Islamic Azad University , Shiraz Branch
نشریه تحلیل مدارها، داده ها و سامانه ها
Journal of Circuits, Data and Systems Analysis
sanad.iau.ir/journal/jcdsa

New neighborhood approaches in memetic algorithm for customer type discovery

Hamed Sherafat-Moula[1], S.Hadi Yaghoubyan[2*], Razieh Malekhosseini[3], Karamollah Bagherifard[4]

[1] Department of Computer Engineering, Yasuj Branch, Islamic Azad University, Yasuj, Iran
Young Researchers and Elite Club, Yasuj Branch, Islamic Azad University, Yasuj, Iran
h.sherafat.m@gmail.com
[2] Department of Computer Engineering, Yasuj Branch, Islamic Azad University, Yasuj, Iran
Young Researchers and Elite Club, Yasuj Branch, Islamic Azad University, Yasuj, Iran
yaghoobian.h@gmail.com
[3] Department of Computer Engineering, Yasuj Branch, Islamic Azad University, Yasuj, Iran
Young Researchers and Elite Club, Yasuj Branch, Islamic Azad University, Yasuj, Iran
malekhoseini.r@gmail.com
[4] Department of Computer Engineering, Yasuj Branch, Islamic Azad University, Yasuj, Iran
Young Researchers and Elite Club, Yasuj Branch, Islamic Azad University, Yasuj, Iran
ka.bagherifard@iau.ac.ir

Abstract: "Revenue management" systems are extensively utilized across various industries today. One of the primary pillars of revenue management lies in demand estimation, which predicts the demand for products and services. Understanding customers and their preferences forms the cornerstone of demand estimation, and this understanding is acquired through solving the "customer type discovery" problem. Recently, this problem has been addressed using the "genetic" meta-heuristic method. In this research, we propose solving this problem utilizing the "memetic" meta-heuristic method, employing alternative approaches to identify the neighborhood. By evaluating real data from five hotels, we demonstrate that our method offers the first viable solution to the problem, resulting in a total of 10.5% fewer iterations compared to the "genetic" method.

Keywords: revenue management, meta-heuristic algorithms, genetic algorithm, memetic algorithm

JCDSA, Vol. 2, No. 5, Spring 2024	Online ISSN: 2981-1295	Journal Homepage: https://sanad.iau.ir/en/Journal/jcdsa
Received: 2024-02-02	Accepted: 2024-05-08	Published: 2024-06-15

CITATION

Sherafat, M.H., et al., *"New neighborhood approaches in memetic algorithm for customer type discovery"*, Journal of Circuits, Data and Systems Analysis (JCDSA), Vol. 2, No. 5, pp. 34-44, 2024.
DOI: 00.00000/0000

COPYRIGHTS

©2024 by the authors. Published by the Islamic Azad University Shiraz Branch. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0)
https://creativecommons.org/licenses/by/4.0

* Corresponding author

Extended Abstract

1- Introduction

Earnings management systems are in use in many industries. Maximum profit or revenue can be achieved using revenue management systems and this is a result of the deregulation of the US airline industry in 1978 [1]. Profit management helps in making important business decisions that directly affect the purchasing behavior of customers and consequently the business revenue.

Finding customer demand is known as “demand estimation” and this prediction describes which customers will come and which products they will buy [2-5]. The product offering decision is known as "bundle optimization" which considers what products should be available to customers in a specific time period [6-9]. Pricing strategy is a term used to refer to the procedure that determines dynamic pricing policies and their relationship to revenue management [10-12]. "Demand estimation" is one of the most important stages of the profit management process. At this stage, we should have a clear view of customers and their list of product preferences.

The amount of demand for each product in a period of time can be found by obtaining two main data "products available for sale" and "types of customers"; Therefore, demand estimation may directly depend on finding exact types of customers. As mentioned, finding exact customer types among all factorial types is classified as NP-hard problems [16]. Note that "Sales Transactions" and "Product Availability" over a period of time are the only available data in this problem, and customer types must be inferred from this data. The name of this problem is "discovery of customer type".

2- Methodology

The problem of "customer type discovery" is investigated with many mathematical approaches. Most of these researches use "choice-based demand estimation" models [16-18]. Recently, the idea of using meta-heuristics to find customer types was presented by Haj Mirzaei et al [19]. They presented a genetic algorithm to find suitable customer types from 5 real hotel datasets and solved the problem of customer type discovery.

Finding an acceptable solution using this approach is time-consuming and may not be suitable for some business applications that require immediate results. The use of "local search" algorithm can provide faster solutions, and the connection of "local search" and "genetic" algorithm can help for "faster convergence". The present research is an attempt to find the first acceptable solution to the customer type discovery problem with the least possible generation in meta-heuristic algorithm. The main idea of this will be through the use of "local search" with the genetic algorithm. In this research, different local search approaches will be presented, so that by using them in the local search of the memetic algorithm, we can obtain the first acceptable solution of the problem faster (by reducing the number of generations) compared to the genetic algorithm.

After entering the type of customer in a time period, the customer chooses the product with the highest rank from his list of preferences in that period. If the customer cannot buy a product, he chooses the zero product and leaves the market. The customer entry rate in each period is discrete and homogeneous with a probability of $0 < \lambda < 1$, which is considered the same in all periods. The term "consistency" is used because no one can truly tell which type of customer entered a period. Therefore, we can only find out which type of customer is compatible.

Evolutionary Algorithms (EA) are suitable for cases like customer type discovery problem. As mentioned in the literature review section, genetic algorithm has recently been able to perform better than mathematical approaches for problem solving. Pure evolutionary algorithms are more efficient in searching the problem space when combined with other techniques, such as local searches [25]-[27]. Therefore, the use of neighborhood makes it possible to use local search together with the genetic algorithm [19]. The memetic algorithm is a combination of genetic and memetic evolution We will use them in this research [28].

3- Results and discussion

For the evaluation, we will use the dataset of five Continental hotels located in the United States. This dataset mainly consists of travel information of people who have booked rooms for business purposes. This dataset is publicly available and used as a benchmark for the customer type discovery problem. The data span is from March 12, 2007 to April 15, 2007, which includes 34 days [23]. Reservations are made through various channels such as travel agencies, hotel operators, personal visits and hotel websites. To evaluate the efficiency of our local search approaches, we implemented the mentioned three approaches with the memetic algorithm and compared it with the genetic method [19].

For both approaches (memetic and genetic), we have created an initial random population with a number of 25. The test is to show the speed of convergence in the initial phase (convergence of the population from unacceptable solutions to acceptable solutions). In this test, when the first acceptable solution is created, we have prevented the algorithm from continuing. Each experiment was repeated 30 times and the average results were reported. The results can be seen in Table (2).

4- Conclusion

If we consider each customer type as a preferred list of products, discovering customer types is the first step in estimating product demand and one of the first steps in pricing strategy. If sales transactions and product availability data are used, finding customer types with meta-heuristic methods is challenging due to lack of data. In some applications, finding the first acceptable solution in a short time is more challenging than finding the optimal solution.

