

طراحی الگوریتمی مبتنی بر اتوماتاهای یادگیر برای حل مسئله حمل و نقل و فروش کالا با تقاضاهای احتمالی

افشین مهرابی^۱، محمد رضا میبدی^۲

چکیده

در هنگام پیاده‌سازی طرحهای حمل و نقل، معمولاً اغلب پارامترهای مسئله به صورت قطعی، معین و مشخص نیستند و همچنین ممکن است این پارامترها در هنگامی که طرحهای حمل و نقل پیاده‌سازی شده‌اند در طول زمان تغییر کنند. لذا حل مسائل حمل و نقل در مقیاسهای بزرگ با پارامترهای تصادفی یکی از چالشهای مهم میباشد. از این رو برای حل این مسئله الگوریتمهای متعددی ارائه شده که در اکثر آنها نوع توزیع متغیرهای تصادفی مسئله، مشخص و پارامترهای توزیع تصادفی به صورت معلوم فرض شده است. از این جهت در این مقاله الگوریتمی مبتنی بر اتوماتای یادگیر برای حل مسئله حمل و نقل و فروش کالا پیشنهاد می - گردد که در این روش نوع توزیع متغیرهای تصادفی به صورت نامشخص فرض شده است. در این روش شبکههای از اتوماتاهای یادگیر که تشکیل یک حلقه را میدهند برای حل مسئله همکاری مینمایند. نتایج شبیه‌سازی نشان داده است که الگوریتم پیشنهادی در مقایسه با الگوریتمهای موجود از کارایی بالاتر و هوشمندی بیشتری برخوردار است و همچنین در مسائل با اندازه های بزرگ دارای سرعت و درصد همگرایی بالایی میباشد.

کلمات کلیدی

مسئله حمل و نقل و فروش کالا، اتوماتاهای یادگیر، مسائل مشکل

Learning Automata based Algorithms for Solving the Transportation and Selling Problem

Afshin mehrabi; mohamad reza meybodi

ABSTRACT

The problem mentioned in this document is using distributed learning automata to solve transportation and selling problem. In this problem each customer requires some quantity of a resource, but each customer's demand is unknown and it is stochastic. The goal of this problem is to select the set of customers that have best utilization. This problem is one of the NP-Complete problems and for this reason many approximate algorithms have been designed for solving it. In this paper an approximate algorithm based on distributed learning automata for solving transportation and selling problem with stochastic demands is proposed.

KEYWORDS

Learning automata, transportation and selling

^۱ کارشناسی ارشد کامپیوتر - نرم افزار، دانشگاه آزاد اسلامی واحد قزوین، afshin.mehrabi@gmail.com

^۲ عضو هیات علمی دانشگاه صنعتی امیر کبیر دانشکده مهندسی کامپیوتر و فناوری، mmeybodi@aut.ac.ir

مسئله حمل و نقل و فروش کالا نوع خاصی از مسئله مسیریابی وسیله نقلیه با درخواستهای احتمالی (VRPSD)^۱ میباشد. مسئله VRP دارای تقاضای احتمالی است چنانچه تقاضاها در مکان های تحویل و یا دریافت کالا یک مقدار تصادفی باشد. اولین الگوریتمی که برای این مسئله ارائه گردیده توسط تیلمن^۲ [۱] معرفی شده است که مبتنی بر الگوریتم کلارک و رایت [۲] می باشد. از دیگر کارهای مهمی که در این زمینه انجام گرفته مطالعات انجام شده توسط استوارت^۳ و گلدن^۴ [۳] و برتسیماس^۵ [۴][۵] می باشد. در کاری که توسط برتسیماس انجام شده از یک تابع ارجاعی جهت بازبینه سازی استفاده شده است. همچنین برتسیماس و سیمچی-لوی^۶ [۶] یک بررسی و مطالعه اجمالی بر روی VRPSD ارائه کرده اند. بیشترین مطالعاتی که اخیراً بر روی VRPSD صورت گرفته از نوع باز-بینه سازی^۷ می باشد که توسط سکومندی^۸ [۷] صورت گرفته است.

از مقالات مهم دیگری که در این زمینه تاکنون ارائه شده است می توان به مقاله های [۸]، [۹] و [۱۰] اشاره کرد. در مقاله [۸] چند الگوریتم مهم مبتنی بر الگوریتم های فرا-اکتشافی برای حل این مسئله نظیر کلونی مورچه ها، جستجوی محلی تکراری^۹ (ILS)، جستجوی تابو، روش سرد شدن فلزات و الگوریتم های ژنتیکی ارائه شده است. برتسیماس و دیگران در مقاله [۱۰] یک الگوریتم اکتشافی با نام اکتشافی دورهای^{۱۰} برای حل VRPSD ارائه کرده اند. در مقاله [۱۱]، یانگ^{۱۱} و دیگران چند نوع از مسائل VRPSD را در حالتی که چندین وسیله نقلیه و یا فقط یک وسیله نقلیه وجود داشته باشد مورد مطالعه خود قرار داده اند. در مقاله [۱۲] مطالعاتی که تا سال ۲۰۰۲ بروی این مسئله صورت گرفته است، جمع آوری شده است، همچنین پارامترهای گوناگونی که در مسئله VRP می تواند تصادفی باشد مورد بررسی قرار گرفته است. اما مسئله حمل و نقل و فروش کالا با مسئله VRPSD تفاوت های عمده ای دارد که این مسئله را از آن متمایز میکند. در مسئله حمل و نقل و فروش کالا، یک شرکت نقل و انتقال و فروش و تحویل کالا در یک بازه زمانی مثلاً در هر روز یک زیر مجموعه از مشتریان معین را برای تحویل کالا انتخاب میکند. سپس مشتریان را به راننده یا رانندگان ارجاع و تخصیص میدهد. در این مسئله ظرفیت و گنجایش شرکت در بازه زمانی مورد نظر، از لحاظ میزان کالا جهت ارائه، محدود و معین میباشد و از طرفی میزان نیاز و تقاضای مشتریان از پیش مشخص نمیشود و به صورت تصادفی میباشد. این مسئله شبیه به مسئله مسیریابی وسیله نقلیه با تقاضاهای احتمالی (VRPSD) می باشد با این تفاوت که در اکثر مواردی که بروی مسئله VRPSD کار شده است فرض بر این است که حتماً می بایست تمامی مشتریان سرویس داده شوند، و حال اگر در زمان سرویس دادن به مشتریان، کالا تمام شود راننده موظف است دوباره به انبار بازگردد و دوباره بارگیری کند تا به مشتریان باقیمانده سرویس داده شود. بنابراین با این ستاریو، مسئله در آنجا، شامل تصمیم گیری در مورد این خواهد بود که چه مواقعی راننده باید به انبار بازگردد و دیگر اینکه مشتریان به چه ترتیبی سرویس داده شوند. در حالی که مسئله در اینجا اندکی با مسئله آنها (VRPSD) متفاوت است و آن این است که در این مسئله، اختیار انتخاب کردن فقط یک زیر مجموعه از مشتریان برای سرویس دهی وجود دارد و هدف انتخاب یک زیر مجموعه بهینه از بین مشتریان میباشد به گونه ای که سود حاصل را افزایش و بهترین بهره وری را با ظرفیت موجود داشته باشد. از طرفی در این مسئله فرض بر این است که شبکه مشتریان جهت انتخاب، بسیار متراکم میباشد (مثلاً یک منطقه شهری پر جمعیت) پس بنابراین بحث انتخاب بهینه، بر مسیریابی و ترتیب مشتریان ارجحیت دارد و مسیریابی و ترتیب مشتریان چندان اهمیتی ندارد (میتوان به این گونه فرض کرد که در یک تور، کل کالاها که میزان آن محدود و مشخص است میبایست به مشتریان ارائه شود، و دیگر نیاز به بازگشت به انبار جهت بارگیری مجدد نیست). توجه شود که در این مسئله می بایست یک مشتری یا به طور کامل سرویس داده شده و نیاز آن برآورده شود و یا اینکه اصلاً سرویس داده نشود.

