. آماده سازی داده ها

در این مرحله داده های بانک طلاعاتی با توجه به اهداف و الگوریتم های بکار گرفته شده آماده سازی شده اند . مفاهیم و پارامترهای مورد استفاده در تابع شباهت عبارتند از :
 G : مجموعه همه گروه قطعات یدکی شرکت (حجم و ابعاد یکسان)
 T^C : بانک اطلاعات مشتریان
 t^C : یک رکورد از جدول T^C است، هر رکورد این جدول شامل کد مشتری، نام مشتری، مختصات جغرافیایی مشتری، نوع و گروه کالای درخواستی، میزان تقاضا، شهر مشتری، زمان تقاضا، ارزش تقاضا و حجم تقاضا میباشد .

C_i شهر i

g_{ia} : گروه کالای مورد نیاز a در شهر i

m_{ia} : مجموع ارزش حجمی کالای درخواستی گروه a در شهر i برای مشاهده رفتار مشتریان نیاز به بازیابی اطلاعات همه مشتریان و کل ارزش حجمی گروه کالای درخواستی توسط آنها در طول پانزده سال گذشته است .

$Citycode_i$: کد شهر i

$goodgorup_i$: مجموعه گروه کالاهایی موردنیاز شهر i

است و براساس ارزش فعلی، ارزش بالقوه و میزان وفاداری، مشتریان بخش بندی شده اند . [۵] سایون و همکاران در سال ۲۰۰۶ با درنظر گرفتن اهمیت ساخت مشتری در جهت ایجاد ارتباط بلندمدت، کسب وفاداری و سودآوری بیشتر مشتری، ارزش هر مشتری را تعیین نموده و سپس با در نظر گرفتن ارزش مشتریان، آنها را بخش بندی نموده و استراتژی های مناسب برای هر بخش را تبیین نموده اند [۶].

تسای و چیو در سال ۲۰۰۴ به بسط متداوله جدید برای بخش بندی بازار بر مبنای متغیرهای خاص از جمله اقلام خریداری شده و میزان درآمد مرتبط با آنها با در نظر گرفتن تراکنش های قبلی مشتریان پرداخته و بدین ترتیب بازار را بخش بندی نمودند [۷].

در برخی مطالعات صورت گرفته در حوزه بخش بندی بازار از روش های خوش بندی مدل FUZZY k-means برای بخش بندی مشتریان استفاده شده است [۸].

همچنین در دیگر تحقیقات مشتریان ، براساس معیار تازگی ، تکرار و ارزش پولی به گروه های همگن بخش بندی شده اند ، سپس سیاست های بهینه بازار یابی برای هر بخش تدوین شده است [۹]. همانطور که اشاره شد بخش بندی مشتریان به شیوه های گوناگون صورت می گیرد . در شرکت ایساکو میزان شباهت رفتار مشتریان با توجه به نوع داده ها و اطلاعات موجود از مشتریان به محل استقرار آنها، شهر مشتری ، گروه کالای مورد نیاز آنها، میزان تقاضای آنها و زمان تقاضای آنها وابسته است . بر این اساس در این مقاله از معیارهای شهر ، فاصله اقلیدسی ، گروه کالای مورد نیاز ، میزان و زمان تقاضا برای استخراج تابع شباهت استفاده شده است . ابتدا دو معیار گروه کالای موردنیاز و حجم تقاضا با استفاده از الگوریتم k-means برای خوش بندی مشتریان استفاده شده است . سپس تابع فاصله اقلیدسی به صورت جداگانه در الگوریتم k-means خوش بندی مشتریان منظور شده است . برای بهبود تابع شباهت معیارهای شهر ، فاصله اقلیدسی، گروه کالای موردنیاز ، میزان و زمان تقاضا با هم ترکیب شده و مجدداً با الگوریتم k-means برای خوش بندی مشتریان بکار رفته است . پس از تعیین بهترین خوش بندی براساس معیار تراکم خوش بندی ، بهترین تعداد خوش بندی را در نظر گرفتن معیار سنجش کیفیت خوش بندی محاسبه شده است .

هدف این مقاله بهینه سازی برنامه توزیع قطعات یدکی شرکت ایران خودرو با رویکرد داده کاوی و خوش بندی فروشگاه های مجاز در نقاط مختلف کشور میباشد . برای این منظور از سه تابع شباهت فاصله اقلیدسی، میزان تقاضای مشتریان و زمان تقاضای مشتریان بصورت همزمان استفاده شده است . سپس با در نظر گرفتن ورن های مختلف برای فضول سال تابع شباهت ترکیبی بهبود داده شده است . ساختار این مقاله در ادامه بشرح زیر است .

همچنین اگر رکورد t_j^c برای مشتریان شهر i شامل $(Citycode_j, goodgorup_j, moneyset_j)$ باشد و گروه کالا و ارزش حجمی هر گروه کالا به ترتیب برابر $\{g_{j1}, g_{j2}, \dots, g_{ji}\}$ باشد شباهت دو شهر i نسبتی تقاضای مشتریان دو شهر به کمک رابطه (۳) زیر محاسبه می شود.

$$\left[\begin{array}{ll} sim(C_i, C_j) = \frac{\sum_{a=1}^s \sum_{b=1}^s m_{ia} \times m_{jb} \times \delta(g_{ia}, g_{jb})}{\sum_{a=1}^s \sum_{b=1}^s m_{ia} \times m_{jb}} & C_i \neq C_j \\ sim(C_i, C_j) = 1 & C_i = C_j \end{array} \right] \quad (3)$$

متعاقباً فاصله دو شهر i ، C_i یا تفاوت رفتار تقاضای مشتریان دو شهر به کمک رابطه (۴) محاسبه می شود.

$$\left[\begin{array}{ll} DCB(C_i, C_j) = 1 - sim(C_i, C_j) & C_i \neq C_j \\ DCB(C_i, C_j) = 0 & C_i = C_j \end{array} \right] \quad (4)$$

دو معادله بالا علاوه بر در نظر گرفتن گروه کالای مورد نیاز مشتریان و ارزش حجمی تقاضای مشتریان هر شهر در میزان کاهش هزینه های حمل و نقل شرکت ایساکو تاثیرگذار است. برای مثال اگر گروه کالای A معمولاً با گروه کالای B ارسال شود و کمتر با گروه کالای C ارسال شود، وابستگی تقاضای دو گروه کالای A, B باید بیشتر از دو گروه کالای C باشد. همچنین نادیده گرفتن این وابستگی ها و یکسان گرفتن رفتار همه گروه کالاهای باعث ایجاد شباهت یک طرفه می شود. در صورتی که باید برای کامل شدن محاسبات از شباهت دو طرفه استفاده نمود. به راین اساس در این مقاله از شباهت دو طرفه گروه کالاهای استفاده شده است.