Islamic Azad University , Shiraz Branch
نشریه تحلیل مدارها، داده ها و سامانه ها
Journal of Circuits, Data and Systems Analysis
sanad.iau.ir/journal/jcdsa

رویکردهای همسایگی جدید در الگوریتم ممتیک برای یافتن نوع مشتری

حامد شرافت‌مولا[1]، هادی یعقوبیان[2*]، راضیه ملک‌حسینی[3]، کرم‌الله باقری‌فرد [4]

۱- گروه مهندسی کامپیوتر، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
باشگاه پژوهشگران جوان و نخبگان، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
h.sherafat.m@gmail.com
۲- گروه مهندسی کامپیوتر، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
باشگاه پژوهشگران جوان و نخبگان، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
yaghoobian.h@gmail.com
۳- گروه مهندسی کامپیوتر، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
باشگاه پژوهشگران جوان و نخبگان، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
malekhoseini.r@gmail.com
۴- گروه مهندسی کامپیوتر، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
باشگاه پژوهشگران جوان و نخبگان، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
ka.bagherifard@iau.ac.ir

چکیده: سیستم‌های «مدیریت سود» امروزه به‌صورت فراوان در صنایع مختلفی استفاده می‌شوند. یکی از پایه‌های اصلی مدیریت سود، «برآورد تقاضا» است که بر اساس آن تقاضای محصولات و خدمات پیش‌بینی می‌شود. شناخت مشتریان و علایق آنها زیربنای برآورد تقاضاست و این شناخت با حل مسئله «کشف نوع مشتری» به‌دست می‌آید. به تازگی این مسئله با استفاده از روش فراابتکاری «ژنتیک» حل‌شده‌است و در این تحقیق با استفاده از رویکردهایی دیگر برای یافتن همسایگی، این مسئله را با روش فراابتکاری «ممتیک» حل خواهیم‌کرد. برای ارزیابی تحقیق خود، از داده‌های واقعی پنج هتل استفاده خواهیم‌کرد و در ادامه نشان می‌دهیم که روش پیشنهادی درمجموع با ۱۰.۵ درصد تعداد نسل کمتر نسبت به روش «ژنتیک» اولین راه‌حل قابل‌قبول مسئله را ارائه می‌دهد.

واژه‌های کلیدی: مدیریت درآمد، الگوریتم‌های فراابتکاری، الگوریتم ژنتیک، الگوریتم ممتیک

نوع مقاله: پژوهشی
DOI: 00.00000/0000
تاریخ ارسال مقاله: ۱۴۰۲/۱۱/۱۳ تاریخ پذیرش مقاله: ۱۴۰۳/۰۲/۱۹ تاریخ چاپ مقاله: ۱۴۰۳/۰۳/۲۶

۱- مقدمه

سیستم‌های «مدیریت سود[1]» در صنایع زیادی در حال استفاده است. حداکثر سود یا درآمد را می‌توان با استفاده از سیستم‌های مدیریت درآمد به‌دست آورد و این امر حاصل مقررات‌زدایی صنعت هواپیمایی ایالات‌متحده در سال ۱۹۷۸ است [۱]. مدیریت سود به تصمیم‌گیری‌های مهمی در کسب‌وکار کمک می‌کند که مستقیماً بر رفتار خرید مشتریان تأثیر می‌گذارد و درنتیجه، بر درآمد کسب‌وکار تأثیر می‌گذارد. این تصمیمات را می‌توان در سه بخش، دسته‌بندی کرد:

۱. تقاضای مشتریان برای محصولات؛
۲. محصولاتی که ارائه می‌شوند؛
۳. قیمت‌هایی که باید ارائه شوند.

یافتن تقاضای مشتریان به‌عنوان «برآورد تقاضا[2]» شناخته می‌شود و این پیش‌بینی توصیف می‌کند که چه مشتریانی خواهند آمد و کدام محصولات را خریداری خواهند کرد [۲-۵]. تصمیم «ارائه محصولات» به‌عنوان «بهینه‌سازی مجموعه‌ای[3]» شناخته می‌شود که در نظر می‌گیرد چه محصولاتی باید در یک دورهٔ زمانی خاص در دسترس مشتریان باشد [۶-۹]. استراتژی قیمت‌گذاری، اصطلاحی است که برای اشاره به رویه‌ای استفاده می‌شود که سیاست‌های قیمت‌گذاری پویا و

[1] Revenue Managment
[2] Demand Estimation
[3] Assortment Optimization

ارتباط آن‌ها با مدیریت درآمد را تعیین می‌کند [۱۰-۱۲]. «برآورد تقاضا» یکی از مهم‌ترین مراحل فرایند مدیریت سود است. در این مرحله باید دید روشنی از مشتریان و لیست ترجیحات آن‌ها از محصولات داشته باشیم. در بازاری با n نوع محصول، هر «نوع مشتری[1]» با فهرستی مرتب از محصولاتی که ترجیح می‌دهد خریداری کند، نشان داده می‌شود [۸، ۱۳-۱۵]. قابل‌ذکر است که هر مشتری می‌تولند کالایی را که در دسترس نیست (به‌دلیل عدم موجودی یا تغییر استراتژی فروش کسب) با کالایی که در دسترس است، جایگزین کند. چالش اصلی در این زمینه، تعداد بالقوهٔ انواع مشتری است که با توجه به تعداد محصولات و در نظر گرفتن تعداد جایگزین برای محصولات به‌صورت فاکتوریل خواهد بود [۱۶-۱۷]. فرض کنید {1,2,3} محصول در کسب‌وکار وجود دارد و سه نوع مشتری $C_1=\{1,2\}$، $C_2=\{1\}$ و $C_3=\{2,3\}$ در یک بازهٔ زمانی وارد بازار می‌شوند. مشتری نوع C_1 در ابتدا ترجیح می‌دهد که محصول ۱ را بخرد؛ اما اگر ۱ در دسترس نباشد، به‌جای آن محصول ۲ را خریداری می‌کند. نکته مهم این است که اگر هر دو محصول ۱ و ۲ در دسترس نباشد، او هیچ محصولی را خریداری نخواهد کرد و بازار را ترک خواهد کرد. «عدم خرید» با خرید محصول صفر (۰) نمایش داده می‌شود. اگر همه محصولات مدنظر نوع مشتری در دسترس نباشند، او ترجیح می‌دهد محصول صفر را خریداری کند. بنابراین، محصول صفر به انتهای هر نوع مشتری اضافه خواهد شد. روش صحیح نمایش انواع مشتریان به این صورت خواهد بود: $C_1=\{1,2,0\}$ $C_2=\{1,0\}$ $C_3=\{2,3,0\}$.

میزان تقاضا برای هر محصول در یک دوره زمانی را می‌توان با به‌دست‌آوردن دو دادهٔ اصلی «محصولات موجود برای فروش» و «انواع مشتری» پیدا کرد؛ بنابراین، برآورد تقاضا به‌طور مستقیم ممکن است به یافتن انواع دقیق مشتری بستگی داشته باشد. همان‌طور که گفته شد، یافتن انواع مشتریان دقیق در میان همهٔ انواع که به‌صورت فاکتوریل است، در دسته مسائل NP-hard دسته‌بندی می‌شود [۱۶]. قابل توجه است که «تراکنش‌های فروش» و در «دسترس بودن محصول» در یک دوره زمانی، تنها داده‌های قابل دسترس در این مسئله هستند و انواع مشتریان باید از این داده‌ها استنتاج شوند. نام این مسئله «کشف نوع مشتری[2]» است. مسئلهٔ کشف نوع مشتری با رویکردهای ریاضی بسیاری مورد بررسی قرار می‌گیرد. اکثر این پژوهش‌ها از مدل‌های «برآورد تقاضا مبتنی بر انتخاب[3]» استفاده می‌کنند [۱۶-۱۸]. اخیراً ایده استفاده از روش فرااِبتکاری[4] برای یافتن انواع مشتری توسط حاج‌میرزایی و همکاران ارائه‌شده است [۱۹]. آنها یک الگوریتم ژنتیک برای یافتن انواع مشتری مناسب از ۵ مجموعه دادهٔ هتل واقعی ارائه دادند و مسئلهٔ کشف نوع مشتری را حل کردند.