این مسئله را می توان به صورت زیر فرمول بندی کرد:

$$M = \text{تعداد مشتریان}$$

$$a = \text{ظرفیت شرکت از لحاظ میزان کالا در بازه زمانی مورد نظر}$$

$$d = \text{میزان جریمه بر واحد وزن برای اضافه ظرفیت}$$

$$r_i = \text{میزان پاداش یا اجرت بر واحد وزن برای مشتری } i$$

$$\mu_i = \text{میزان تقاضای مشتری } i \text{ برای کالا که به صورت یک مقدار تصادفی می باشد.}$$

$$X_i = \{ \text{مجموعه مشتریانی که جهت ارسال کالا انتخاب شده اند.} \}$$

$$S = \{ a | E[S] = \text{مقدار مورد انتظار برای اضافه ظرفیت برای مجموعه } S \text{ از مشتریان (انتظار اینکه میزان تقاضای مشتریانی که انتخاب شده اند،}$$

از ظرفیت مجاز فراتر رود)

حال با نمادگذاری های ذکر شده می توان هدف مسئله را به صورت زیر فرموله کرد.

$$\max(\sum_{i=1}^m r_i \mu_i x_i - d \times E[> a | \{x_i\}]) \quad (1)$$

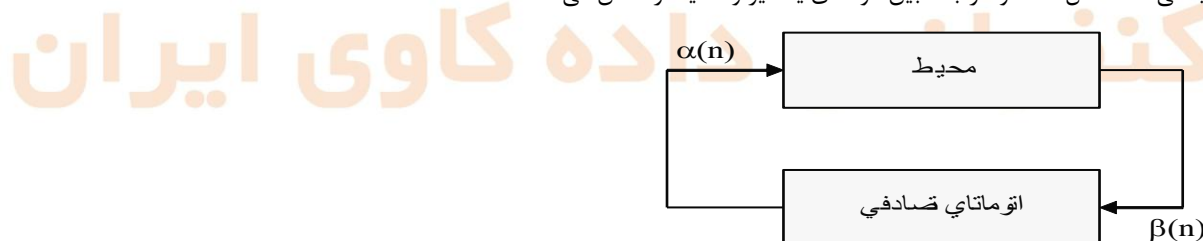
$$s.t. \quad x_i \in \{0,1\}$$

در نمادهای بالا d میزان جریمه بر واحد وزن میباشد. یعنی زمانی که میزان تقاضای مشتریانی که انتخاب شدهاند، از ظرفیت مجاز فراتر رود، به میزان هر واحد وزن که فراتر از ظرفیت مجاز میباشد جریمه‌های معادل d در نظر گرفته شده است. که این جریمه می‌تواند در دنیای واقعی به صورت هزینه مربوط به قصور در تکمیل یک معامله یا پروژه همانطور که تعهد شده است باشد و یا هزینه خریدن منابع اضافی غیر معمول با یک هزینه بالاتر باشد. به عنوان نمونه‌هایی برای این مسئله میتوان یک شرکت نقل و انتقال و فروش مواد سوختی و گرمایشی منزل را در نظر گرفت و یا یک آژانس تاکسی را که میتواند به تعدادی از خانواده‌ها در بازه‌های زمانی خاصی متعهد شود بدون آنکه اطلاع دقیقی از میزان مسافرت هر خانواده داشته باشد. الگوریتم‌های مختلفی برای مسائلی شبیه به این مسئله ارائه شده است اما ماینویل کن^{۱۳} [۱۳] در تحقیقات خود منحصرأ به این مسئله پرداخته و الگوریتمی دقیق به روش شاخه و قید برای آن ارائه داده است. او در مورد مسئله فرض را بر این گرفته است که میزان تقاضای مشتریان به صورت تصادفی با توزیع نرمال بوده و پارامترهای میانگین و واریانس متغیر تصادفی معلوم میباشد و بر اساس آن مسئله تصادفی را به یک مسئله معین تبدیل کرده و برای آن راه حل ارائه داده است. از طرفی الگوریتم‌های تقریبی متعددی برای حل مسائلی بدین گونه به صورت تصادفی ارائه شده‌اند. از جمله الگوریتم‌های تقریبی، الگوریتم‌های تکرار شونده^{۱۳} هستند. در این روشها رسیدن به یک پاسخ بهینه تضمین نمیشود اما در اغلب موارد جوابهای تقریبی قابل قبولی تولید میکنند. از جمله الگوریتم‌های تکرار شونده میتوان به الگوریتم ژنتیکی یوشیتومی^{۱۴} و همکارانش [۱۴] اشاره کرد.

در این مقاله یک الگوریتم تکرارشونده مبتنی بر اتوماتاهای یادگیر برای حل مسئله حمل و نقل و فروش کالا پیشنهاد میشود. در این الگوریتم، شبکه‌ای از اتوماتاهای یادگیر که تشکیل یک حلقه را میدهند برای حل مسئله همکاری می‌نمایند. در این الگوریتم به تعداد مشتریها، اتوماتای یادگیر در نظر گرفته شده است که هر اتوماتا متناظر با یک مشتری خواهد بود. هر اتوماتا انتخاب یا عدم انتخاب مشتری متناظر را مشخص میکند. در این شبکه در هر مرحله یک اتوماتای یادگیر فعال شده و سپس یکی از اعمال خود را انتخاب میکنند. اعمال انتخاب شده در محیط اجرا و با توجه به نتیجه عمل، به آنها پاداش و یا جریمه داده میشود و در نتیجه بردار احتمالات اتوماتاهای یادگیر بروز میشود و در خاتمه الگوریتم، بهترین راه حل بدست آمده به عنوان جواب نهایی انتخاب میشود. نتایج شبیه سازی انجام گرفته کارایی الگوریتم پیشنهادی را در مقایسه با دیگر الگوریتم‌های ارائه شده را نشان میدهد. ادامه مقاله به این صورت سازماندهی شده است، در بخش ۲ اتوماتاهای یادگیر به صورت اجمالی معرفی میگردد. در بخش ۳، الگوریتم پیشنهادی برای حل مسئله حمل و نقل و فروش کالا ارائه میشود. در بخش ۴، ارزیابیهای انجام شده را بررسی کرده و در بخش ۵، نتیجه را میتوان مشاهده کرد.

۲- اتوماتاهای یادگیر

اتوماتای یادگیر یک مدل انتزاعی است که بطور تصادفی یک اقدام از مجموعه متناهی اقدامهای خود را انتخاب کرده و بر محیط اعمال می‌کند. محیط اقدام انتخاب شده توسط اتوماتای یادگیر را ارزیابی کرده و نتیجه ارزیابی خود را توسط یک سیگنال تقویتی به اتوماتای یادگیر اطلاع می‌دهد. اتوماتای یادگیر با دریافت سیگنال و با توجه به آخرین اقدام انجام شده، وضعیت داخلی خود را بروز کرده و اقدام بعدی خود را انتخاب می‌کند. شکل ۱، نحوه ارتباط بین اتوماتای یادگیر و محیط را نشان می‌دهد.