۳-۲. خوش بندی

خوش بندی وظیفه تقسیم یک گروه ناهمگن را به چندین زیر گروه که به اصطلاح به آن خوش می گویند را به عهده دارد. در این تحقیق از الگوریتم k-means برای خوش بندی مشتریان استفاده شده است. برای سنجش فاصله مشتریان دو شهر در الگوریتم k-means ازتابع DCB تابع فاصله اقلیدسی و سپس با ترکیب این دو تابع و تاثیر داده زمان سفارش مشتریان تابع DCB بهبود داده شده است.

$$goodgorup_i = \{g_{ia} / g_{ia} \in G\}$$

: مجموعه ارزش حجمی کالای درخواستی مشتریان شهر i
 $moneyset_i = \{m_{ia} | a=1,2,\dots, | / goodgorup_i //\}$

$GISplan_i$: مختصات جغرافیایی محل استقرار مشتری i
 $Demandtime_i$: زمان و تاریخ سفارش مشتری

بنابراین کالای درخواستی مشتریان شهر i را می توان به صورت t^C_i نشان داده و در بانک اطلاعات مشتریان (T^C) ذخیره نمود.
 $T^C_i = \{Citycode_i, goodgorup_i, moneyset_i, Location plan_i, Demandtime_i\}$

۲-۲. تابع فاصله تفاوت رفتار تقاضای مشتریان (DCB)

برای اندازه گیری تفاوت رفتار مشتریان در هر شهر، ابتدا وابستگی هر گروه کالا با گروه کالاهای دیگر براساس مفهوم پشتیبانی در قواعد انجمن محاسبه شده است. برای محاسبه این تفاوت از رابطه (۱) استفاده شده است.

$$\|t^c \in T^C_i / t^c \text{ Contains } \{g_{ia}, g_{ib}\} \| \quad (1)$$

$$S(g_{ia}, g_{ib}) = \frac{1}{\|T^C_i\|} \quad (2)$$

$S(g_{ia}, g_{ib})$: نسبت تعداد تراکنش هایی که در آن دو گروه کالای a, b از شهر i وجود دارد به کل تعداد تراکنش های شهر i را نشان می دهد. اگر دو گروه کالا به ندرت با هم سفارش دهنند در نتیجه مقدار پشتیبانی آنها خیلی کم خواهد شد [۲].

درتابع DCB فاصله براساس میزان شباهت گروه کالا می باشد و نیاز به در نظر گرفتن وابستگی دوطرفه انواع گروه قطعات یدکی دارد. لذا برای محاسبه وابستگی دو طرفه گروه قطعات یدکی ضریب (۷) به کمک رابطه (۲) محاسبه می نماییم [۱۳].

$$S \{g_{ia}, g_{ib}\} \quad (2)$$

$$\gamma(g_{ia}, g_{ib}) = \frac{1}{S(g_{ia}) + S(g_{ib}) - S(g_{ia}, g_{ib})} \quad (2)$$

مقدار $\gamma(g_{ia}, g_{ib})$ بین صفر و یک است. اگر $g_{ia} = g_{ib}$ باشد آنگاه $\gamma(g_{ia}, g_{ib}) = 1$ می باشد. حال میزان شباهت تقاضای مشتریان دو شهر به صورت زیر ارزیابی می شود.

اگر رکورد t^c_i برای مشتریان شهر i شامل $(Citycode_i, goodgorup_i, moneyset_i)$ باشد و گروه کالا و ارزش حجمی هر گروه کالا به ترتیب برابر $\{g_{il}, g_{i2}, \dots, g_{is}\}$ باشد $moneyset_i = \{m_{il}, m_{i2}, \dots, m_{is}\}$ و g_{i2}, \dots, g_{is}

محل استقرار مشتریان شهر z باشد. فاصله یا عدم شباهت مشتریان دو شهر با استفاده فاصله اقلیدسی به کمک رابطه (۵) محاسبه می‌شود.

$$Dis(C_i, C_j) = \sqrt{\sum_{l=1}^L [(X_{il} - X_{jl})^2 + (Y_{il} - Y_{jl})^2]} \quad (5)$$

در این تحقیق ابتدا تابع فاصله در الگوریتم k-means مطابق تابع فاصله اقلیدسی به صورت معادله فوق در نظر گرفته شده است و سپس تابع DCB به دو صورت ساده و ترکیبی (بهبود یافته) در نظر گرفته شده است.

۵-۱. الگوریتم خوشبندی k-means تابع DCB
در این خوشبندی در الگوریتم k-means برای محاسبه فاصله از معادله (۴) استفاده شده است. فاصله مشتریان هر شهر به صورت دوبعدی محاسبه شده است و در ماتریس به شکل زیر قرار گرفته است. در این ماتریس n تعداد مشتریان است، d_{ij} نشان دهنده فاصله بین مشتریان C_i و C_j ($i, j \leq n$) است. با توجه به معادله (۴) قطر اصلی این ماتریس یعنی فاصله هر مشتری با خودش صفر است.