یافتن راه‌حل قابل قبول با استفاده از این رویکرد، زمان‌بر است و ممکن است برای برخی از برنامه‌های کاربردی تجاری که احتیاج فوری به نتیجه دارند، مناسب نباشد. استفاده از الگوریتم «جستجوی محلی[5]» می‌تواند راه‌حل های سریعتری ارائه دهد و اتصال الگوریتم «جستجوی محلی» و «ژنتیک» می‌تواند برای «همگرایی سریعتر» کمک کند. تحقیق حاضر تلاشی برای یافتن اولین راه‌حل قابل قبول مسئله کشف نوع مشتری با کمترین نسل ممکن در الگوریتم فرا ابتکاری است. ایده اصلی این امر، به‌کارگیری «جستجوی محلی» با الگوریتم ژنتیک خواهد بود. در این تحقیق رویکردهای مختلف جستجوی محلی ارائه خواهد شد تا با استفاده از آن‌ها در جستجوی محلی الگوریتم ممتیک، بتوانیم اولین راه‌حل قابل قبول مسئله را سریع‌تر (از طریق کم‌کردن تعداد نسل‌های لازم) نسبت به الگوریتم ژنتیک به‌دست آورده شود.

در ادامه این تحقیق به پیشینه تحقیق پرداخته می‌شود. سپس مسئله «کشف نوع مشتری» مورد بررسی قرارخواهد گرفت و بعد از آن به نحوه حل مسئله با کمک الگوریتم ممتیک و بررسی رویکردهای جدید همسایگی اشاره می‌شود و در بخش بعدی، رویکردهای مدنظر مورد مقایسه قرار می‌گیرند و بخش نتیجه‌گیری پایان این تحقیق خواهد بود.

۲- ادبیات و پیشینه تحقیق

ون رایزین و ولکانو [۱۶] مدل برآورد تقاضای غیرپارامتریک کلی را برای پیداکردن انواع مشتری از تراکنش‌های فروش و داده‌های در دسترس بودن محصول، پیشنهاد کردند که شامل سه مرحلهٔ اصلی است. مرحله اول این مدل از یک لیست محصول تکی برابر با تعداد محصولات برای حل اولیهٔ رویکرد استفاده می‌کند. هر محصول تکی نشان دهنده یک نوع مشتری است که فقط یک محصول و صفر را شامل می‌شود. در مرحله دوم، نویسندگان از مدل تخمین تقاضا استفاده می‌کنند که بر اساس مدل تخمین، حداکثر احتمال برای ارزیابی برازندگی[6] هر راه‌حل محاسبه می‌شود. در مرحله سوم، برنامهٔ عدد صحیح مختلط[7] (MIP) برای یافتن یک نوع مشتری جدید استفاده می‌شود که با اضافه کردن به آن، بهبود برازندگی راه‌حل فعلی را تضمین می‌کند. مراحل ۲ و ۳ تکرار می‌شوند تا زمانی که رویکرد نتواند انواع مشتریان جدیدی را که برازندگی راه‌حل را بهبود می‌بخشد شناسایی کند. همچنین آنها در سال ۲۰۱۷ سعی کردند با استفاده از حداکثرسازی انتظار[8] (EM) و حل مسئله به صورت حداکثر احتمال، کارایی زمانی کار قبلی خود را بهبود بخشند [۱۷] و رویکرد آنها به عنوان حل مسئله کشف نوع مشتری درنظر گرفته شد. حاج‌میرزایی و همکاران [۱۹] از الگوریتم ژنتیک برای پیداکردن انواع مشتری از ۵ مجموعه دادهٔ هتل استفاده کردند. نویسندگان از دو تابع برازندگی برای ارزیابی جمعیت استفاده کردند. تابع برازندگی اصلی، تابع حداکثر

[1] Customer Type
[2] Customer Type Discovery
[3] choice-based demand estimation
[4] Metaheuristics
[5] Local Search
[6] Fitness
[7] Mixed-integer program
[8] Expectation maximization

احتمال است که مشابه MLE است که در [۱۶] برای راه‌حل‌های قابل قبول استفاده می‌شود. برای راه‌حل‌های غیرقابل‌قبول، نویسندگان از یک تابع برازندگی دیگری استفاده کردند. نویسندگان همچنین سعی کردند عملگرهای نمایش و بازتولید بهینه را برای کاربرد خود پیشنهاد کنند و توانستند دقت بهتری نسبت به روش ون رایزن داشته باشند.

در [۲۰]، مدل تخمین تقاضای مبتنی بر انتخاب برای پیداکردن انواع مشتری در سه مرحله استفاده می‌شود: ۱) ساخت نمودار غیر چرخه‌ای جهت‌دار برای هر نوع مشتری، ۲) خوشه‌بندی آن، و ۳) تخمین توزیع‌ها روی فهرست ترجیحی هر نوع مشتری. داده‌های مورد استفاده برای این منظور از نوع تابلویی[1] هستند. در برخی بازارها، مانند بازارهای آنلاین و آن‌هایی که از برخی برنامه‌ها برای ثبت رفتار خرید مشتریان استفاده می‌کنند (برای مثال باشگاه وفاداری)، داده‌های موجود علاوه بر دقت، بسیار بیشتر از معاملات فروش و داده‌های موجود بودن محصول هستند. این نوع داده‌ها به داده‌های تابلویی معروف‌اند و برای مجموعه‌هایی با داده‌های کم، این رویکرد قابل استفاده نیست.

مدل «در نظر بگیرید؛ سپس انتخاب کنید» برای کشف انواع مشتری توسط جاگاباتولا و همکاران استفاده شد [۱۸]. برای هر نوع مشتری، یک مجموعهٔ «درنظرگرفتن توسط مشتری» برای استفاده از روش یادگیری ماشینی مبتنی بر داده استفاده شد و برای کشف هر مجموعهٔ «درنظرگرفتن» از داده‌های خرده‌فروشی اشتراکی خودرو و داده‌های در دسترس بودن محصول استفاده شد و رویکردشان برای مواردی که احتیاج به دسته‌بندی مشتریان بر اساس مکان‌های جغرافیایی است مناسب خواهدبود. لی و یون [۲۱] رویکردی مبتنی بر مدل انتخاب برای یافتن انواع مشتریان ناهمگن با استفاده از داده‌های فروش و در دسترس بودن محصول پیشنهاد کردند. این رویکرد با اندازهٔ ۱ آغاز می‌شود و سپس سعی می‌کند انواع مشتریان جدید را به این راه‌حل اضافه کند تا راه‌حل بهینه را بیابد. اندازه راه‌حل بهینه، تعداد انواع مشتری را با توجه به مجموعهٔ دادهٔ استفاده‌شده تعیین می‌کند و از بیشینه‌سازی انتظارات[2] برای تخمین پارامترهای مدل انتخابی استفاده می‌شود. حاج‌میرزایی و همکاران [۲۲] رویکردی برای کشف نوع مشتری را بر اساس مسئلهٔ سفارش خطی کلاسیک پیشنهاد کردند که از سه بخش تشکیل‌شده است. بخش ۱ یک راه‌حل اولیه ایجاد می‌کند. بخش ۲ احتمال هر نوع مشتری (PMF) را تخمین می‌زند. درنهایت، بخش ۳ مشتری جدیدی پیدا می‌کند. بخش‌های ۲ و ۳ تا رسیدن به معیارهای توقف تکرار می شوند.