شکل ۱: ارتباط اتوماتای یادگیر با محیط

محیط را می‌توان توسط سه تایی $E = \{\alpha, \beta, c\}$ نشان داد که در آن $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ مجموعه ورودیه‌ها، $\beta = \{\beta_1, \beta_2, \dots, \beta_r\}$ مجموعه خروجیه‌ها و $c = \{c_1, c_2, \dots, c_r\}$ مجموعه احتمالات جریمه می‌باشد. هرگاه β مجموعه دو عضوی باشد، محیط از نوع P می‌باشد. در چنین محیطی $\beta_1 = 1$ به عنوان جریمه و $\beta_2 = 0$ به عنوان پاداش در نظر گرفته می‌شود. در محیط از نوع Q ، مجموعه β دارای تعداد متناهی عضو می‌باشد و در محیط از نوع S ، تعداد اعضا مجموعه β نامتناهی است. c_i نشان دهنده احتمال نامطلوب

بودن سیگنال تقویتی محیط در پاسخ به اقدام α_i می‌باشد. در یک محیط ایستا^{۱۵} مقادیر C_i ثابت هستند، حال آنکه در یک محیط غیر ایستا^{۱۶} این مقادیر در طی زمان تغییر می‌کنند. بر اساس اینکه تابع بروز رسانی وضعیت اتوماتای یادگیر (که با اطلاع از اقدام انتخاب شده و سیگنال تقویتی β ، وضعیت بعدی اتوماتای یادگیر را محاسبه می‌کند) ثابت یا متغیر باشد، اتوماتای یادگیر به دو دسته اتوماتای یادگیر با ساختار ثابت و اتوماتای یادگیر با ساختار متغیر تقسیم می‌گردند. در این مقاله از اتوماتای یادگیر با ساختار متغیر استفاده شده است که در ادامه معرفی می‌شود. اتوماتای یادگیر با ساختار متغیر توسط چهار تایی $\{\alpha, \beta, p, T\}$ نشان داده می‌شود که در آن $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ مجموعه اقدام‌های اتوماتای یادگیر، $\beta = \{\beta_1, \beta_2, \dots, \beta_r\}$ مجموعه ورودی‌های اتوماتای یادگیر، $p = \{p_1, p_2, \dots, p_r\}$ بردار احتمال انتخاب هر یک از اقدام‌ها و $T[p(n+1), \beta(n), p(n)]$ الگوریتم یادگیری اتوماتای یادگیر می‌باشد. الگوریتم‌های یادگیری متنوعی برای اتوماتای یادگیر ارائه شده است که در ادامه یک الگوریتم یادگیری خطی برای اتوماتای یادگیر بیان می‌گردد. فرض کنید اتوماتای یادگیر در مرحله n اقدام α_i خود را انتخاب نموده و محیط ارزیابی خود را توسط سیگنال تقویتی $\beta(n)$ به اتوماتای یادگیر اعلام کند. با استفاده از الگوریتم یادگیری خطی، اتوماتای یادگیر بردار احتمال انتخاب اقدام‌های خود را مطابق رابطه زیر تنظیم می‌کند.

الف- پاسخ مطلوب از محیط

$$\begin{aligned} p_i(n+1) &= p_i(n) + a[1 - p_i(n)] \\ p_j(n+1) &= (1-a)p_j(n) \quad \forall j, j \neq i \end{aligned} \quad (2)$$

ب- پاسخ نامطلوب از محیط

$$\begin{aligned} p_i(n+1) &= (1-b)p_i(n) \\ p_j(n+1) &= \frac{b}{r-1} + (1-b)p_j(n) \quad \forall j, j \neq i \end{aligned} \quad (3)$$

که a پارامتر پاداش و b پارامتر جریمه می‌باشد. اگر a و b با هم برابر باشند، الگوریتم L_{R-P} ^{۱۷}، اگر a از b خیلی کوچکتر باشد، الگوریتم L_{R-EP} ^{۱۸} و اگر b صفر باشد، الگوریتم L_{R-I} ^{۱۹} نام دارد [۱۵].

۳- حل مسئله حمل و نقل و فروش کالا توسط اتوماتاهای یادگیر

در این بخش الگوریتم پیشنهادی برای حل مسئله حمل و نقل و فروش کالا ارائه می‌گردد. در این الگوریتم به تعداد مشتریها، اتوماتای یادگیر در نظر گرفته شده است، یعنی هر اتوماتا متناظر با یک مشتری می‌باشد. هر اتوماتا دارای دو عمل است که یک عمل انتخاب مشتری متناظر و عمل دیگر عدم انتخاب مشتری متناظر می‌باشد. در ابتدا این دو عمل دارای احتمال برابر و مساوی 0.5 می‌باشند. اتوماتاها با همدیگر تشکیل یک حلقه را می‌دهند. در هر زمان فقط یک اتوماتا درون این شبکه فعال می‌باشد. اولین اتوماتا به صورت تصادفی یکنواخت انتخاب می‌شود از آن پس مسیر حلقه برخلاف حرکت عقربه‌های ساعت از چپ به راست طی می‌شود زمانی که یک اتوماتا فعال می‌شود از بین دو عمل خود یکی را بنا به احتمال انتخاب می‌کند. زمانی که یک اتوماتا عمل خود را انتخاب می‌کند اتوماتای همسایه سمت راست این اتوماتا فعال می‌شود و آن نیز به همین ترتیب یکی از اعمال خود را انتخاب می‌کند. طی شدن این حلقه تا زمانی که ظرفیت پر شود ادامه پیدا می‌کند. اعمال انتخاب شده در محیط اجرا می‌شود و با توجه به نتیجه اعمال، به آنها پاداش و یا جریمه داده می‌شود. الگوریتم پیشنهاد شده سعی می‌کند تا بر اساس یادگیری، هوش عملیاتی را بالا برده تا در نهایت از بین آیتمها با ارزشترین آنها انتخاب شود.

الگوریتم پیشنهادی را $DLA-TPSD$ ^{۲۰} می‌نامیم. مراحل این الگوریتم به شرح زیر است. مرحله ۱: یک شبکه از اتوماتاهای یادگیر متناظر با مشتریها ایجاد می‌شود. یعنی هر اتوماتا متناظر با یک مشتری می‌باشد. این شبکه تشکیل یک حلقه را می‌دهد. هر اتوماتا دارای دو عمل است که یک عمل انتخاب مشتری متناظر و عمل دیگر عدم انتخاب مشتری متناظر می‌باشد. در ابتدا این دو عمل دارای احتمال برابر و مساوی 0.5 می‌باشند. توجه شود که هر مشتری را فقط و فقط یک بار میتوان انتخاب کرد. در این مسئله ارزش و میزان تقاضای مشتریها نامشخص است و زمانی میتوان به ارزش و میزان تقاضای مشتری پی برد که آن مشتری انتخاب شود. در ضمن میزان ظرفیت و گنجایش مشخص می‌باشد. اتوماتاهای یادگیر از الگوریتم یادگیر L_{R-EP} ^{۲۱} استفاده می‌کنند. بردار احتمال و عملهای اتوماتاهای یادگیر به ترتیب $p_m = (p_1, p_2)$ و $\alpha_m = \{\alpha_1, \alpha_2\}$ می‌باشد. مقدار اولیه اعضای بردار احتمال برابر و مساوی 0.5 در نظر گرفته می‌شود. محیط به صورت مدل P پیاده‌سازی شده است. ورودی محیط یک مجموعه از اعمال انتخاب شده توسط اتوماتاها می‌باشد. خروجی (پاسخ) محیط به اعمال، توسط β_i مشخص می‌شود. β_i یک پاسخ دودویی می‌باشد، در چنین محیطی $\beta_i(n) = 1$ بعنوان پاسخ نامطلوب یا شکست و $\beta_i(n) = 0$ بعنوان پاسخ مطلوب یا موفقیت در نظر گرفته می‌شود.