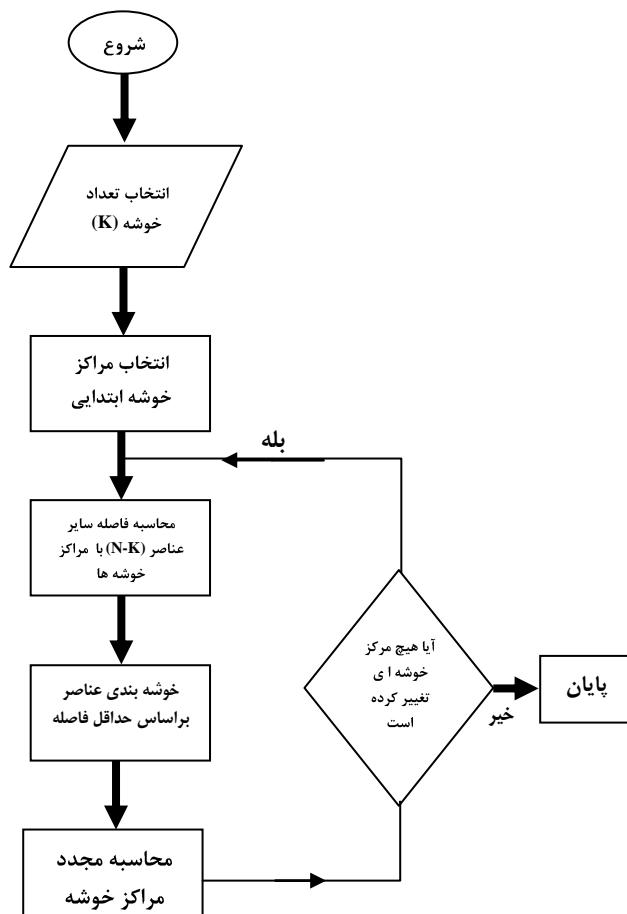
$$Dis = \begin{bmatrix} 0 & d_{i2} & d_{ij} & d_{in} \\ d_{21} & 0 & d_{2j} & d_{2n} \\ d_{i1} & d_{i2} & 0 & d_{in} \\ d_{n1} & d_{n2} & d_{nj} & 0 \end{bmatrix} \quad (6)$$

این ماتریس به عنوان ورودی به الگوریتم k-means در نرم افزار R داده شده است. در واقع تابع فواصل یا عدم شباهت بین مشتریان یکبار براساس تابع فاصله اقلیدسی، یکبار براساس تابع فاصله DCB و یکبار هم بر اساس DCB بهبود یافته محاسبه شده است.

۶-۱. ارزیابی کیفیت خوشبندی
معیارهای زیادی برای محاسبه کیفیت خوشباهای حاصل از خوشبندی وجود دارد. در خوشبندی مطلوب باید شباهت درون خوشباهای حداکثر و عدم شباهت بین خوشباهای حداکثر شود. کلیه معیارهای سنجش کیفیت مبتنی بر این دو مفهوم هستند. برخی از آنها فقط می‌توانند همگنی درون خوشباهای یا ناهمگنی بروند خوشباهای را تضمین کنند ولی برخی دیگر هردو همگنی بروند خوشباهای را تضمین می‌کنند [۱۳].

۴-۲. الگوریتم خوشبندی k-means براساس تابع فاصله اقلیدسی

در الگوریتم k-means ابتدا K عضو به صورت تصادفی از میان N عضو باقی مانده به نزدیکترین خوشبندی تخصیص می‌یابند. بعد از تخصیص همه اعضاء مراکز خوشبندی مجدد محاسبه می‌شوند و اعضاء با توجه به مراکز جدید به خوشبندی تخصیص می‌یابند و این کار تا زمانی که مراکز هر خوشبندی ادامه می‌یابد خلاصه این مراحل در نمودار (۱) نشان داده شده است [۲].



شکل ۱. الگوریتم خوشبندی k-means

اگر رکورد i برای مشتریان شهر C_i شامل و

$(Citycode_i, Location plan_i, Demondtime_i)$

$Location plan =$

$[(X_{i1}, Y_{i1}), (X_{i2}, Y_{i2}), \dots, (X_{iL}, Y_{iL})]_i$

محل استقرار مشتریان شهر i باشد و رکورد j برای مشتریان شهر C_j شامل

$(Citycode_j, Location plan_j, Demondtime_j)$

و

$Location plan_i, \dots, (X_{iL}, Y_{iL})] = [(X_{j1}, Y_{j1}), (X_{j2}, Y_{j2})]$

$$\delta_{nm} = sim(C^n, C^m) \quad (12)$$

معادله (۱۰)، μ_n را به عنوان میانگین شباهت بین مرکز خوشه C^n و همه مشتریان خوشه O^n تعریف می‌کند.
معادله (۱۱) توضیح می‌دهد که μ_m میانگین شباهت بین مرکز خوشه C^m و همه مشتریان در خوشه O^m است. معادله (۱۲) را به عنوان شباهت C^n و C^m تعریف می‌کند.

مقدار $p(k)$ با توجه بهتابع کیفیت خوشه بندی در معادله (۸)، برای یک محدوده پیشنهادی (حد بالای t و حد پائین s) محاسبه شده است. و به ازای هر تعداد خوشه‌ای که $p(k)$ حداکثر گردیده است، مقدار K به عنوان تعداد خوشه در نظر گرفته شده است. در این تحقیق محدوده پیشنهادی K با توجه به نظر افراد خبره به صورت $30 \leq k \leq 50$ در نظر گرفته شده است.

$$k = \max_{s \leq K \leq t} \{P(k)\} \quad (13)$$

بدین ترتیب می‌توان تعداد خوشه‌ای را انتخاب نمود که به ازای آن فاصله بین مراکز خوشه‌ها و شباهت مراکز خوشه با اعضا درون هر خوشه حداکثر گردد [۷].