در [۳۰] به وظیفه چالش برانگیز افزایش رشد بازاریابی دیجیتال از طریق توسعه یک رویکرد تقسیم‌بندی مشتری کارآمد مبتنی بر هوش مصنوعی می‌پردازد. این مطالعه با استفاده از یک مدل یادگیری عمیق بدون نظارت برای شناسایی ویژگی‌های رفتاری کلیدی در داده‌های مشتری استفاده می‌کند ولی روش آنها نوع مشتری‌ها را پیدا نمی‌کند. در این تحقیق از مجموعه داده کگل استفاده شده و هر فاکتور شامل چندین خرید می‌شود و روش آنها برای محصولاتی که فاسدپذیری دارند قابل استفاده نیست. همچنین در [۳۱]، با در نظر گرفتن اطلاعات نمونه، از تکنیک‌های خوشه‌بندی مانند K-means برای شناسایی بخش‌های مشتری متمایز استفاده کردند. این تکنیک‌ها اجازه می‌دهد که با گروه‌بندی مشتریان بر اساس ویژگی‌های مشترک، از جمله جنسیت، سن، علایق و عادت‌های خرج کردن، بر پایه‌های کاربران بالقوه تمرکز کرد. آنها از پروفایل کاربر برای کار خود به استفاده کردند و کار آنها کارایی لازم را برای مجموعه داده‌هایی که داده‌های کمی دارند قابلیت استفاده ندارد. در [۳۲] بر نقش حیاتی تقسیم‌بندی مشتری در هدایت شخصی‌سازی، تبلیغات هدفمند و استراتژی‌های تبلیغاتی تأکید می‌کنند و روندهای ارزشمندی را هم به محققان ارائه می‌کنند. ادبیات موضوعی موجود در تحقیق آنها عمدتاً بر انواع تقسیم‌بندی منحصربه‌فرد، به‌ویژه تقسیم‌بندی رفتاری مبتنی بر پروفایل کاربر تکیه دارد که دید جامع نسبت به مشتری که متکی بر داده‌های خرید باشد را محدود می‌کند.

۳- شرح مسئله

این بخش، مسئلهٔ «کشف نوع مشتری» را توضیح می‌دهد. جهت راحتی خواننده از نماد و اصطلاحات [۱۹] استفاده می‌گردد. فرض کنید $N = \{0,1, \dots, n\}$ محصولات خریداری‌شده در بازه[3]های زمانی $t = \{1,2, \dots, T\}$ توسط انواع مشتریان ناشناخته که در تلاش برای یافتن آن‌ها هستیم، هستند. انتخاب محصول صفر توسط مشتری به معنای خریدن نکردن و خروج از بازار است. همانطور که قبلاً اشاره شد، هر نوع مشتری با لیست ترجیحی از محصولاتی که در آن محصول h بالاتر از j قرار دارد، نشان داده می‌شود $C(j) < C(h)$؛ یعنی مشتری خرید محصول h را به j ترجیح می‌دهد. لیست محصولات ارائه‌شده در هر دورهٔ زمانی با $S_t \in 0\ N, \subseteq S_t$ ، که در آن $S_t \geq 2$ است (محصول صفر و حداقل یک محصول دیگر) مشخص می شود. پس از ورود نوع مشتری در یک بازهٔ زمانی، مشتری محصولی را با بالاترین رتبه از لیست ترجیحات خود در آن دوره انتخاب می‌کند. اگر مشتری نتواند محصولی را بخرد، محصول صفر را انتخاب کرده و بازار را ترک می‌کند. نرخ ورود مشتریان در هر دوره گسسته و همگن با احتمال $0 < \lambda < 1$ است که در تمامی دوره‌ها یکسان در نظر گرفته می‌شود.

جدول (۱) نمونه‌ای از تراکنش‌های فروش و داده‌های در دسترس بودن محصول را پس از پیش‌پردازش مجموعهٔ داده نشان می‌دهد. در این جدول ۵ دوره زمانی و ۵ محصول وجود دارد. در دوره ۱، سه محصول برای فروش در دسترس (موجود) است و محصول ۱ توسط مشتری خریداری‌شده است. در دوره ۲ محصولات ۱ و ۲ و ۵ موجود است و محصول ۵ خریداری می‌شود. فرض کنید سه نوع مشتری وارد بازار می‌شوند: $C_1 = \{1,2,3,4,5\}$، $C_2 = \{4,5\}$ و $C_3 = \{2\}$. در دوره ۱، مشتری C_1 در این دوره سازگار[4] است، یعنی اگر C_1 وارد این

[1] Panel
[2] Expectation maximization
[3] Period
[4] Compatible

دوره شود، محصول ۱ خریداری می شود. اصطلاح «سازگاری» به این دلیل به کار می‌رود که هیچ‌کس نمی‌تواند به‌طور واقعی تشخیص دهد کدام نوع مشتری در یک دوره واردشده است. بنابراین، فقط می‌توانیم بفهمیم کدام نوع مشتری سازگار است. در دوره ۲، C_2 تنها نوع مشتری سازگار است. به این مورد دقت شود که در دوره ۲، C_1 سازگار نیست؛ زیرا اگر C_1 در این دوره می‌رسید، مشتری محصول ۱ را خریداری می‌کرد. در دوره ۴، هیچ نوع مشتری وارد نشده است. در دوره ۵، هیچ محصولی خریداری نشده چرا که فقط نوع ۲ و ۳ مشتری سازگار باشد زیرا اگر C_1 به این دوره برسد، محصول ۳ را خریداری می‌کند. یافتن سازگاری انواع مشتری برای هر دوره، اولین گام فرایند در رویکردهای کشف بازار است. انواع مشتریان سازگار در دورهٔ t با $M_t(j_t, S_t)$ نشان داده می‌شوند. برای مجموعه کامل انواع مشتری $\{\mathbb{C}_1, \dots, \mathbb{C}_k\}$ و $j_t \in S_t$ که در آن S_t محصولات ارائه شده و j_t تراکنش دوره t است، انواع مشتری سازگار به صورت زیر پیدا می شود:

$$M_t(j_t, S_t) = \{i: \mathbb{C}_i(j_t) < \mathbb{C}_i(k) \;\; \forall k \in S_t, k \neq j_t\} \quad (۱)$$

۳-۱- مدل مبتنی بر انتخاب[1]

مدل انتخاب گسسته کلی در [۱۵] پیشنهادشده است که از آن به‌عنوان یک مدل غیرپارامتریک برای یافتن احتمال خرید محصول j_t در بازه زمانی t استفاده می‌گردد:

$$\begin{cases} \lambda p(j_t|S_t), j_t \in S_t & purchase\ probability \\ p(0|S_t) + (1-y) & No-purchase\ probability \end{cases} \quad (۲)$$

که در آن منظور از عدم خرید، خرید محصول صفر یا عدم ورود مشتری است. برآوردگر حداکثر درستنمایی[2] (MLE) برای برآورد تقاضا به شرح زیر استفاده می‌شود:

$$\mathcal{L}(X|\lambda) = \prod_{t=1}^{T} \lambda p(j_t|S_t) \quad (۳)$$

جدول (۱): در دسترس بودن محصول و نمونه تراکنش فروش

محصولات	دوره ۱	۲	۳	۴	۵
۱	بله	بله	خیر	خیر	خیر
۲	بله	بله	خیر	خیر	خیر
۳	بله	خیر	خیر	خیر	بله
۴	خیر	خیر	بله	خیر	خیر
۵	خیر	بله	بله	بله	خیر
تراکنش	۱	۵	۴	NA	۰
سازگاری	{1}	{2}	{1,2}	-	{2,3}
نوع تراکنش	T_P	T_P	T_P	T_{NA}	T_{NP}

" P "، "NP"، "NA"" به ترتیب مخفف Purchase، No-Purchase و No-Arrival است.