مرحله ۲: اولین اتوماتا به صورت تصادفی یکنواخت انتخاب شده و فعال میشود و یک عمل از مجموعه اعمال خود را به صورت تصادفی و مطابق بردار احتمال $p(n)$ ، انتخاب میکند. به عبارت دیگر در این مرحله مشتری متناظر با اتوماتا یا انتخاب میشود و یا انتخاب نمیشود. در صورتیکه عمل انتخاب مشتری برگزیده شود، اگر مقدار مورد انتظار (امید ریاضی) میزان تقاضای مشتری از ظرفیت بیشتر نشود آن مشتری انتخاب میشود. (بر اساس تجربیات گذشته، امید ریاضی میزان تقاضای مشتری محاسبه میشود) با انتخاب مشتری ارزش و میزان تقاضای واقعی مشتری نمایان میشود. در این مرحله از تابع برازش $f(i) = r_i \mu_i$ به منظور سنجش ارزش مشتری استفاده شده است و مقدار محاسبه شده، به مقدار برازش کل انتخابهای موجود، که آن را با $C(K)$ نشان میدهیم جمع بسته میشود.

$$C(K) = \sum f_i \quad : i \in \{1, 2, \dots, m\} \quad (4)$$

مرحله ۳: از این پس مسیر حلقه برخلاف حرکت عقربه‌های ساعت از چپ به راست طی میشود. زمانی که یک اتوماتا عمل خود را انتخاب میکند اتوماتای همسایه سمت راست این اتوماتا فعال میشود و آن نیز به همین ترتیب یکی از اعمال خود را انتخاب میکند. مرحله ۴: مرحله ۳ تا زمانی که گنجایش به اتمام نرسیده باشد یا اینکه دیگر مشتری ای باقی نمانده باشد که میزان تقاضای مورد انتظار آن، از گنجایش باقی مانده، کمتر باشد تکرار میشود. بعد از اتمام مرحله انتخاب، میزان اضافه ظرفیت چک میشود، که به ازاء هر واحد اضافه ظرفیت، به میزان d از مقدار ارزش کل انتخابهای موجود کاسته میشود. پس اگر اضافه ظرفیت را با Z نشان دهیم، خواهیم داشت:

$$C(K) = \sum f_i - (d \times Z) \quad : i \in \{1, 2, \dots, m\} \quad (5)$$

مرحله ۵: بعد از اتمام ظرفیت، مقدار $C(K)$ با میانگین برازش انتخابهای قبلی اتوماتاها در دفعات قبل که پاداش دریافت کرده‌اند، مقایسه میشود. میانگین برازش انتخابهای قبلی را با $AVC(K_n)$ ^{۲۲} نشان می دهیم. اگر مقدار $C(K)$ کوچکتر از میزان $AVC(K_n)$ باشد به عمل انتخاب شده توسط اتوماتاهای یادگیر جریمه و در غیر این صورت به اعمال انتخاب شده توسط اتوماتاهای یادگیر پاداش تعلق میگیرد. در پایان این مرحله میزان $AVC(K_n)$ دوباره محاسبه میشود که فرمول آن به قرار رابطه (۶) است:

$$AVC(K_n) = \frac{\sum_{l=1}^n C(k_l)}{n} \quad (6)$$

در اینجا n ، تعداد دفعات قبلی میباشد که کوله پشتی پر شده و نیز نتیجه مطلوب بوده و منجر به دادن پاداش به اعمال انتخاب شده توسط اتوماتاهای یادگیر شده است و $C(K_l)$ مقدار برازش حاصل از انتخاب مشتریان، در دفعه L ام که اعمال انتخاب شده پاداش دریافت کرده‌اند، میباشد.

مرحله ۶: اگر تعداد اجرای الگوریتم از یک تعداد از پیش تعیین شده بیشتر شود الگوریتم خاتمه می‌یابد و در غیر اینصورت مرحله ۲ تا ۶ تکرار می‌گردد. شبه کد الگوریتم پیشنهادی در شکل ۲ آمده است.

کنفرانس داده کاوی ایران

```

Procedure DLA-TPSD
Input: costumers from dataset
Output: best selected costumers, fitness of selected costumers
Begin
    Construct array of learning automata as a ring //the length of array is equal costumers number
    Repeat
        auto_id = Random(max value= automata number)
        Repeat
            For all automata in ring do
                Automata[auto_id].select_action()
                auto_id = next_id()
            End For
        Until (capacity_is_full)
        If ( $C^{(K)} < AVC(K_n)$ ) then
            For all automata in ring do
                Penalize(selected_action)
            Else
                For all automata in ring do
                    Reward(selected_action)
                End For
            Compute ( $AVC(K_n)$ )
            Insert selected costumers and fitness of them into database
        Until (current repeat  $\geq$  iteration number)
    End. DLA-TPSD

```

شکل ۲: شبه کد الگوریتم پیشنهادی DLA-TPSD برای حل مسئله TPSD

۴- ارزیابی

الگوریتم پیشنهادی DLA-TPSD، بروی مجموعه مشتریهایی با مشخصات متفاوت از لحاظ ارزش، میزان تقاضا و میزان ظرفیت، آزمایش گردیده است و با الگوریتم شاخه‌بند به روش اکتشافی که توسط ماینویل کن، روی همین مسئله ارائه گردیده است [۱۳]، و نیز الگوریتم ژنتیک که توسط یوشیتومی^{۲۳} و هیروکو^{۲۴} برای محیطهای غیر قطعی و مسائل تصادفی ارائه گردیده است [۱۴]، مقایسه میشود. به منظور ارزیابی از دادههایی که در [۱۳] ذکر شده است استفاده میشود، که در ادامه توضیح مختصری در مورد این دادهها، بیان میشود. همانطور که در جدول ۱ مشاهده میشود مجموعه های متفاوتی از لحاظ تعداد و نوع متغیرهای تصادفی در نظر گرفته شده است که ارزیابیها روی آنها انجام میشود. در این مجموعهها متغیرهای تصادفی پیوسته به صورت توزیع نرمال در نظر گرفته شدهاند. اگر σ^2_i واریانس متغیر تصادفی میزان تقاضای مشتری i ام باشد، $\max \sigma^2$ ماکزیمم واریانس و $\min \sigma^2$ مینیمم واریانس توزیعهای احتمالی میزان تقاضای مشتریهای مجموعه می باشد. اگر μ_i میانگین توزیع احتمال برای مشتری i ام باشد، در این صورت $\max \mu$ ماکزیمم میانگین و $\min \mu$ مینیمم میانگین توزیعهای احتمالی میزان تقاضای مشتریهای مجموعه می باشد. در مجموعه ۱، میزان تقاضای مشتریان بر اساس توزیع نرمال میباشد که در آن میانگینها متفاوت و واریانس نیز متفاوت میباشد فاکتور اجرت هر مشتری نیز به صورت تصادفی گسسته میباشد که نوع توزیع احتمال آن در

جدول ۲ قابل مشاهده است. در مجموعه ۲، میانگین و واریانس میزان تقاضای مشتریان متفاوت می باشد و میزان فاکتور اجرت به صورت ثابت و یکی از مقادیر ۱ تا ۳ میباشد. در مجموعه ۳، میانگین میزان تقاضای هر مشتری متفاوت ولی واریانس آنها برابر میباشد و میزان فاکتور اجرت به صورت ثابت و یکی از مقادیر ۱ تا ۴ میباشد. در مجموعه ۴، میانگین میزان تقاضای هر مشتری برابر ولی واریانس آنها متفاوت است و میزان فاکتور اجرت هر مشتری به صورت ثابت و یکی از مقادیر ۲ یا ۴ میباشد.