۲-۸. بهبود تابع DEB

برای اندازه گیری میزان شباهت مشتریان /مناطق باید همه عوامل شباهت به صورت هم زمان در نظر گرفته شود. در تابع DCB دو عامل گروه کالای مشابه و ارزش حجمی در نظر گرفته شده‌اند. در تابع فاصله اقلیدسی عامل فاصله مکانی بین دو مشتری / منطقه در نظر گرفته شده است. این سه عامل به کمک رابطه (۱۴) ترکیب شده‌اند.

$$\alpha(C_i, C_j) = \frac{Maxd(C_i, C_j) - d(C_i, C_j)}{Maxd(C_i, C_j)} \quad (14)$$

مقدار α همواره بین صفر و یک می‌باشد. اگر $C_i = C_j$ باشد مقدار $\alpha = 1$ است و اگر دو مشتری /منطقه اگر C_i, C_j حداکثر فاصله را داشته باشند مقدار $\alpha = 0$ است. میزان شباهت دو مشتری /منطقه اگر C_i, C_j با توجه به سه عامل فاصله مکانی، گروه کالای مشابه و میزان ارزش حجمی سفارشات به صورت رابطه (۱۵) حاصل می‌شود.

یکی از روش‌های ارزیابی عملکرد خوشه بندی، روش سنجش تراکم درون خوشه‌ای است. این معیار میزان تراکم خوشه‌ها را زمانی که تعداد خوشه‌ها ثابت است نشان می‌دهد. در این معیار باید متغیرها در یک محدوده باشند. با توجه به این که در این تحقیق داده‌ها گستته هستند برای محاسبه تراکم یک خوشه، فاصله مشتریان (مناطق) درون یک خوشه باید با مرکز خوشه سنجیده شود. تراکم خوشه I_n ، $I_n = \sum_{c_i \in O^n} Dist(C_i, C^n)$ به صورت رابطه (۷) تعریف شده است:

$$I_n = \sum_{c_i \in O^n} Dist(C_i, C^n) \quad (7)$$

سپس $F(K)$ برای K خوشه برابر است با:

$$F(K) = \frac{\mathbf{1}}{K} \sum_{n=1}^k \sum_{c_i \in O^n} Dist(C_i, C^n) \quad (8)$$

در واقع $F(K)$ میانگین مربعات فاصله اقلیدسی بین هر مشاهده و میانه خوشه‌ای که به آن تعلق دارد، است هرچه، $F(K)$ کمتر باشد نشان می‌دهد که خوشه‌ها متراکم‌ترند و خوشه بندی بهتر صورت گرفته است [۸].

۲-۷. انتخاب بهترین تعداد خوشه

همانطور که بیشتر اشاره شد، بهترین خوشه بندی، خوشه بندی است که مجموع تشابه بین مرکز خوشه و همه شهرهای عضو خوشه مربوطه حداکثر و مجموع تشابه بین مراکز خوشه را حداقل نماید.

اگر $O = \{C^n | n = 1, \dots, k\}$ مجموعه مراکز خوشه‌ها و C^n مراکز خوشه‌ها باشد و $O^n = \{C_i | i = 1, \dots, ||T^n - O||\}$ مجموعه باقیمانده مشتریان (مناطق) که به عنوان مرکز خوشه انتخاب نشده اند و T^n مجموعه کلیه شهرهایی که خوشه بندی روی آنها صورت گرفته است باشد، کیفیت نتایج خوشه بندی با k خوشه می‌تواند به صورت معادله (۹) تعریف شود:

$$p(K) = \frac{\mathbf{1}}{K} \sum_{n=1}^k (\min_{1 \leq m \leq k, m \neq n} \left\{ \frac{\mu_n + \mu_m}{\delta_m} \right\}) \quad (9)$$

$$\mu_n = \frac{\mathbf{1}}{\|O^n\|} \sum_{c_i \in O^n} sim(C_i, C^n) \quad (10)$$

$$\mu_m = \frac{\mathbf{1}}{\|O^m\|} \sum_{c_j \in O^m} sim(C_j, C^m) \quad (11)$$

برمبنای این معیار ، ارزش هر مشتری / منطقه را می‌توان به وسیله رابطه (۱۸) تعیین نمود.

$$V(C_i) = W^D \cdot D(C_i) + W^T \cdot T(C_i) + W^M \cdot M(C_i) \quad (18)$$

در معادله (۱۸) مقادیر $D(C_i)$ ، $T(C_i)$ و $M(C_i)$ به ترتیب بیانگر امتیازات مشتری / منطقه (C_i) با توجه به معیارهای D ، T و M است. W^D ، W^T و W^M اهمیت وزن ها برای معیارهای D ، T و M را به ترتیب نشان می‌دهد. بعلاوه $W^M + W^T + W^D = 1$ می‌باشد [۹]. امتیازات "عموماً" به نوع کار بردها و رویکرد امتیاز دهی وابسته است. امتیازاتی را که از بانک اطلاعاتی استخراج می‌نماییم قبل از محاسبه ارزش هر مشتری / منطقه ابتدا نرمال می‌گردد. بنابر این $(D(C_i), T(C_i), M(C_i))$ را به وسیله روابط (۱۹) تعریف می‌کیم.

$$\begin{aligned} D(C_i) &= \frac{Q^D - Q^D_{Min}}{Q^D_{Max} - Q^D_{Min}} \\ T(C_i) &= \frac{Q^T - Q^T_{Min}}{Q^T_{Max} - Q^T_{Min}} \\ M(C_i) &= \frac{Q^M - Q^M_{Min}}{Q^M_{Max} - Q^M_{Min}} \end{aligned} \quad (19)$$

در روابط (۱۹) مقادیر Q^D و Q^T و Q^M مقدار اصلی امتیازات مشتری / منطقه (C_i) مطابق با تعریف D ، T و M هستند. مقادیر Q^D_{Min} و Q^D_{Max} حداقل مقدار D و Q^T_{Min} و Q^T_{Max} حداکثر مقدار T و Q^M_{Min} و Q^M_{Max} حداکثر مقدار M همه مشتریان / مناطق را نشان می‌دهند. سود آوری هر خوشه Q^n با محاسبه ارزش همه مشتریان / مناطق خوشه n ام حاصل می‌شود. این ارزش به صورت معادلات (۲۰) و (۲۱) محاسبه می‌شود:

$$V(O^n) = W^D \cdot W(O^n) + W^T \cdot W(O^n) + W^M \cdot M(O^n) \quad (20)$$

$$\begin{aligned} W(O^n) &= \frac{\sum D(C_i)}{\| O^n \|} & T(O^n) &= \frac{\sum T(C_i)}{\| O^n \|} \\ M(O^n) &= \frac{\sum M(C_i)}{\| O^n \|} \end{aligned} \quad (21)$$