همان‌طور که قبلا ذکر شد، بازه T کل شامل سه نوع مختلف است:

$$\mathcal{L}(X|\lambda) = \prod_{t\epsilon T_P} \lambda p(j_t|S_t) \cdot \prod_{t\epsilon T_{NP}} \lambda p(0|S_t) \cdot \prod_{t\epsilon T_{NA}} (1-\lambda) \quad (۴)$$

و حداکثر تخمین log-likelihood به شرح زیر است:

$$\mathcal{L}_l(X|\lambda) = \sum_{t\epsilon T_P} \big(\log\lambda + \log P\,(j_t|S_t)\big) + \sum_{t\epsilon T_{NP}} \big(\log\lambda + \log P\,(0|S_t)\big) + \sum_{t\epsilon T_{NA}} \log(1-\lambda) \quad (۵)$$

هنگامی S_t ارائه می‌شود، احتمال انتخاب یک محصول j_t توسط مشتریان در دوره t به صورت $\sum_{i\epsilon M_t(j_t,S_t)} x_i$ است. بدیهی است که در صورت عدم ارائه j_t احتمال صفر خواهد بود. با جایگزین کردن تابع احتمال در معادله (۵)، می‌توانیم داشته باشیم:

$$\mathcal{L}_l(X|\lambda) = \sum_{t\epsilon T_P} \left(\log\lambda + \log\left(\sum_{i\epsilon M_t(j_t|S_t)} x_i\right)\right) + \sum_{t\epsilon T_{NP}} \left(\log\lambda + \log\left(\sum_{i\epsilon M_t(0|S_t)} x_i\right)\right) + \sum_{t\epsilon T_{NA}} \log(1-\lambda) \quad (۶)$$

و پس از خلاصه‌سازی (۶):

$$\mathcal{L}_l(X|\lambda) = \sum_{t\epsilon T_P} \log\left(\sum_{i\epsilon M_t(j_t|S_t)} x_i\right) + \sum_{t\epsilon T_{NA}} \log\left(\sum_{i\epsilon M_t(0|S_t)} x_i\right) + |T_{NA}|\log(1-\lambda) + (|T_P| + |T_{NP}|)\log\lambda \quad (۷)$$

در (۷)، عبارت اول نشان‌دهنده log-likelihood خرید کالا در دوره‌های خرید، عبارت دوم برای دوره‌های بدون خرید و عبارت آخر برای دوره‌های بدون ورود است. این تابع log-likelihood مشابه [۱۶] است. معادله در x و λ قابل‌تفکیک است و به‌صورت سراسری در (x,λ) مقعر است. بنابراین، به‌راحتی می‌توانیم $\lambda^* = \frac{|T_P|+|T_{NP}|}{T}$ را پیدا کنیم.

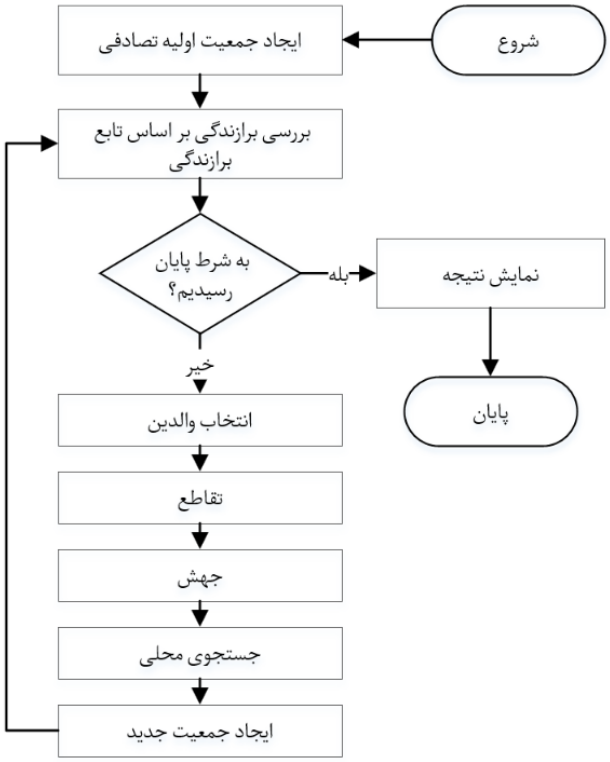
۴- ممتیک و حل مسئله «کشف نوع مشتری»

الگوریتم‌های تکاملی[3] (EA) برای موارد مشابه مسئله کشف نوع مشتری مناسب هستند. همان‌طور که در بخش بررسی ادبیات ذکر شد، اخیراً الگوریتم ژنتیک توانسته از رویکردهای ریاضی برای حل مسئله بهتر عمل کند. الگوریتم‌های تکاملی خالص در جستجوی فضای مسئله

[1] Choice-based model
[2] Maximum likelihood estimator
[3] Evolutionary algorithms

زمانی که با تکنیک‌های دیگر، مانند جستجوهای محلی ترکیب شوند، کارآمدتر هستند [۲۵-۲۷]. بنابراین، استفاده از همسایگی این امکان را می‌دهد از جستجوی محلی همراه با الگوریتم ژنتیک [۱۹] استفاده شود. الگوریتم ممتیک، ترکیبی از تکاملی‌های ژنتیک و ممتیک است و در این تحقیق از آنها استفاده خواهیم کرد [۲۸]. شکل (۱) روند کلی ممتیک را نمایش می‌دهد.

در این بخش ساختار ممتیک مورد استفاده برای حل مسئله «کشف نوع مشتری» را شرح خواهیم داد. قابل توجه است که هدف اصلی این مقاله، نشان‌دادن تأثیر رویکردهایی است که برای جستجوی محلی در حل این مسئله استفاده‌شده است. برای درک این فرآیند، باید بین راه‌حل قابل‌قبول و غیرقابل‌قبول تمایز قائل شویم. فرض کنید راه‌حلی وجود دارد که شامل برخی از انواع مشتری است. برای ارزیابی این راه‌حل، ابتدا باید سازگاری هر دوره را با استفاده از $M_t(j_t, S_t)$ پیدا کنیم. پس از آن، اگر راه‌حلی همه دوره‌ها را پوشش دهد و در همه دوره‌ها سازگار باشد، راه‌حلی قابل‌قبول است. ازنقطه‌نظر امکان‌سنجی، می‌توان فرآیند الگوریتم ممتیک (یا سایر تکاملی‌ها) را برای حل مسئله کشف نوع مشتری به دو فاز مختلف طبقه‌بندی کرد. فاز اول، از راه‌حل‌های غیرقابل‌قبول به راه‌حل‌های قابل‌قبول همگرا می‌شود و فاز دوم یافتن راه‌حل بهینه با استفاده از راه‌حل‌های قابل قبول است. در تحقیق فعلی، اولین هدف تسریع فرآیند فاز اول باکمک جستجوی محلی است. هدف دوم در حیطه‌ی این تحقیق نمی‌گنجد. شبه کد الگوریتم ممتیک استفاده‌شده به شرح زیر است:



شکل (۱): الگوریتم ممتیک [۲۸]

1	Generate random initial ***population***
2	Local search
3	Fitness evaluation
4	**Repeat until** (termination condition)
5	Single-point crossover
	Reproduction operators
6	Local search
7	Fitness evaluation
8	Survival selection
9	**End repeat**

۴-۱- نحوه نمایش

یک راه‌حل، شامل لیستی از انواع مشتری با اندازه حداقل یک است. هر نوع مشتری یک لیست ترجیحی از محصولات است که به محصول صفر ختم می‌شود. به‌عنوان‌مثال، راه‌حل زیر شامل سه نوع مشتری با اندازه‌های مختلف است.

$$solution = [\{2,7,3,4,6,0\}, \{1,2,0\}, \{5,4,2,1,0\}]$$

۴-۲- جمعیت اولیه

در ابتدای کار الگوریتم، جمعیت اولیه تصادفی با اندازه ثابت تولید شده است. هر راه‌حل از جمعیت اولیه شامل تعدادی تصادفی از انواع مشتری و هر نوع مشتری شامل تعداد تصادفی از محصولات است. محصولات تکراری در نوع مشتری و محصول صفر در ابتدای نوع مشتری مجاز نیستند.