جدول (۱) مشخصات مجموعه دادههای مورد استفاده در آزمایشات

شماره مجموعه	تعداد مشتری	میزان تقاضای مشتریان			میزان پارامتر اجرت مشتری		ظرفیت
		$\min \mu - \max \mu$	$\min \sigma^2 - \max \sigma^2$	نوع متغیر تصادفی	ثابت	تصادفی	
مجموعه ۱	10	170 – 225	5 -50	پیوسته	-	گسسته	1100
مجموعه ۲	15	180 -260	5 -50	پیوسته	ثابت 3-1	-	2000
مجموعه ۳	20	280 -700	25	پیوسته	ثابت 4-1	-	3850
مجموعه ۴	40	200	1 -145	پیوسته	ثابت 4-2	-	4400

جدول (۲) تابع توزیع احتمال میزان اجرت مشتریان در مجموعه ۱

items	cost	
	value	probability
Costumer1	1, 2, 3	0.2, 0.4, 0.4
Costumer2	2	1
Costumer3	1, 2	0.3, 0.7
Costumer4	1	1
Costumer5	3, 4	0.4, 0.6
Costumer6	2, 3	0.2, 0.8
Costumer7	2, 3, 4	0.3, 0.3, 0.4
Costumer8	3	1
Costumer9	2, 4	0.5, 0.5
Costumer10	1, 3	0.4, 0.6

برای بررسی عملکرد الگوریتم پیشنهادی و مقایسه آن با الگوریتمهای موجود پنج معیار زیر استفاده می شود :

حاصلضرب احتمالات انتخاب بهینه از بین مشتریان

میانگین ارزش انتخابهای حاصل از اجرای الگوریتم (Av-Cost)

درصد اجراهای همگرا شده که با PC^{25} نشان میدهم

میانگین تعداد تکرار برای اجراهای همگرا شده که با AVI^{26} نشان میدهم.

میانگین تعداد نمونهگیریهای انجام شده از کل مشتریان که با Av-Sam No نشان میدهم.

آزمایشاتی جهت بررسی عملکرد الگوریتمهای پیشنهادی طراحی شده است. در این آزمایشات هرکدام از معیارهای معرفی شده مورد بررسی

قرار گرفته است. نتایج حاصل از اجرای هر الگوریتم در دیاگرام و جداولی که نشاندهنده میزان معیارهای ذکر شده است، نشان داده می شود.

با توجه به تعدد حالات ممکن در آزمایشها، ما یک استاندارد ساده برای نامگذاری آنها ارائه کرده ایم. نام هر آزمایش از دو قسمت که با خط

تیره (-) از هم جدا شدهاند تشکیل شده است. به ترتیب از سمت چپ به راست، هر بخش از این اسم به معنای آنچه که در ادامه آمده است می

باشد:

نوع مجموعه آیتمهای مورد استفاده در آزمایش (مجموعه ۱=۱، مجموعه ۲=۲، مجموعه ۳=۳ و مجموعه ۴=۴)

نوع الگوریتم یادگیر و میزان پارامتر پاداش و جریمه که به صورت جدول ۳ که در ادامه آمده است، تعیین گردیده اند. (A, B, C,

D, E

جدول (۳) میزان پارامتر پاداش و جریمه

میزان پارامتر جریمه (b)		میزان پارامتر پاداش (a)		شناسه
$L_{R\&P}$	A	0.08	0.01	
	B	0.08	0.002	
	C	0.06	0.009	
	D	0.06	0.002	
L_{Ri}	E	0.08	0	

در ضمن شرط همگرایی در همه این آزمایشها، بزرگتر بودن حاصلضرب احتمالات انتخاب مشتریهای انتخاب بهینه، از (۰.۹) میباشد.

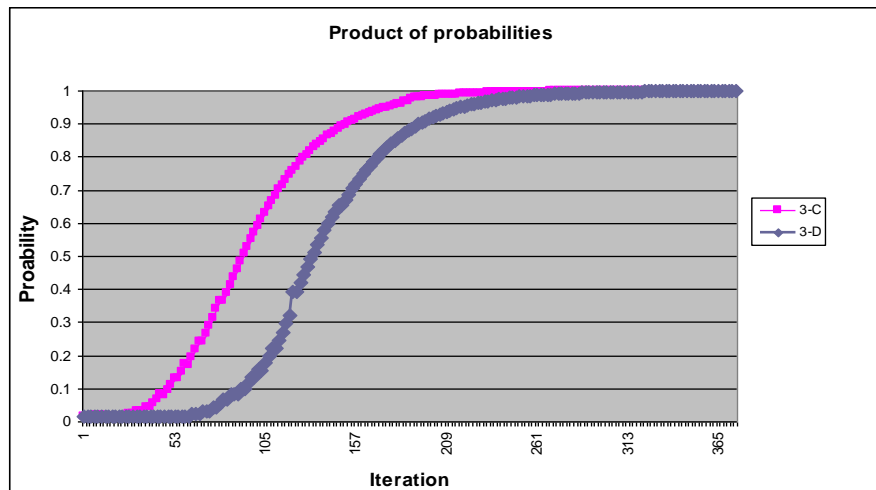
همچنین تعداد تکرار در حالت عادی ۲۵۰۰ می باشد که هر آزمایش را ۱۰ بار اجرا نموده و خروجی الگوریتمهای مختلف را در جداول و نمودارهایی که در ادامه مشاهده میشود لحاظ نموده ایم.

جدول (4) جزئیات نتایج حاصل از اجرای الگوریتم پیشنهاد شده در آزمایشهای ذکر شده

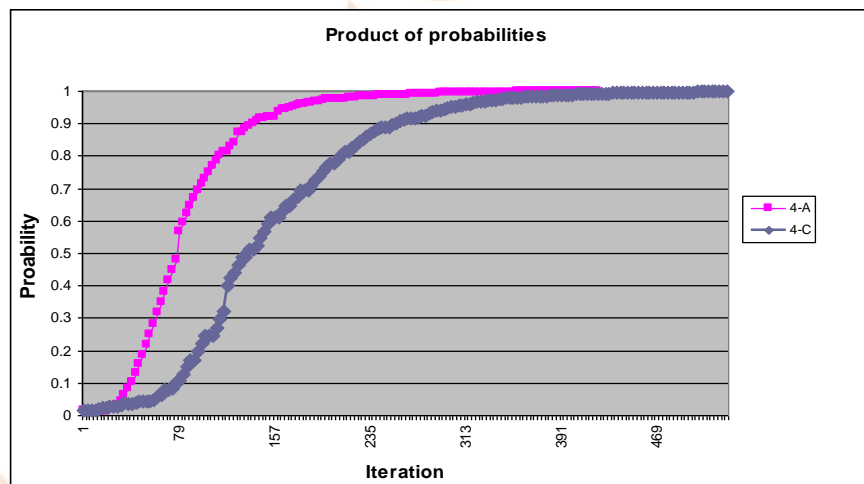
Method	PC	AVI	Av-Sam No
1-A	0	0	10803.2
1-B	0	0	12260
1-C	0	0	11090.3
1-D	0	0	12240.4
1-E	90	679	11650
2-A	60	1151.1	23200.5
2-B	10	960.2	20865
2-C	40	805	21570.2
2-D	10	1067.5	20490.1
2-E	0	0	20782
3-A	10	1034	21640.6
3-B	40	381.2	22440.1
3-C	80	266.6	22666
3-D	70	287	22574.4
3-E	10	458	21236.7
4-A	100	264.4	53520
4-B	60	973.1	53720.2
4-C	90	278.4	53272.5
4-D	50	986	53560
4-E	40	1027.4	52974.4