در معادلات (۲۱) مقادیر $D(O^n)$ ، $T(O^n)$ و $M(O^n)$ بیانگر امتیازات خوشه n با توجه معیارهای D ، T و M است. پس از مشخص شدن سودآوری هر خوشه به راین اساس مرتب می‌شود و

$$\left[\begin{array}{ll} sim(C_i, C_j) = & \begin{cases} \frac{\sum\limits_{a=1}^s \sum\limits_{b=1}^t m_{ia} \times m_{jb} \times \text{Min}\{\delta(g_{ia}, g_{jb}), \alpha(C_i, C_j)\}}{\sum\limits_{a=1}^s \sum\limits_{b=1}^t m_{ia} \cdot m_{jb}} & C_i \neq C_j \\ 1 & C_i = C_j \end{cases} \\ \end{array} \right] \quad (15)$$

متعاقباً فاصله دو شهر / منطقه یا تفاوت رفتار آنها به کمک رابطه (۴) محاسبه می‌شود. همچنین برای محاسبه شباهت کلیه داده‌های ۳۶ ماه گذشته با یکدیگر جمع شده‌اند و زمان که عامل مهمی است با مجتمع کردن داده‌ها "عملماً" نادیده گرفته شده است و وابستگی بین اقلام صادرشده در سالها و فصل‌های مختلف در نظر گرفته نمی‌شود. لذا در این مرحله برای محاسبه ماتریس فاصله ، داده‌های هر ماه جداگانه مورد بررسی قرار گرفته و برای هر ماه ماتریس فاصله جداگانه محاسبه شده است . سپس با استفاده از روش دلفی و نظرات خبرگان ضرایب وزنی هرماه محاسبه گردیده و ماتریس فاصله کل برای ۳۶ ماه برحسبتابع بهبود یافته DCB محاسبه شده است. ماتریس فاصله در ماه آخر ماتریس فاصله در ماه اول

$$\left[\begin{array}{cccc} d_{11} & d_{i2} & d_{ij} & d_{1n} \\ d_{21} & d_{22} & d_{2j} & d_{2n} \\ d_{i1} & d_{i2} & d_{ij} & d_{in} \\ d_{n1} & d_{n2} & d_{nj} & d_{nn} \end{array} \right] \dots \left[\begin{array}{cccc} d_{11} & d_{i2} & d_{ij} & d_{1n} \\ d_{21} & d_{22} & d_{2j} & d_{2n} \\ d_{i1} & d_{i2} & d_{ij} & d_{in} \\ d_{n1} & d_{n2} & d_{nj} & d_{nn} \end{array} \right] \quad (16)$$

مقدار ماتریس فاصله کل ۳۶ ماه (DIS) از فرمول (۱۷) محاسبه می‌شود.

$$Dis = \sum_{Y=1}^Y W_Y \cdot D_Y \quad (17)$$

D_Y ماتریس فاصله ماه y و W_Y ضریب ماه y بین $1 \leq Y \leq 36$ می‌باشد.

۳. مدل DTM برای ارزیابی خوشه‌ها

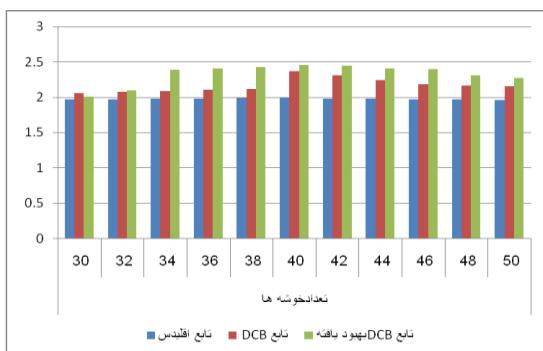
مدل DTM ارزش مشتریان / مناطق را برمبنای سه معیار فاصله مکانی (D) ، فاصله زمانی بین سفارشات (T) و ارزش حجمی (M) اندازه گیری می‌کند. متغیر فاصله مکانی، میانگین فاصله بین مشتریان و مرکز خوشه را اندازه می‌گیرد. متغیر فاصله زمانی بین سفارشات، متوسط زمانی بین سفارشات در هر خوشه را نشان می‌دهد. متغیر ارزش حجمی بیانگر متوسط حجم کالای در خواستی در هر خوشه می‌باشد. مدل DTM به سازمان‌ها کمک می‌کند تا بتوانند ارزش خوشه‌های خود را در این مدل مشخص نمایند و خوشه‌های هدف را جهت برنامه ریزی تعیین نموده و اقدامات لازم در هر خوشه را برای کاهش هزینه تعیین نمایند.

۴-۴. محاسبه تعداد خوشها

برای تعیین تعداد خوشها با استفاده از معادله (۹) مقدار $F(k)$ برای مقادیر مختلف (K) محاسبه شده است مقدار $F(k)$ حداکثر نشان دهنده تعداد مناسب خوشهاست . در این تحقیق $F(k)$ برای مقادیر $50 \leq K \leq 30$ محاسبه شده است که این بازه با توجه به نظر خبرگان و تعداد دادهها تعیین شده است . بدین منظور به کمک نرم افزار R و ماتریس فاصله محاسبه شده برای $30 \leq K \leq 50$ خوش بندی صورت گرفته است. کیفیت خوشها در به اره $30 \leq K \leq 50$ در جدول (۲) نشان داده شده است.

