

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

افشین مهرابی<sup>۱</sup>، محمد رضا میبدی<sup>۲</sup>

## 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 تصادفی به صورت معلوم فرض شده است. از این جهت در این مقاله الگوریتمی مبتنی بر آtomataهای یادگیر برای حل مسئله حمل و نقل و فروش کالا پیشنهاد می‌گردد که در این روش نوع توزیع متغیرهای تصادفی به صورت نامشخص فرض شده است. در این روش شبکه‌های از آtomataهای 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 اندازه‌های بزرگ دارای سرعت و درصد همگرایی بالایی می‌باشد.

## کلمات کلیدی

مسئله حمل و نقل و فروش کالا، آtomataهای 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

<sup>۱</sup> کارشناسی ارشد کامپیوتر - نرم افزار، دانشگاه آزاد اسلامی واحد قزوین.

<sup>۲</sup> عضو هیات علمی دانشگاه صنعتی امیر کبیر دانشکده مهندسی کامپیوتر و فناوری،

مسئله حمل و نقل و فروش کالا نوع خاصی از مسئله مسیریابی وسیله نقلیه با درخواستهای احتمالی (VRPSD)<sup>۱</sup> میباشد.

مسئله VRP دارای تقاضای احتمالی است چنانچه تقاضاهای در مکان های تحویل و یا دریافت کالا یک مقدار تصادفی باشد. اولین الگوریتمی که برای این مسئله ارائه گردیده توسط تیلمن<sup>[۱]</sup> معرفی شده است که مبتنی بر الگوریتم کلارک و رایت<sup>[۲]</sup> میباشد. از دیگر کارهای مهمی که در این زمینه انجام گرفته مطالعات انجام شده توسط استوارت<sup>۳</sup> و گلدن<sup>۴</sup> [۳][۴] و برتسیماس<sup>۵</sup> [۵] میباشد. در کاری که توسط برتسیماس انجام شده از یکتابع ارجاعی جهت بازبینیه سازی استفاده شده است. همچنین برتسیماس و سیمچی-لوی<sup>۶</sup> [۶] یک بررسی و مطالعه اجمالی بروی VRPSD ارائه کرده‌اند. بیشترین مطالعاتی که اخیراً بروی VRPSD صورت گرفته از نوع باز-بینیه سازی<sup>۷</sup> میباشد که توسط سکوندی<sup>۸</sup> [۷] صورت گرفته است.

از مقالات مهم دیگری که در این زمینه تاکنون ارائه شده است میتوان به مقاله های [۸]، [۹] و [۱۰] اشاره کرد. در مقاله [۸] چند الگوریتم مهم مبتنی بر الگوریتم‌های فرا-اکتشافی برای حل این مسئله نظیر کلونی مورچه‌ها، جستجوی محلی تکراری<sup>۹</sup> (ILS)، جستجوی تابو، روش سرد شدن فلزات و الگوریتم‌های زنتیکی ارائه شده است. برتسیماس و دیگران در مقاله [۱۰] یک الگوریتم اکتشافی با نام اکتشافی دورهای<sup>۱۰</sup> برای حل VRPSD ارائه کرده‌اند. در مقاله [۱۱]، یانگ<sup>۱۱</sup> 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 داده‌اند. 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 میتواند تصادفی باشد مورد بررسی قرار گرفته است. اما مسئله حمل و نقل و فروش کالا با مسئله VRPSD تفاوت‌های عمدی 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 نمیباشد و به صورت تصادفی میباشد. این مسئله شبیه به مسئله مسیریابی وسیله نقلیه با تقاضاهای احتمالی (VRPSD) میباشد با این تفاوت که در اکثر مواردی که بروی مسئله VRPSD کار شده است فرض بر این است که حتماً میباشد 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 سناریو، مسئله در آنجا، شامل تصمیم گیری در مورد این خواهد بود که چه موقعی 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 گونه‌ای که سود حاصل را افزایش و بهترین بهره‌وری 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 کالاهای که میزان آن محدود و مشخص است میباشد به مشتریان ارائه شود، و دیگر نیاز به بازگشت به انبار جهت بارگیری مجدد نیست). توجه شود که در این مسئله میباشد یک مشتری یا به طور کامل سرویس داده شده و نیاز آن برآورده شود و یا اینکه اصلاً سرویس داده نشود.