با مشاهده نتایج بدست آمده و دقت در آنها میتوان متوجه شد که در مجموعه دادههای شماره ۱، از بین حالتهای مختلف جهت دادن پاداش و جریمه، تنها گونه E بوده است که جواب مطلوبی داشته است و بقیه گونهها اصلاً بروی جواب بهینه همگرا نشدهاند. این بدان جهت است که مقادیر پارامتر اجرت مشتریان به صورت تصادفی گسسته بودهاند و چون پراکندگی و میزان واریانس این مقادیر برای هر مشتری زیاد است لذا جواب بهینه که شامل مشتریان ۵، ۶، ۷، ۸ و ۹ میباشد نیز همین پراکندگی زیاد را در میزان ارزش خود خواهد داشت. لذا در بعضی مواقع میزان ارزش جواب بهینه بالاتر از میانگین ارزش انتخابهای قبلی (حد آستانه) قرار میگیرد و در بعضی مواقع ارزش همان جواب بهینه پایینتر از میانگین ارزش انتخابهای قبلی قرار میگیرد، از این رو انتخاب جواب بهینه دائماً با پاداش و جریمه روبرو میشود و بدین سبب است که گونه های A الی D جهت دادن پاداش و جریمه در رابطه با این مجموعه مناسب نیستند و تنها گونه E که فاقد جریمه است برای این مجموعه از دادهها، مناسب است. در مجموعه دادههای شماره ۲، گونه A، جهت دادن پاداش و جریمه بهتر از بقیه گونه ها بوده است و با انتخاب گونه A درصد همگرایی بیشتر بوده است و گونه C، تنها رقیب آن جهت همگرایی بوده است هر چند که در گونه A، تعداد تکرارهای بیشتری جهت همگرایی، لازم بوده است. در مجموعه ۳، گونه C و D بهتر از بقیه گونهها جواب دادهاند. با مقایسه نتایج بروی مجموعه های ۲ و ۳، می توان مشاهده کرد که الگوریتم پیشنهاد شده بروی مجموعه دادههای ۳ بهتر عمل کرده است چرا که در مجموعه ۳، واریانس میزان تقاضای مشتریان یکسان است در حالی که واریانس میزان تقاضای مشتریان در مجموعه ۲، متفاوت و نیز بیشتر است. و از طرفی اختلاف میانگین در میزان تقاضای مشتریان در مجموعه ۳، بیشتر از مجموعه ۲ میباشد. الگوریتم بروی مجموعه دادههای شماره ۴، نسبت به بقیه مجموعهها، بهتر جواب داده است، چرا که میانگین میزان تقاضای مشتریان یکسان است هر چند که دارای بیشترین واریانس در بین مجموعهها است ولی در نهایت در این مجموعه میزان ارزش هر مشتری به میزان اجرت آن مشتری بستگی خواهد داشت. گونه A، در این مجموعه بهتر از بقیه گونهها جواب داده است. با مقایسه کلی این نتایج میتوان نتیجه گرفت که تعداد و کمیت مشتریان در رسیدن یا نرسیدن به جواب بهینه تاثیر چندانی نداشته بلکه میزان پراکندگی مقادیر تصادفی و نوع دادههای تصادفی بوده است که در رسیدن به جواب بهینه و تعیین کارایی الگوریتم موثر است.

در ادامه حاصلضرب احتمالات انتخاب بهینه از بین مشتری ها در طول اجرای الگوریتمهای پیشنهاد شده، مورد بررسی قرار گرفته شده است. مقدار حاصلضرب احتمالات میبایست با تکرارهای بیشتر و متوالی، به سمت یک، میل کند. که به عنوان نمونه نمودارهای مربوط به آزمایشات ۳-C، ۳-D، ۴-A، ۴-C و ۴-C نشان داده شده است.



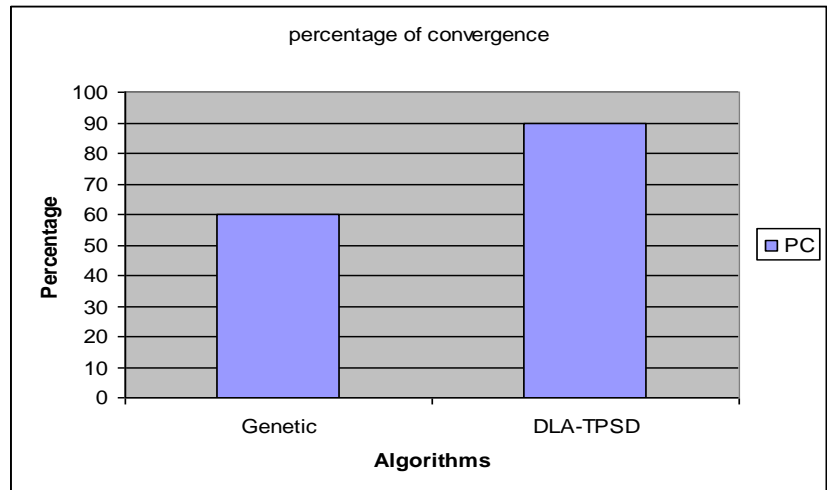
شکل 3: حاصلضرب احتمالات انتخاب بهینه از بین مشتریها در الگوریتم پیشنهادی، در آزمایش 3-C و 3-D



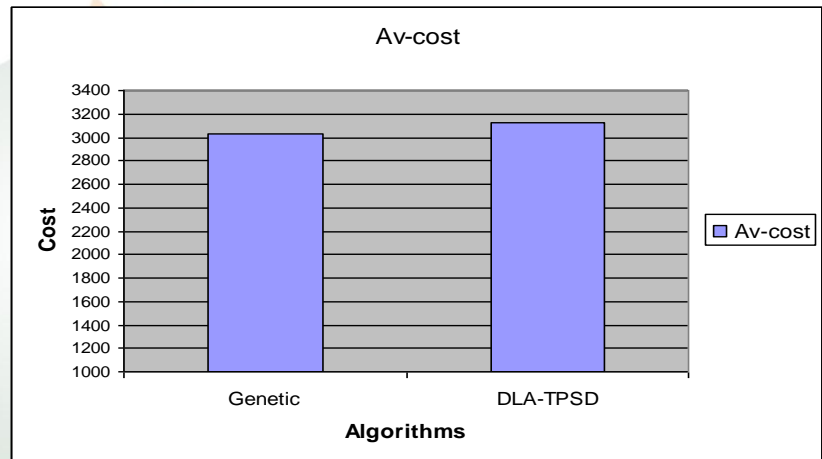
شکل 4: حاصلضرب احتمالات انتخاب بهینه از بین مشتریها در الگوریتم پیشنهادی، در آزمایش 4-A و 4-C

آزمایشات نشان داده است که الگوریتم ذکر شده در مجموعه دادههای متفاوت چه از لحاظ تعداد مشتریها، میزان ظرفیت و تصادفی بودن میزان تقاضای مشتریان یا اجرت آنها یا هر دو و چه از لحاظ پراکنگی میزان متغیرهای تصادفی میزان تقاضا و اجرت مشتریان و حتی نوع توزیع از لحاظ پیوسته یا گسسته بودن، نتایج مطلوبی در کل داشته است.