جدول ۲، کیفیت خوشها در بازه $30 \leq K \leq 50$

تعداد خوش	تابع اقلیدسی	تابع DCB	تابع DCB بهبود یافته
۳۰	۱/۹۷۰	۲/۰۵۱	۲/۰۱۱
۳۲	۱/۹۷۲	۲/۰۷۲	۲/۰۹۵
۳۴	۱/۹۷۸	۲/۰۸۰	۲/۳۹۰
۳۶	۱/۹۸۰	۲/۱۰۷	۲/۴۰۲
۳۸	۱/۹۸۳	۲/۱۱۱	۲/۴۳۰
۴۰	۱/۹۹۲*	۲/۳۶۹*	۲/۴۵۲*
۴۲	۱/۹۸۱	۲/۳۱۰	۲/۴۴۰
۴۴	۱/۹۷۵	۲/۲۴۲	۲/۴۰۱
۴۶	۱/۹۷۲	۲/۱۸۲	۲/۳۹۵
۴۸	۱/۹۶۵	۲/۱۶۱	۲/۳۱۲
۵۰	۱/۹۶۰	۲/۱۵۴	۲/۲۷۰



نمودار ۱. مقایسه کیفیت و خوشها

۴-۳. تحلیل خوشها در مدل DTM

برای تحلیل خوشها در مدل DTM باید معیارهای M,T,D برای مشتریان / مناطق و خوشها محاسبه گردد. مقدار متغیر فاصله اقلیدسی بین مشتریان / مناطق نشان دهنده فاصله اقلیدسی است که بین دو مشتری / منطقه وجود دارد. این متغیر برای هر مشتری / منطقه با توجه به رابطه

$$D(C_i) = \frac{Q^D - Q^D_{Min}}{Q^D_{Max} - Q^D_{Min}}$$

با ارزش ترین آنها مشخص شده و به شرکت در برنامه ریزی تعیین استراتژی توزیع هر خوش و خدمت رسانی به مشتریان هدف شرکت کمک می نماید.

۴. مطالعه موردی

مورد مطالعاتی در این تحقیق شرکت ایساکو می باشد. شرکت ایساکو در راستای برنامه های کلان گروه صنعتی ایران خودرو در پی توسعه و تقویت سیستم حمل و نقل قطعات یدکی و دستیابی به سهم بیشتر بازار هدف به شکل اثر بخش می باشد. این شرکت با بکارگیری دانش و فناوری جدید به دنبال ایجاد بستر مناسب جهت تقویت زیربنای لازم برای تسهیل و توسعه سیستم حمل و نقل شرکت است. مطالعه صورت گرفته برروی داده های ایساکو به مشتریان / مناطق شهر تهران در طی ۳۶ ماه گذشته است . این شرکت طیف وسیعی از کالاهای متنوع را بین مشتریان / مناطق توزیع می کند. تعداد اقلام مورد بررسی ۷۵ قلم کالای مهم بوده که در ۸ گروه دسته بندی شده اند . داده های شرکت ایساکو در طی سه سال گذشته با توجه به ۷۰۰ نمایندگی و ۸۰۰ فروشگاه شامل $=75000$ (50)(800) رکورد می باشد. هر مشتری حداقل هفتادی یک سفارش دارد و هر سال میانگین (۵۰) هفته فرض شده است.

۱-۴. خوش بندی مناطق / مشتریان

بعد از انجام مراحل آماده سازی، یکباره گنجایش و استخراج داده ها با فرمت مناسب با استفاده از الگوریتم k-means مناطق کشور خوش بندی شده اند. همانطور که توضیح داده شد خوش بندی یک بار بر اساس فاصله اقلیدسی، یک بار بر اساس تابع DCB و یکبار بر اساس تابع DCB بهبود یافته صورت گرفته است . تعداد خوشها برابر ۴۰ در نظر گرفته شده است که در ادامه علت این انتخاب توضیح داده شده است. تراکم خوشها حاصل شده در سه روش با توجه به معادله (۸) محاسبه شده است. هرچه مقدار $F(K)$ در معادله (۸) کمتر باشد میزان تراکم خوشها بیشتر است. جدول (۱) میزان تراکم خوشها با استفاده از سه روش پیشنهادی را نشان می دهد. در جدول زیر مشاهده می شود که تراکم خوشها در روش تابع DCB بهبود بیشتر از دوتابع دیگر است. یعنی این تابع نسبت به توابع دیگر عملکرد بهتری داشته است.

جدول ۱. مقایسه میزان تراکم خوشها از توابع مختلف

تراکم خوش	تابع فاصله خوش بندی
۱/۲۹۴	تابع فاصله اقلیدسی
۰/۱۱۹	تابع DCB
۰/۰۷۴	تابع DCB بهبود یافته

به اهمیت این متغیرها به رای متغیرهای M, T, D به ترتیب وزن‌های $0.5, 0.3, 0.2$ درنظر گرفته شده است. سپس با معادلات $(M(O^n), T(O^n), M(O^n))$ مقدار فاصله مکانی، فاصله زمانی و ارزش حجمی هر خوشة محاسبه و با استفاده از رابطه $V(O^n)$ ارزش هر خوشه تعیین می‌گردد. نتایج حاصل به صورت جدول (۳) آمده است:

محاسبه می‌شود. متغیر فاصله زمانی بین سفارشات مشتریان / مناطق نشان دهنده فاصله بین دو سفارش متوالی مشتریان / مناطق است. مقدار این متغیر برای هر مشتری / منطقه با توجه به رابطه $T(C_i)$ محاسبه می‌شود. متغیر ارزش حجمی کالاهای درخواستی نیز بیانگر میزان حجم کالای درخواستی توسط مشتری / منطقه می‌باشد که توسط رابطه $M(C_i)$ محاسبه می‌شود. وزن متغیرها براساس نظرات خبرگان تعیین گردیده است. با توجه