۴-۳- تابع برازندگی

قبل از استفاده از تابع برازندگی برای ارزیابی یک راه‌حل، باید سازگاری آن راه‌حل را در تمام دوره‌ها پیدا کنیم. اگر حداقل یک دوره ناسازگار وجود داشته باشد، به معنای غیرقابل‌قبول بودن راه‌حل است و اگر دوره‌های ناسازگار وجود نداشته باشد، به معنای قابل‌قبول بودن راه‌حل است. تابع برازندگی زیر از دو بخش مجزا تشکیل‌شده است. قسمت اول $|T_{inc}| > 0$ برای راه‌حل های غیرقابل قبول و قسمت دوم برای راه‌حل‌های قابل‌قبول استفاده می‌شود. بدیهی است که تابع برازندگی را پس از ارزیابی سازگاری دوره ها برای یک راه‌حل انتخاب می کنیم.

$$Fitness = \begin{cases} \mu \times |T_{inc}|, & |T_{inc}| > 0 \\ \max_{x \geq 0} \mathcal{L}_l(x), s.t. \sum_{i=1}^{k} x_i = 1, & |T_{inc}| = 0 \end{cases} \quad (۸)$$

کل دوره‌ها $T = T_c + T_{inc}$ می‌باشد که در آن $|T_c|$ و $|T_{inc}|$ به ترتیب تعداد دوره های سازگار و ناسازگار را نشان می دهد. قابل توجه است که T شامل دوره‌های بدون ورود نمی‌شود. پارامتر $\mu > 0$ ثابت تنبیه است. در بخش دوم تابع برازندگی، $\mathcal{L}_l(x)$ تابع (۳) است که در بخش قبل توضیح داده شد.

۴-۴- عملگر ترکیب تک نقطه ای

برای اعمال عملگر ترکیب تک نقطه‌ای[1] روی دو راه‌حل، نیمی از طول راه‌حل را به‌عنوان نقطه انتخاب کردیم. با استفاده از این کراس‌اُوور، دو فرزند به وجود می‌آید.

۴-۵- عملگرهای تولید مثل

انتخاب چرخ رولت برای انتخاب یکی از اپراتورهای تولید مثل زیر و اعمال آن برای فرزندان اعمال می شود:

۱. اضافه کردن یک نوع مشتری تصادفی با طول تصادفی به یک راه‌حل.
۲. حذف یک نوع مشتری تصادفی از یک راه‌حل.
۳. حذف یک محصول تصادفی از نوع مشتری تصادفی در یک راه‌حل.
۴. جایگزینی یک محصول تصادفی با محصول دیگری از نوع مشتری تصادفی در یک راه‌حل.
۵. اعمال یک کراس‌اوور تک نقطه‌ای برروی دو نوع تصادفی مشتری از یک راه‌حل.

۴-۶- انتخاب بازماندگان

جمعیت زنده مانده شامل ۸۰ درصد بهترین و ۲۰ درصد از بدترین راه‌حل‌ها از جمعیت وللدین و فرزندان است. در این تحقیق، جمعیت باقی‌مانده را قبل از استفاده در تکرار بعدی به هم می‌ریزیم.

۴-۷- جستجوی محلی

ادعا می‌شود که موفقیت ممتیک به‌دلیل مبادله بین قابلیت‌های اکتشاف و توانایی بهره‌برداری از طریق جستجوی محلی است که مورداستفاده قرار می‌گیرد [۲۹]. طراحی یک روش جستجوی محلی کارآمد مهم است. ضمناً در این تحقیق هدف اصلی پیداکردن اولین پاسخ قابل‌قبول با کمتر کردن تعداد نسل‌های موردنیاز فرا ابتکاری است نه حل و یافتن راه‌حل بهینه مسئله کشف نوع مشتری. به‌طورخلاصه در این بخش، یک‌سری رویکرد را معرفی می‌کنیم و سپس این رویکردها را ارزیابی می‌کنیم تا یکی از آنها را به‌عنوان رویکرد جستجوی محلی الگوریتم ممتیک انتخاب کنیم.

همان‌طور که در بخش قبل توضیح داده شد، یک راه‌حل شامل برخی از انواع مشتری است. فرض کنید که راه‌حل $solution\ (j) = [\{ctype^{(1)}\}, \dots, \{ctype^{(l)}\}]$ که در آن $ctype^{(i)}, 1 < i < l$ یک نوع مشتری است. دیدگاه اصلی جهت یافتن همسایه بهتر برای $ctype^{(i)} \in solution(j)$ است. در ذیل، سه رویکرد مختلف برای جستجوی محلی با توجه به این دیدگاه پیشنهاد شده است.

- رویکرد اول: همان‌طور که قبلا ذکر شد، $ctype^{(i)}$ لیستی از برخی محصولات است. در این رویکرد، جایگشت‌های دیگر این محصولات را به عنوان $neighbor^{(i)}$ در نظر می‌گیریم.
- رویکرد دوم: تعداد دوره‌های ناسازگار با $ctype^{(i)}$ را درنظر می‌گیریم. محصولی که باعث ایجاد بیشترین ناسازگاری شده است را پیدا می‌کنیم. سایر جایگشت های این نوع مشتری را بدون محصولی که باعث ناسازگاری شده را به عنوان $neighbor^{(i)}$ در نظر می‌گیریم.
- رویکرد سوم: سایر جایگشت های $ctype^{(i)}$ را ایجاد می‌کنیم، سپس مقدار تابع برازندگی را به ازای همه محاسبه می‌کنیم و بهترین نوع مشتری را به عنوان $neighbor^{(i)}$ معرفی می‌کنیم.

۵- آزمایشات

برای ارزیابی، از مجموعهٔ داده مربوط به پنج هتل Continental واقع در ایالات‌متحده استفاده شد. این مجموعه داده عمدتاً شامل اطلاعات سفر افرادی است که برای اهداف تجاری اتاق رزرو کرده‌اند. این مجموعه‌داده به‌صورت عمومی در دسترس است و به‌عنوان بنچمارک برای مسئلهٔ کشف نوع مشتری استفاده می‌شود. بازه زمانی داده‌ها از ۱۲ مارس ۲۰۰۷ تا ۱۵ آوریل ۲۰۰۷ است که ۳۴ روز را شامل می‌شود [۲۳]. رزروها از طریق کانال‌های مختلفی مانند آژانس‌های مسافرتی، اپراتورهای هتل، بازدیدهای شخصی و وب‌سایت هتل‌ها انجام‌شده است. برای ارزیابی کارایی رویکردهای جستجوی محلی خود، سه رویکرد ذکر شده با الگوریتم ممتیک، پیاده‌سازی و روش ژنتیک [۱۹] مقایسه شد. برای هر دو رویکرد (ممتیک و ژنتیک) جمعیت اولیه تصادفی با تعداد ۲۵ ایجاد شد. آزمایش برای نمایش سرعت همگرایی در فاز اولیه (همگرایی جمعیت از راه‌حل‌های غیرقابل‌قبول به راه‌حل‌های قابل‌قبول) است. در این آزمایش هنگامی که اولین راه‌حل قابل‌قبول ایجادشده است، از ادامه کار الگوریتم جلوگیری شد. هر آزمایش ۳۰ بار تکرار شده و میانگین نتایج در جدول (۲) گزارش‌شده است. در ستون «دیتاست» شماره مجموعه داده هتل و در ستون رویکرد، ژنتیک و سه رویکرد پیشنهادی برای همسایگی در ممتیک مشاهده می‌شود. منظور از اندازه راه‌حل، تعداد نوع‌های مشتری موجود در راه‌حل است. همچنین در شکل‌های (۲-۳) درصد بهبود نسبت به ژنتیک مشاهده می‌شود. همان‌طور که از این شکل‌ها مشخص است، تنها رویکرد سوم می‌تواند میزان همگرایی بهتری در زمینه همسایگی اعمال کند. همچنین برای مقایسه بهتر، میزان شباهت که در تابع برازندگی استفاده و توسط (۵) محاسبه می‌شود نیز با دو روش دیگر نیز مقایسه شد که در جدول (۳) قابل مشاهده است. این مقایسه جهت بررسی این موضوع است که آیا ارزش راه‌حل ها در صورت استفاده از همسایگی کمتر می‌شود یا نه. در این جدول با توجه به اینکه میزان شباهت هر دو الگوریتم مشابه بوده‌است، هر دو روش را در یک سطر قرار داده‌ایم. به این مورد دقت شود که در این تحقیق، میزان برتری الگوریتم ممتیک نسبت به ژنتیک