این مسئله را میتوان به صورت زیر فرمول بندی کرد:

$$M = \text{تعداد مشتریان}$$

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

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

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

$$m_i = \text{میزان تقاضای مشتری } i \text{ برای کالا که به صورت یک مقدار تصادفی میباشد.}$$

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

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

از ظرفیت مجاز فراتر رود)

حال با نمادگذاری‌های ذکر شده میتوان 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.       $x_i \in \{0,1\}$

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 جریمه‌ای 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 آزادس تاکسی را که میتواند به تعدادی از خانواده‌ها در بازهای زمانی خاصی متوجه شود بدون آنکه اطلاع دقیقی از میزان مسافت هر خانواده داشته باشد. الگوریتم‌های مختلفی برای مسائلی شبیه به این مسئله ارائه شده است اما مایبیویل کن [۱۳] در تحقیقات خود منحصراً 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 [۱۴] 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 زنگیکی یوشیتومی [۱۴] و همکارانش [۱۴] اشاره کرد.

در این مقاله یک الگوریتم تکرار شونده مبتنی بر آtomاتاهای یادگیر برای حل مسئله حمل و نقل و فروش کالا پیشنهاد میشود. در این الگوریتم، شبکه‌ای از آtomاتاهای یادگیر که تشکیل یک حلقه را میدهند برای حل مسئله همکاری می‌نمایند. در این الگوریتم به تعداد مشتریها، آtomاتای یادگیر در نظر گرفته شده است که هر آtomاتا متناظر با یک مشتری خواهد بود. هر آtomاتا انتخاب یا عدم انتخاب مشتری متناظر را مشخص میکند. در این شبکه در هر مرحله یک آtomاتای 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 آtomاتاهای یادگیر بروز میشود و در خاتمه الگوریتم، بهترین راه حل بدست آمد که 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 آtomاتاهای 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- آtomاتاهای یادگیر

آtomاتای 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 آtomاتای یادگیر را ارزیابی کرده و نتیجه ارزیابی خود را توسط یک سیگنال تقویتی به آtomاتای یادگیر اطلاع می‌دهد. آtomاتای 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 آtomاتای یادگیر و محیط را نشان می‌دهد.



شکل ۱: ارتباط آtomاتای 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 ورودیها،  $\beta = \{\beta_1, \beta_2, \dots, \beta_r\}$  مجموعه خروجیها و  $c = \{c_1, c_2, \dots, c_r\}$  مجموعه احتمالات جریمه می‌باشد. هرگاه  $\beta$  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  $\beta$  دارای تعداد متناهی عضو می‌باشد و در محیط از نوع  $S$ ، تعداد اعضا مجموعه  $\beta$  نامتناهی است.  $c_i$  نشان دهنده احتمال نامطلوب

بودن سیگنال تقویتی محیط در پاسخ به اقدام  $\alpha$  می‌باشد. در یک محیط ایستا<sup>۱۵</sup> مقادیر  $c_i$  ها ثابت هستند، حال آنکه در یک محیط غیر ایستا<sup>۱۶</sup> 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  $\beta$ ، 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 می‌گردد. در این مقاله از اتوماتای یادگیر با ساختار متغیر استفاده شده است که در ادامه معرفی می‌شود. اتوماتای یادگیر با ساختار متغیر توسط چهارتایی  $\{\alpha, \beta, p, T\}$  نشان داده می‌شود که در آن  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  مجموعه اقدام‌های اتوماتای یادگیر،  $p = \{p_1, p_2, \dots, p_r\}$  بردار احتمال انتخاب هر یک از اقدام‌ها و  $T[\alpha(n), \beta(n), p(n)]$ .<sup>۱۷</sup> 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  $a_i$  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  $L_{R-P}$ <sup>۱۸</sup>، اگر  $b$  از  $a$  خیلی کوچکتر باشد، الگوریتم  $L_{R-EP}$ <sup>۱۹</sup> و اگر  $b$  صفر باشد، الگوریتم  $L_{R-I}$ <sup>۲۰</sup> نام دارد [۱۵].

### ۳ حل مسئله حمل و نقل و فروش کالا توسط آتوماتاهای 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  $DLA-TPSD$ <sup>۲۱</sup> 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 استفاده می‌کنند. بردار احتمال و عملهای اتوماتاهای یادگیر به ترتیب  $(p_1, p_2, \dots, p_m) = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$  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 اتوماتاهای می‌باشد. خروجی (پاسخ) محیط به اعمال، توسط  $\beta_i$  مشخص می‌شود. یک پاسخ دودویی می‌باشد، در چنین محیطی  $\beta(n) = 1$  بعنوان پاسخ نامطلوب یا شکست و  $\beta(n) = 0$  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 ازتابع برآش  $f(i) = r_i \mu_i$  به منظور سنجش ارزش مشتری استفاده شده است و مقدار محاسبه شده، به مقدار برآش کل انتخابهای موجود، که آن را با  $C(K)$  نشان میدهیم جمع بسته میشود.