جدول ۳. ارزش مشتریان / مناطق هر خوش

شماره خوش	میانگین فاصله مکانی ($D(O'')$)	میانگین فاصله زمانی ($T(O'')$)	میانگین ارزش حجمی ($M(O'')$)	میانگین ارزش خوش	تعداد مشتری / منطقه در خوش
۱۷	.۰/۴۱۷	.۰/۲۲۱	.۰/۵۴۰	.۰/۵۳۲۸	۱۵
۲۹	.۰/۹۱۸	.۰/۱۵۱	.۰/۲۷۱	.۰/۵۵۸۵	۱۸
۸	.۰/۸۴۱	.۰/۲۷۴	.۰/۲۵۲	.۰/۵۵۳۱	۲۵
۳۰	.۰/۵۷۱	.۰/۱۷۱	.۰/۴۷۱	.۰/۴۳۱	۴۲
۱۸	.۰/۹۴۱	.۰/۲۴۸	.۰/۵۴۲	.۰/۰۳۳	۱۱
۳۱	.۰/۵۱۷	.۰/۲۱۵	.۰/۷۲۱	.۰/۴۶۷۲	۵
۷	.۰/۶۱۸	.۰/۲۱۸	.۰/۵۱۱	.۰/۴۷۶۶	۲
۲۷	.۰/۴۲۵	.۰/۴۱۸	.۰/۹۲۱	.۰/۵۲۲۱	۷
۹	.۰/۴۱۹	.۰/۴۷۱	.۰/۱۲۱	.۰/۴۳۵	۴۸
۳۲	.۰/۲۷۱	.۰/۵۷۱	.۰/۱۰۹	.۰/۳۴۸۶	۱۵
۱۹	.۰/۲۱۵	.۰/۲۴۸	.۰/۶۲۱	.۰/۳۰۶۱	۱۸
۴	.۰/۴۹۵	.۰/۵۷۱	.۰/۵۲۱	.۰/۵۲۳	۱۹
۲۶	.۰/۷۱۵	.۰/۶۹۱	.۰/۷۴۱	.۰/۷۱۳	۲۱
۱۰	.۰/۸۲۱	.۰/۵۷۱	.۰/۸۱۱	.۰/۷۴۴	۲۲
۳۳	.۰/۹۴۲	.۰/۸۲۱	.۰/۹۱۵	.۰/۹۰۰۳	۱۷
۳	.۰/۹۱۲	.۰/۲۸۱	.۰/۹۴۷	.۰/۷۲۹۷	۱۶
۳۴	.۰/۱۵۰	.۰/۲۱۵	.۰/۰۱۷	.۰/۱۴۲۹	۱۹
۱۱	.۰/۱۲۷	.۰/۲۵۲	.۰/۲۱۵	.۰/۱۸۲۱	۱۱
۳۵	.۰/۹۱۸	.۰/۹۴۷	.۰/۵۶۱	.۰/۸۵۴۳	۱۷
۲	.۰/۱۹۷	.۰/۱۵۷	.۰/۱۴۸	.۰/۱۹۵۲	۲۵
۲۰	.۰/۸۱۷	.۰/۸۱۵	.۰/۱۱۱	.۰/۶۷۵۲	۴۲
۱	.۰/۲۴۸	.۰/۲۴۹	.۰/۱۵۲	.۰/۲۲۹۱	۵۱
۲۱	.۰/۲۵۲	.۰/۰۵۰	.۰/۱۱۷	.۰/۱۶۴۴	۷۱
۱۲	.۰/۰۴۲	.۰/۰۴۸	.۰/۹۴۸	.۰/۴۷۵	۱۵
۴۰	.۰/۷۴۲	.۰/۰۲۵	.۰/۵۲۱	.۰/۴۸۲۷	۶۸
۲۲	.۰/۵۴۱	.۰/۲۱۵	.۰/۷۱۸	.۰/۴۷۸۶	۱۵
۵	.۰/۸۲۱	.۰/۱۸۷	.۰/۱۵۱	.۰/۴۹۶۸	۱۷
۲۴	.۰/۲۷۱	.۰/۹۴۱	.۰/۱۷۲	.۰/۴۵۲۲	۱۶
۳۹	.۰/۲۵۷	.۰/۱۵۱	.۰/۱۵۱	.۰/۲۰۴	۵
۱۴	.۰/۴۲۲	.۰/۲۴۸	.۰/۹۴۱	.۰/۴۷۳۶	۷
۳۶	.۰/۴۴۷	.۰/۶۴۷	.۰/۱۵۷	.۰/۴۴۹	۹
۶	.۰/۵۷۴	.۰/۱۴۸	.۰/۲۴۸	.۰/۳۸۱	۸
۳۷	.۰/۱۱۵	.۰/۰۰۵	.۰/۷۴۸	.۰/۲۰۸۶	۴
۲۳	.۰/۵۴۲	.۰/۱۵۶	.۰/۱۵۱	.۰/۳۴۸	۳
۱۶	.۰/۹۷۱	.۰/۱۷۸	.۰/۱۷۵	.۰/۲۲۳۹	۱۵
۲۵	.۰/۱۲۵	.۰/۲۴۹	.۰/۱۶۱	.۰/۱۶۹۴	۱۷
۳۸	.۰/۷۶۴	.۰/۳۴۸	.۰/۷۴۰	.۰/۶۳۴۴	۱۴
۱۳	.۰/۱۱۷	.۰/۷۱۵	.۰/۷۴۵	.۰/۴۲۲	۲۵
۲۸	.۰/۹۱۵	.۰/۸۴۸	.۰/۲۱۱	.۰/۷۵۴۱	۱۶
۱۵	.۰/۸۲۱	.۰/۲۴۱	.۰/۹۱۱	.۰/۶۶۵	۱۹

استراتژی‌های توزیع در هر خوش میانگین ارزش حجمی هر گروه کالا را به صورت جداگانه محاسبه و سپس بر میانگین ارزش

پس از تحلیل انجام شده برروی خوشه‌ها و محاسبه ارزش هر خوشه استراتژی‌های توزیع استخراج می‌گردد. پس از تعیین

بهره گرفت. همچنین می‌توان تحقیقات را ببروی گروه کالای خاص با شرایط مختلف انجام داد و نتایج را ببروی هریک از گروه محصولات خاص بررسی نمود.