[1] Crossover

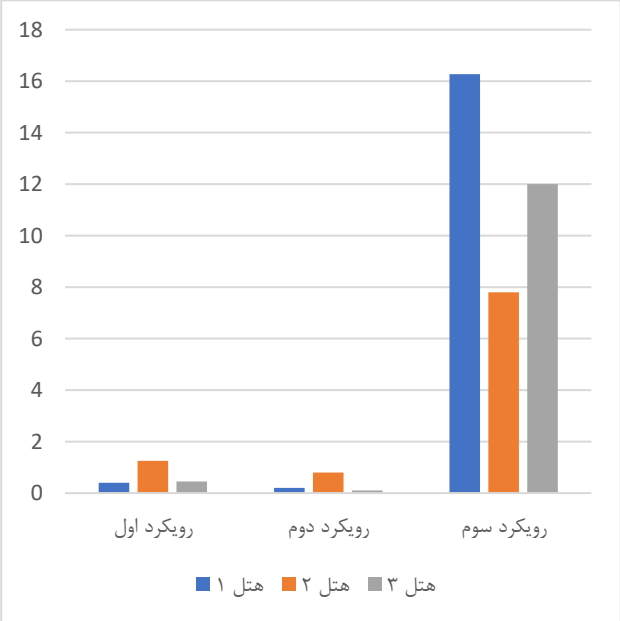
[۱۹] در کم شدن تعداد نسل مورد نیاز برای رسیدن به راه‌حل بوده است. در جدول (۳) روش ID[1] در [۱۶] و روش RMD[2] در [۱۷] نمایش داده شده است. رویکرد همسایگی باعث کم ارزش‌تر شدن راه‌حل های نهایی نیست.

جدول (۲): نمایش تعداد نسل موردنیاز برای ایجاد اولین راه‌حل قابل‌قبول اندازه راه‌حل

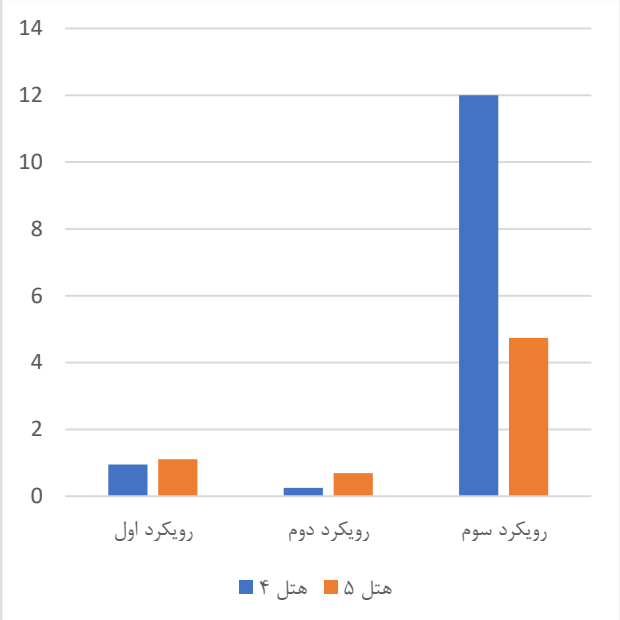
دیتاست	رویکرد	اندازه راه‌حل	تعداد نسل
هتل ۱	ژنتیک	۱۲.۲	۹۸.۳
	ممتیک، رویکرد اول	۱۲.۸	۹۷.۹
	ممتیک رویکرد دوم	۱۲.۶	۹۸.۱
	ممتیک رویکرد سوم	۱۲.۵	**۸۲.۳**
هتل ۲	ژنتیک	۱۲.۷	۱۹۸.۸
	ممتیک، رویکرد اول	۱۱.۸	۱۹۶.۳
	ممتیک رویکرد دوم	۱۱.۷	۱۹۷.۲
	ممتیک رویکرد سوم	۱۱.۶	**۱۸۳.۲**
هتل ۳	ژنتیک	۱۲.۶	۱۹۷.۰
	ممتیک، رویکرد اول	۱۳.۳	۱۹۶.۱
	ممتیک رویکرد دوم	۱۳.۵	۱۹۶.۸
	ممتیک رویکرد سوم	۱۳.۶	**۱۷۳.۱**
هتل ۴	ژنتیک	۹.۶	۱۱۵.۴
	ممتیک، رویکرد اول	۱۰.۰	۱۱۴.۳
	ممتیک رویکرد دوم	۱۰.۳	۱۱۵.۱
	ممتیک رویکرد سوم	۱۰.۲	**۱۰۰.۹**
هتل ۵	ژنتیک	۱۱.۱	۷۱.۷
	ممتیک، رویکرد اول	۱۰.۶	۷۰.۹
	ممتیک رویکرد دوم	۱۰.۷	۷۱.۲
	ممتیک رویکرد سوم	۱۰.۵	**۶۸.۳**

جدول (۳): نمایش تعداد نسل موردنیاز برای ایجاد اولین راه‌حل قابل‌قبول اندازه راه‌حل

دیتاست	روش	Log-Value
هتل ۱	I.D	-۲۶۲۱
	RMD	-۲۴۲۶
	ژنتیک و ممتیک	**-۲۰۹۵**
هتل ۲	I.D	-۷۷۸
	RMD	-۵۳۶
	ژنتیک و ممتیک	**-۴۹۱**
هتل ۳	I.D	-۱۵۶۹
	RMD	-۱۱۲۹
	ژنتیک و ممتیک	**-۱۰۷۳**
هتل ۴	I.D	-۶۱۶
	RMD	-۴۰۱
	ژنتیک و ممتیک	**-۳۵۲**
هتل ۵	I.D	-۴۰۶
	RMD	-۳۷۰
	ژنتیک و ممتیک	**-۳۳۹**



شکل (۲): درصد بهبود رویکردهای همسایگی ممتیک نسبت به ژنتیک در هتل‌های ۱ الی ۳



شکل (۳): درصد بهبود رویکردهای همسایگی ممتیک نسبت به ژنتیک در هتل‌های ۴ و ۵

۶- نتیجه‌گیری

اگر هر نوع مشتری را به‌عنوان یک لیست ترجیحی از محصولات در نظر بگیریم، کشف انواع مشتری گام ابتدایی برای برآورد تقاضای محصولات و یکی از اولین مراحل استراتژی قیمت‌گذاری است. در صورت استفاده از معاملات فروش و داده‌های موجود بودن محصول، یافتن انواع مشتری با روش‌های فراابتکاری به‌دلیل کمبود داده چالش‌برانگیز است. در برخی از برنامه‌های کاربردی، یافتن اولین راه‌حل قابل‌قبول در یک‌زمان کوتاه چالش‌برانگیزتر از یافتن راه‌حل بهینه است.

[1] Independent Demand

[2] Ranked-based Market Discovery

در این تحقیق برای پیداکردن سریعتر اولین راه‌حل قابل‌قبول در الگوریتم ژنتیک به سراغ الگوریتم ممتیک رفته‌ایم. زیرا با ترکیب این الگوریتم با الگوریتم جستجوی محلی میتوان انتظار داشت که این الگوریتم سریعتر به اولین راه‌حل قابل‌قبول برسد، لذا در این تحقیق سه رویکرد برای جستجوی محلی ارائه کردیم، سپس این سه رویکرد را پیاده‌سازی و با الگوریتم ژنتیک مقایسه کردیم، این مقایسه را با استفاده از داده‌های واقعی پنج هتل انجام داده‌ایم.