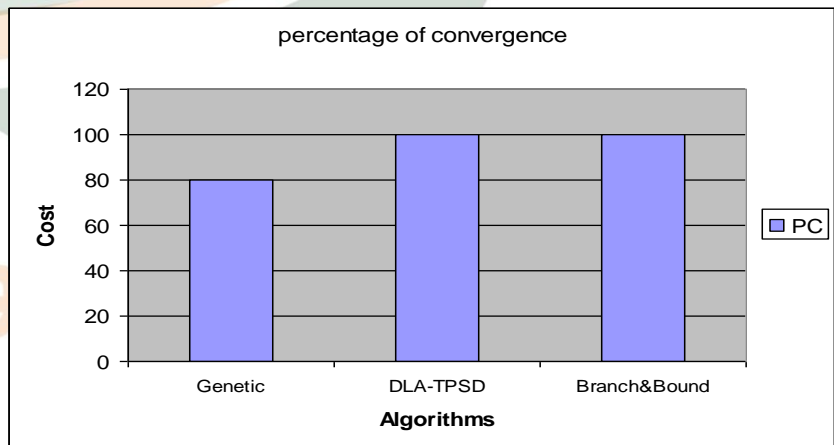
حال جهت بررسی عملکرد الگوریتم پیشنهادی در مقایسه با الگوریتم ژنتیک که توسط یوشیتومی و هیروکو ارائه گردیده است [۱۴]، و نیز الگوریتم شاخهوقید به روش اکتشافی که توسط ماینویل کن، روی همین مسئله ارائه گردیده است [۱۳]، آزمایشاتی طراحی شده است. در این آزمایشها معیار میانگین ارزش نمونههای انتخاب شده توسط الگوریتم و نیز درصد اجراهایی که به پاسخ بهینه خواهند رسید در اجراهای متوالی مورد بررسی قرار گرفته است. در این آزمایشات، در الگوریتم ژنتیک سائز جمعیت برابر ۵۰۰، تعداد نسل برابر ۱۵۰۰، احتمال **Cross over** برابر ۰/۶ و احتمال **mutation** برابر ۰/۱ در نظر گرفته شده است. از آنجایی که در این الگوریتم ژنتیک در هر بار اجرا، یک جواب با بیشترین تکرار در کل نسلا را، به عنوان تنها جواب بهینه معرفی میکند، این الگوریتم بروی دادههای یاد شده، ۱۰ مرتبه اجرا میشود و نتایج حاصل با نتایج الگوریتمهای پیشنهاد شده که آنها نیز بروی دادههای مشابه ۱۰ مرتبه اجرا شدهاند، مقایسه شده است. نتایج در شکلهای ۵، ۶، ۷ و ۸ قابل مشاهده است. در این شکلهای، نتایج آزمایشات بروی مجموعه ۱ و ۴ نشان داده شده است. نتایج آزمایشات بروی مجموعه ۲ و ۳ تقریباً مشابه با این نتایج می باشد.



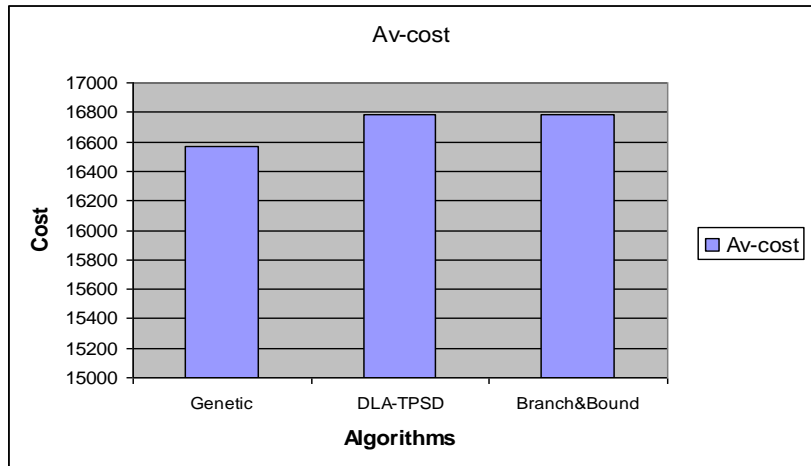
شکل 5: مقایسه الگوریتم پیشنهاد شده با الگوریتم ژنتیک بر اساس درصد تکرارهایی که به جواب بهینه رسیده‌اند در مجموعه 1



شکل 6: مقایسه الگوریتم پیشنهاد شده با الگوریتم ژنتیک بر اساس میانگین ارزش مشتریهای انتخاب شده در مجموعه 1

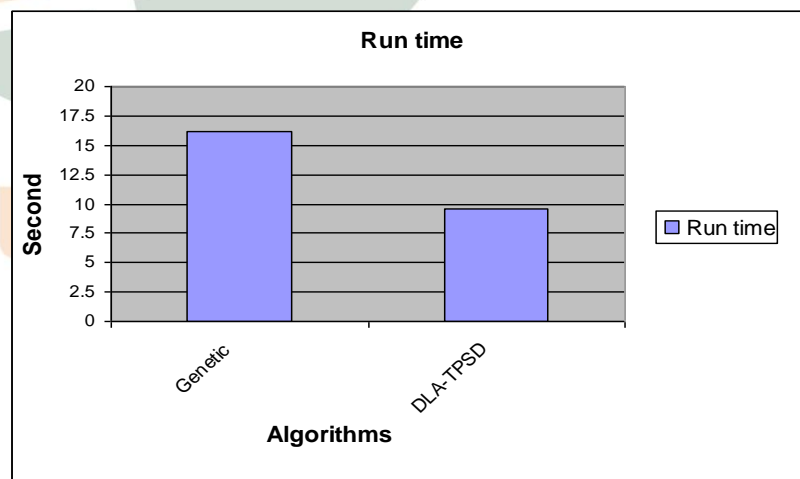


شکل 7: مقایسه الگوریتمها بر اساس درصد تکرارهایی که به جواب بهینه رسیده‌اند در مجموعه 4



شکل 8: مقایسه الگوریتمها بر اساس میانگین ارزش مشتریهای انتخاب شده در مجموعه 4

همانطور که مشاهده میشود از لحاظ ارزش نمونههای انتخاب شده و درصد همگرایی، الگوریتم پیشنهاد شده کارایی بالاتری به نسبت الگوریتم ژنتیک داشته است. نقطه ضعف الگوریتم ژنتیک در این است که ارزش نمونههای انتخاب شده در طول زمان بهبود نمییابد و کاملاً انتخابها به صورت تصادفی انجام میشود اما در مقابل، در الگوریتم پیشنهاد شده، افزایش هوشمندی در طول زمان کاملاً مشهود است و انتخاب های بعدی نسبت به انتخابهای قبلی، خیلی هوشمندانهتر انتخاب شده و از حالت تصادفی بودن خارج میشود و این برتری الگوریتم DLA_TPSD را نشان میدهد. اما در الگوریتم شاخه و قید که جزء الگوریتم های قطعی میباشد نسبت به بقیه الگوریتمها کارتر است، اما در مقابل دو نقطه ضعف بسیار بزرگ دارد و آن این است که در این الگوریتم فرض بر این است که میزان تقاضای مشتریان بر اساس توزیع نرمال است و پارامترهای توزیع رکن مشخص است و اگر نوع توزیع احتمال میزان تقاضا، شناخته شده نباشد این الگوریتم کارا نخواهد بود و دیگر اینکه در این الگوریتم فرض بر این است که میزان اجرت مشتری ثابت و معین است و اگر میزان اجرت به صورت تصادفی باشد باز این الگوریتم کارا نخواهد بود. در ادامه زمان اجرای الگوریتم پیشنهاد شده با زمان اجرای الگوریتم ژنتیک مقایسه شده است. این مقایسه بروی یک کامپیوتر شخصی با مشخصات RAM ۲GB، CPU ۳GHz و سیستم عامل Win-server ۲۰۰۳ انجام شده است. در شکل ۹، نتیجه مقایسه زمان اجرای الگوریتم پیشنهاد شده با زمان اجرای الگوریتم ژنتیک بروی مجموعه ۴ از مجموعه مشتریان، قابل مشاهده است. همان طور که مشاهده می شود زمان اجرای الگوریتم پیشنهاد شده به مراتب از زمان اجرای الگوریتم ژنتیک کمتر میباشد. همانند همین آزمایش بروی مجموعههای ۱، ۲ و ۳ نیز به انجام رسیده است که نتیجه بدست آمده بروی آنها نیز تقریباً مانند این آزمایش میباشد و در آنها نیز زمان اجرای الگوریتم پیشنهاد شده از الگوریتم ژنتیک بسیار کمتر میباشد.