منابع

- [1] Alex Berson, Stephen Smith, kurt Thearling, "Building Data Mining Application for CRM", McGraw-Hill, 2001.
- [2] Nong, Y., "The Handbook of Data Minig", LAWRENCE ERLBAUM ASSOCIATES PUBLISHERS Mahwah, New Jersey London, 2003.
- [3] Hsiao-Fan Wang, Wei-Kuo Hong, 1. "Managing Customer Profitability in a Competitive Market by Continuous Data Mining", Department of Industrial Engineering and Engineering Management. National Tsing Hua University, Hsinchu, Taiwan, ROC, 2005.
- [4] Chris Rygielski, a., Jyun-Cheng, Wang, b., David, C., Yen, a., "Data Mining Techniques for Customer Relationship Management", Technology in Society 24, 2002, pp. 483-502
- [5] Hyunseok Hwang, Taesoo Jung, Euiho Suh, "An LTV Model and Customer Segmentation Based on Customer Value: a Case Study on the Wireless Telecommunication Industry", Expert Systems with Applications 26, 2004, pp. 181-188.
- [6] Su-Yeon kim, Tae-see Jung, Eui-Ho Suh, Hyun-seok Hwang, "Customer Segmentation and Strategy Development Based on Customer Lifetime Value: A Case Study y", Expert Systems with Applications 31, 2006, pp. 101-107
- [7] Tsai, C.Y., chiu, C.C., "A Purchase – Based Market Segmentation Methodology ", Expert Systems with Applications 27, 2004, pp. 265-276.
- [8] Shina, H.W., Sohn, S.Y., "Segmentation of Stock Trading Customers According to Potential Value:", Expert Systems with Applications 27, 2004, pp. 27-33.
- [9] Jedid-Jah Jonkera, Nanda Piersmab, Dirk Van den Poelc, "Joint Optimization of Customer Segmentation and Marketing Policy to Maximize Long-Term Profitability", Expert Systems with Applications 27, 2004, pp. 159-168.
- [10] Pauline, A., Wilcox, Calin Gurau, "Business Modeling with UML: the Implementation of CRM System for Online Retailing ", Journal of Retailing and Consumer Services 10, 2003, pp. 181-191.
- [11] Wegner, A., kamakura, Michel Wedel, Fernando de Rosa, Jose Afonso Mazzon, "Cross-Selling Through Database Marketing: a to Mixed Data Factor Analyzer for Data Augmentation and Prediction", International journal of Research in Marketing 20, 2003, pp. 45-65.

حجمی کل تقسیم می‌کنیم. در اینصورت ارزش حجمی هر گروه کالا نسبت به کل اقلام به کمک رابطه (۲۲) محاسبه می‌شود. میانگین ارزش گروه کالای a برای هر خوش

میانگین ارزش برای هر خوش

$$W_a = \frac{W_{a1} + W_{a2} + \dots + W_{av}}{W_{a1} + W_{a2} + \dots + W_{av}} = 1 \quad (22)$$

براساس مقادیر W_a ظرفیت حجمی وسائل توزیع بین گروه کالاهای مختلف در هر خوش به صورت مناسب تقسیم می‌شود. محدودیت ظرفیت وسائل نقلیه یک محدودیت مشترک بین همه گروه کالاهای مورد برنامه ریزی می‌باشد. این نسبت‌ها بعذار هر بار اجرای برنامه با توجه به اطلاعات جدید به هنگام می‌شود. بدین منظور ابتدا ظرفیت باقیمانده وسائل نقلیه برای هر گروه کالا در هر خوش محاسبه می‌شود. سپس از جمع این ظرفیت‌های باقیمانده برای هر گروه کالا کل ظرفیت باقیمانده محاسبه می‌شود. ظرفیت باقیمانده برای گروه کالاهایی که زیاد محاسبه شده کم می‌شود و به گروه کالاهای که ظرفیت باقیمانده کمتری دارند تخصیص داده می‌شود. بدین صورت ظرفیت باقیمانده بین گروه‌های مختلف کالا تسطیح واژ ظرفیت وسائل نقلیه حداکثر استفاده به عمل می‌آید.

۵. نتیجه گیری

با افزایش اهمیت رضایت مشتری در محیط تجاری امروز، بسیاری از سازمان‌ها ببروی مباحثت مرتبه با شناخت مشتری، وفاداری و سودآوری مشتری برای افزایش سهم بازار خود و کسب رضایت مشتری تمرکز نموده‌اند. مدیریت ارتباط با مشتری به عنوان یک مزیت رقابتی برای سازمان‌ها محسوب می‌گردد. یکی از روش‌های شناخت مشتریان، بخش بندی مشتریان به گروه‌های همگن و اتخاذ سیاست‌های بازاریابی مناسب با هر بخش است. در این مقاله یک تابع فاصله رفتار مشتریان (DCB) براساس ترکیب قواعد تعريف گردیده است. و از آن به عنوان تابع سنجش فاصله در الگوریتم خوش بندی K-means استفاده شده است که نتایج را نسبت به تابع فاصله اقلیدسی به میزان قابل توجهی بهبود داده است. خوش بندی صورت گرفته با تابع سنجش تراکم با یکدیگر مقایسه شده‌اند و تعداد خوش‌ها با استفاده از تابع سنجش کیفیت مشخص شده است. پس از بخش بندی مشتریان /مناطق با استفاده‌های مدل DTM ارزش هر خوش تعیین و سیاست‌های مناسب با هر بخش تعیین گردیده است. در تحقیقات آتی می‌توان از داده‌ای توصیفی و جمعیت شناختی مشتریان /مناطق نیز در تعريف و بهبود تابع

[12] Arindam Banerjee, joydeep Ghosh, "Clickstream Clustering using Weighted Longest Common Subsequences" Dep of Electrical Engineering university of Texas at Austin, 2002.

[13] Jiawei Han, Micheline Kamber, Department of computer Science, University of Illinois at Urbana – champaign, www.cs.Uiuc.edu/~hanj.