نتایج مقایسه نشان داد که اگر روند همسایگی به نحوی باشد که تابع برازندگی را به‌سمت نتایج بهتر سوق دهد می‌توان همگرایی سریع‌تری در الگوریتم داشت و سریع‌تر به اولین راه‌حل قابل‌قبول مسئله برسیم. همچنین روش پیشنهادی راه‌حل‌هایی با ارزش بالاتر نسبت به روش‌های گذشته و هم‌ارزش با روش موجود ارائه می‌دهد و نشان می‌دهد که همسایگی تاثیر منفی بر روی ارزش راه‌حل‌های ارائه شده ندارد. می‌توان رویکردهای دیگری را به عنوان همسایگی در نظر گرفت؛ این رویکردها خود می‌توانند الگوریتم‌های فراابتکاری (نظیر PSO) باشد تا بتوان همسایگی بهتری را برای الگوریتم ممتیک به‌دست آورد و ارزش راه‌حل‌های نهایی را بالاتر برد. این موضوع را می‌توان به عنوان پیشنهاد برای ادامه این تحقیق درنظر گرفت.

مراجع

[1] K. T. Talluri and G. J. Van Ryzin, *The Theory and Practice of Revenue Management*, vol. 68. Boston, MA: Springer US, 2004.

[2] P. Liu and S. Smith, "Estimating unconstrained hotel demand based on censored booking data," *Journal of Revenue and ...*, July 01, 2002. http://link.springer.com/10.1057/palgrave.rpm.5170015 (accessed February 21, 2018).

[3] A. Nikseresht and K. Ziarati, "Estimating True Demand in Airline's Revenue Management Systems using Observed Sales," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 7, pp. 361–369, 2017, doi: 10.14569/ijacsa.2017.080748.

[4] C. Y. Goh, C. Yan, and P. Jaillet, "Estimating Primary Demand in Bike-sharing Systems," *SSRN Electron. J.*, Jan. 2019, doi: 10.2139/ssrn.3311371.

[5] J. P. Newman, M. E. Ferguson, L. A. Garrow, and T. L. Jacobs, "Estimation of Choice-Based Models Using Sales Data from a Single Firm," *Manuf. Serv. Oper. Manag.*, vol. 16, no. 2, pp. 184–197, May 2014, doi: 10.1287/msom.2014.0475.

[6] A. Aouad, V. F. Farias, and R. Levi, "Assortment Optimization under Consider-then-Rank Choice Models," *SSRN Electron. J.*, Jun. 2015, doi: 10.2139/ssrn.2618823.

[7] A. Aouad, V. Farias, R. Levi, and D. Segev, "The approximability of assortment optimization under ranking preferences," *Oper. Res.*, vol. 66, no. 6, pp. 1661–1669, Nov. 2018, doi: 10.1287/opre.2018.1754.

[8] D. Bertsimas and V. V Mišic, "Data-driven assortment optimization," *Manage. Sci.*, vol. 1, pp. 1–35, 2015.

[9] S. Jagabathula, "Assortment Optimization Under General Choice," *Ssrn*, pp. 1–51, 2014, doi: 10.2139/ssrn.2512831.

[10] S. Jagabathula and P. Rusmevichientong, "A Nonparametric Joint Assortment and Price Choice Model," *Ssrn*, no. July, 2013, doi: 10.2139/ssrn.2286923.

[11] G. Gallego, H. Topaloglu, and others, *Revenue management and pricing analytics*, vol. 209. Springer, 2019.

[12] G. Bitran and R. Caldentey, "An overview of pricing models for revenue management," *Manuf. Serv. Oper. Manag.*, vol. 5, no. 3, pp. 203–229, 2003.

[13] S. Kunnumkal, "Randomization Approaches for Network Revenue Management with Customer Choice Behavior," *Prod. Oper. Manag.*, vol. 23, no. 9, pp. 1617–1633, Sep. 2014, doi: 10.1111/poms.12164.

[14] L. Chen and T. Homem-de-Mello, "Mathematical programming models for revenue management under customer choice," *Eur. J. Oper. Res.*, vol. 203, no. 2, pp. 294–305, Jun. 2010, doi: 10.1016/J.EJOR.2009.07.029.

[15] G. Vulcano, G. van Ryzin, and R. Ratliff, "Estimating Primary Demand for Substitutable Products from Sales Transaction Data," *Ssrn*, no. August 2015, 2011, doi: 10.2139/ssrn.1923711.

[16] G. van Ryzin and G. Vulcano, "A Market Discovery Algorithm to Estimate a General Class of Nonparametric Choice Models," *Manage. Sci.*, vol. 61, no. 2, pp. 281–300, 2015, doi: 10.1287/mnsc.2014.2040.

[17] G. van Ryzin and G. Vulcano, "Technical Note—An Expectation-Maximization Method to Estimate a Rank-Based Choice Model of Demand," *Oper. Res.*, vol. 65, no. 2, pp. 396–407, 2017, doi: 10.1287/opre.2016.1559.

[18] S. Jagabathula, D. Mitrofanov, and G. Vulcano, "Inferring Consideration Sets from Sales Transaction Data," *SSRN Electron. J.*, 2019, doi: 10.2139/ssrn.3410019.

[19] M. HajMirzaei, K. Ziarati, and A. Nikseresht, "Discovering customer types using sales transactions and product availability data of 5 hotel datasets with genetic algorithm," *J. Revenue Pricing Manag.*, 2020, doi: 10.1057/s41272-020-00245-3.

[20] S. Jagabathula and G. Vulcano, "A Partial-order-based Model to Estimate Individual Preferences Using Panel Data," *SSRN Electron. J.*, no. April, 2017, doi: 10.2139/ssrn.2560994.

[21] H. Lee and Y. Eun, "Discovering heterogeneous consumer groups from sales transaction data," *Eur. J. Oper. Res.*, vol. 280, no. 1, pp. 338–350, Jan. 2020, doi: 10.1016/J.EJOR.2019.05.043.

[22] M. HajMirzaei, K. Ziarati, and A. Nikseresht, "A customer type discovery algorithm in hotel revenue management systems," *J. Revenue Pricing Manag.*, 2021, doi: 10.1057/s41272-020-00273-z.

[23] T. Bodea, M. Ferguson, and L. Garrow, "Data Set — Choice-Based Revenue Management: Data from a Major Hotel Chain ," *Manuf. Serv. Oper. Manag.*, vol. 11, no. 2, pp. 356–361, 2008, doi: 10.1287/msom.1080.0231.

[24] L. Davis, Handbook of genetic algorithms. 1996.

[25] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, 1997.

[26] J. C. Culberson, "On the futility of blind search: An algorithmic view of 'no free lunch,'" *Evol. Comput.*, vol. 6, no. 2, pp. 109–127, 1998.

[27] D. E. Goldberg and S. Voessner, "Optimizing global-local search hybrids.," in *GECCO*, 1999, vol. 99, pp. 220–228.

[28] P. Moscato, "On evolution, search, optimization, GAs and martial arts: toward memetic algorithms. California Inst. Technol., Pasadena," 1989.

[29] N. Krasnogor and J. Smith, "A tutorial for competent memetic algorithms: model, taxonomy, and design issues," *IEEE Trans. Evol. Comput.*, vol. 9, no. 5, pp. 474–488, 2005.

[30] Wang, C. (2022). Efficient customer segmentation in digital marketing using deep learning with swarm

intelligence approach. *Information Processing & Management*, *59*(6), 103085.

[31] Narayana, V. L., Sirisha, S., Divya, G., Pooja, N. L. S., & Nouf, S. A. (2022, March). Mall customer segmentation using machine learning. In *2022 International Conference on Electronics and Renewable Systems (ICEARS)* (pp. 1280-1288). IEEE.

[32] Griva, A., Zampou, E., Stavrou, V., Papakiriakopoulos, D., & Doukidis, G. (2023). A two-stage business analytics approach to perform behavioural and geographic customer segmentation using e-commerce delivery data. *Journal of Decision Systems*, 1-29.