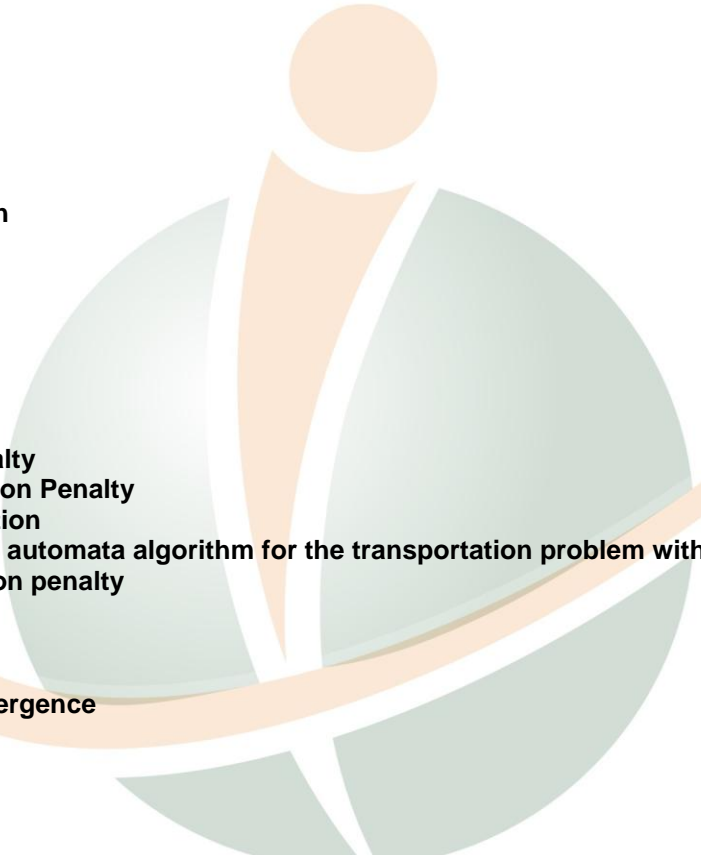


شکل 9: مقایسه الگوریتم پیشنهاد شده با الگوریتم ژنتیک بر اساس زمان اجرا بروی مجموعه 4

در این مقاله یک الگوریتم جدید مبتنی بر اتوماتاهای یادگیر برای حل مسئله حمل و نقل و فروش کالا پیشنهاد گردید. در این الگوریتم شبکه‌ای از اتوماتاهای یادگیر برای حل مسئله همکاری می‌نمایند. در این شبکه در هر مرحله یک اتوماتای یادگیر فعال شده و سپس یکی از اعمال خود را انتخاب می‌کنند. اعمال انتخاب شده در محیط اجرا و با توجه به نتیجه عمل، به آنها پاداش و یا جریمه داده می‌شود. الگوریتم پیشنهاد شده سعی می‌کند که با حداقل تعداد نمونه‌گیری از آیتها در شرایطی که توزیع احتمال میزان تقاضا یا اجرت مشتریها از قبل شناخته شده نیست بهترین مشتریها را برای پر کردن ظرفیت انتخاب نماید. الگوریتم پیشنهاد شده سعی می‌کند تا بر اساس یادگیری، هوش عملیاتی را بالا برده تا در نهایت با ارزشترین مشتریها انتخاب شود. به منظور بررسی کارایی، الگوریتم پیشنهادی بر روی داده‌های استاندارد، آزمایش و نتایج آن بروی شیوه‌های مختلفی که ارائه شده است نشان داده شده است. نتایج بدست آمده با نتایج حاصل از الگوریتم‌های موجود مقایسه شده است. نتایج مقایسه حاکی از هوشمندی و نیز کارایی بالای الگوریتم پیشنهادی در مقایسه با الگوریتم‌های موجود می‌باشد.

مراجع

- [1] F. A. Tillman, "The Multiple Terminal Deliver Problem with Probabilistic Demands", Transportation Science, Volume 3, Pages 192-204, 1969.
- [2] G. Clarke and J. Wright, "Scheduling of Vehicles from a Central Depot to a Number of Delivery Points", Operations Research, Volume 12, No. 4, Pages 568-581, 1964.
- [3] W. Stewart and B. Golden, "Stochastic Vehicle Routing: A Comprehensive Approach", European Journal of Operational Research, Volume 14, 1983.
- [4] D. J. Bertsimas, "A Vehicle Routing Problem with Stochastic Demand", Operations Research, Volume 40, No. 3, Pages 574-585, 1992.
- [5] D. P. Bertsekas and J. N. Tsitsiklis, "An Analysis of Stochastic Shortest Path Problems", Mathematics of Operations Research, Volume 16, No. 3, Pages 580-595, August 1991.
- [6] D. J. Bertsimas and D. Simchi-Levi, "A New Generation of Vehicle Routing Research: Robust Algorithms, Addressing Uncertainty", Operations Research, Volume 44, No. 2, Pages 286-304, 1996.
- [7] N. Secomandi, "A Rollout Policy for the Vehicle Routing Problem with Stochastic Demands", Operations Research, Volume 49, No. 5, Pages 796-802, 2001.
- [8] L. Bianchi, M. Birattari, M. Chiarandini, M. Manfrin, M. Mastrolilli, L. Paquete, O. Rossi-Doria and T. Schiavinotto, "Meta Heuristics for the Vehicle Routing Problem with Stochastic Demands", Proceedings of Lecture Notes in Computer Science, Springer-Verlag, Heidelberg, Germany, Volume 3242, Pages 450-460, 2004.
- [9] W. Yang, K. Mathur and R. H. Ballou, "Stochastic Vehicle Routing Problem with Restocking", Transportation Science, Volume 34, No. 1, Pages 99-112, 2000.
- [10] D. J. Bertsimas, P. Chervi and M. Peterson, "Computational Approaches to Stochastic Vehicle Routing Problems", Transportation Science, Volume 29, No. 4, Pages 342-352, 1995.
- [11] G. B. Dantzig and R.H. Ramser, "The Truck Dispatching Problem", Management Science, Volume 6, Pages 80-91, 1959.
- [12] A. Kenyon and D. P. Morton, "A Survey on Stochastic Location and Routing Problems", Central European Journal of Operations Research, Volume 9, Pages 277-328, 2002.
- [13] A. M. Cohn, "The Stochastic Knapsack Problem with Random Weights: A Heuristic Approach to Robust Transportation Planning", Cynthia Barnhart, MIT, Massachusetts Institute of Technology Cambridge, MA 02139, 2000.
- [14] Y. Yoshitomi, H. Ikenoue, T. Takeba, S. Tomita, "Genetic Algorithm in Uncertain Environments for Solving Stochastic Programming Problem", Journal of the Operations Research, Society of Japan, Volume 43, No. 2, June 2000.
- [15] K. S. Narendra, M. A. L. Thathachar, "Learning Automata: An Introduction", Prentice Hall, 1989.



¹ Vehicle Routing Problem with Stochastic Demand
² Tillman
³ Stewart
⁴ Golden
⁵ Bertsimas
⁶ Simchi-Levi
⁷ Re-optimization
⁸ Secomandi
⁹ Iterated Local Search
¹⁰ Cyclic Heuristic
¹¹ Yang
¹² Mainville Cohn
¹³ Iterative
¹⁴ Yoshitomi
¹⁵ Stationary
¹⁶ Non-Stationary
¹⁷ Linear Reward-Penalty
¹⁸ Linear Reward epsilon Penalty
¹⁹ Linear Reward Inaction
²⁰ Distributed learning automata algorithm for the transportation problem with stochastic demand
²¹ Linear reward epsilon penalty
²² Average cost
²³ Yoshitomi
²⁴ Hiroko
²⁵ Percentage of convergence
²⁶ Average of iteration

کنفرانس داده کاوی ایران