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

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:

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

مرحله ۵: بعد از اتمام ظرفیت، مقدار  $C(K)$  با میانگین برآش انتخابهای قبلی اتوماتاهای در دفعات قبل که پاداش دریافت کردهاند، مقایسه میشود. میانگین برآش انتخابهای قبلی را با  $AVC(K_n)$ <sup>۲۲</sup> نشان می دهیم. اگر مقدار  $C(K)$  کوچکتر از میزان  $AVC(K_n)$  باشد به عمل انتخاب شده توسط اتوماتاهای یادگیر جرمیه 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 برآش 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 الگوریم خاتمه می بارد و در غیر اینصورت مرحله ۲ تا ۶ تکرار می گردد. شبیه 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)**

**Compute ( $AVC(K_n)$ )**

**Insert selected costumers and fitness of them into database**

**Until (current repeat  $\geq$  iteration number)**

**End. DLA-TPSD**

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

#### ۴- ارزیابی

الگوریتم پیشنهادی DLA-TPSD، بروی مجموعه مشتریهایی با مشخصات متفاوت از لحاظ ارزش، میزان تقاضا و میزان ظرفیت، آرمايش گردیده است و با الگوریتم ساخته‌قید به روش اکتشافی که توسط ماینوبیل کن، روی همین مسئله ارائه گردیده است [۱۳]، و نیز الگوریتم ژنتیک که توسط یوشیتومی <sup>۳۳</sup> و هیروکو <sup>۳۴</sup> 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  $\sigma^2$  واریانس متغیر تصادفی میزان تقاضای مشتری  $\lambda$  باشد،  $\max \sigma^2$  ماکریم واریانس و  $\min \sigma^2$  مینیمم واریانس توزیع‌های احتمالی میزان تقاضای مشتریهای مجموعه می‌باشد. اگر  $\mu_i$  میانگین توزیع احتمال برای مشتری  $\lambda$  باشد، در این صورت  $\max \mu$  ماکریم میانگین و  $\min \mu$  مینیمم میانگین توزیعهای احتمالی میزان تقاضای مشتریهای مجموعه می‌باشد. 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(1) مشخصات مجموعه داده‌های مورد استفاده در آزمایشات

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

جدول (2) تابع توزیع احتمال میزان اجرت مشتریها 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 مشتریها

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

درصد اجراهای همگرا شده که با  $^{35}PC$  نشان میدهیم

میانگین تعداد تکرار برای اجراهای همگرا شده که با  $^{36}AVI$  نشان میدهیم.

میانگین تعداد نمونه‌گیریهای انجام شده از کل مشتریها که با Av-Sam No نشان میدهیم.

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

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

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

تیره (-) از هم جدا شده تشكیل 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، مجموعه ۲، مجموعه ۳ = ۴ و مجموعه ۴ = ۴)

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

جدول(3) میزان پارامتر پاداش و جریمه

شناسه	(a) میزان پارامتر پادash	(b) میزان پارامتر جریمه
$L_{Rep}$	A	0.08
	B	0.08
	C	0.06
	D	0.06
$L_{Ri}$	E	0.08

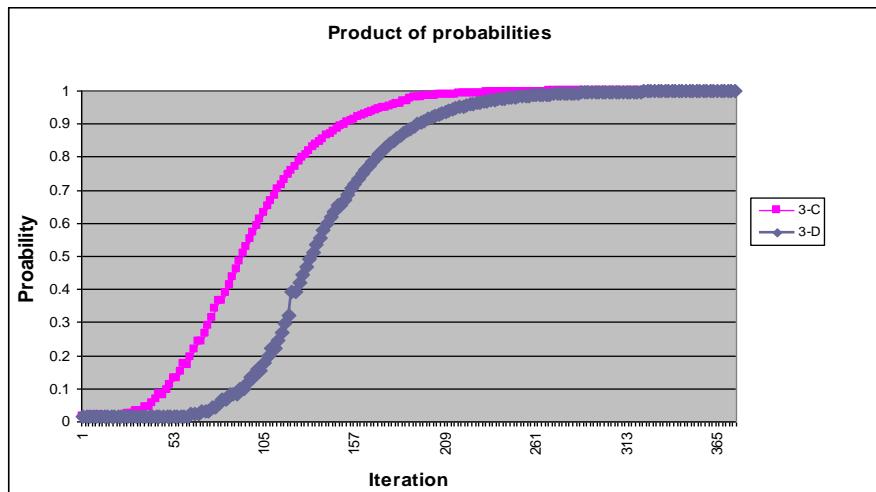
در ضمن شرط همگرایی در همه این آزمایشها، بزرگتر بودن حاصلضرب احتمالات انتخاب مشتریهای انتخاب بهینه، از (۰.۹) میباشد.

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

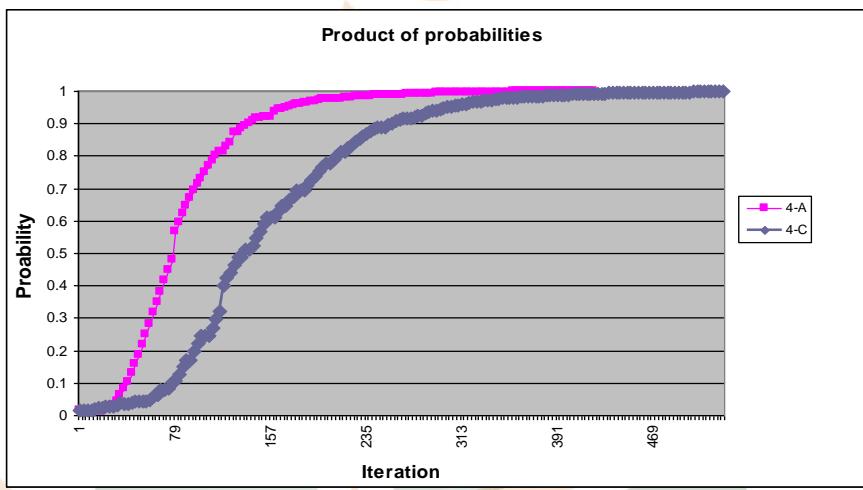
**جدول (4) جزئیات نتایج حاصل از اجرای الگوریتم پیشنهاد شده در آزمایشهای ذکر شده**

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 دادهای 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 گسترش 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 دائمًا با پاداش و جریمه روبرو میشود و بدین سبب است که گونه‌های A تا 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 دادهای شماره ۲، گونه A، جهت دادن پاداش و جریمه بهتر از بقیه گونه‌ها 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، می‌توان مشاهده کرد که الگوریتم پیشنهاد شده بروی مجموعه دادهای 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 دادهای 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 مؤثر است. در ادامه حاصلضرب احتمالات انتخاب بهینه از بین مشتری‌ها در طول اجرای الگوریتمهای پیشنهاد شده، مورد بررسی قرار گرفته شده است. مقدار حاصلضرب احتمالات میباشد که تکرارهای بیشتر و متوالی، به سمت یک، میل کند. که به عنوان نمونه نمودارهای مربوط به آزمایشات ۴-C، ۳-D، ۳-C و ۴-A نشان داده شده است.



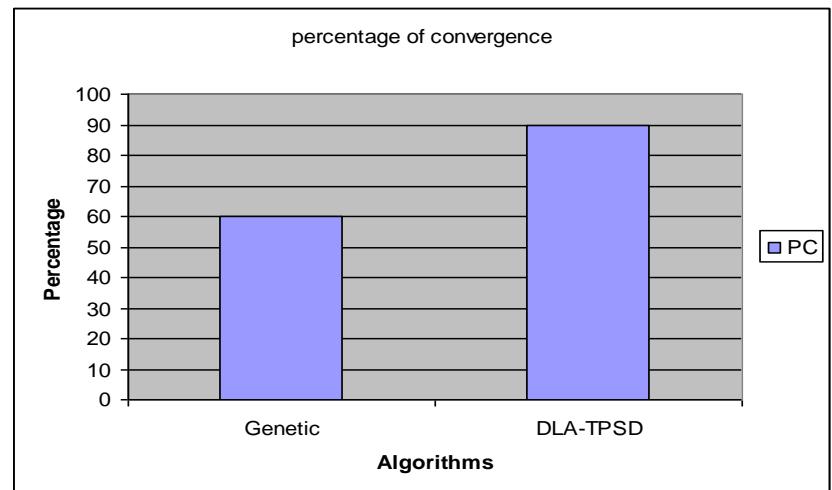
شکل ۳: حاصلضرب احتمالات انتخاب بهینه از بین مشتریها در الگوریتم پیشنهادی، در آزمایش ۳-С و ۳-Д



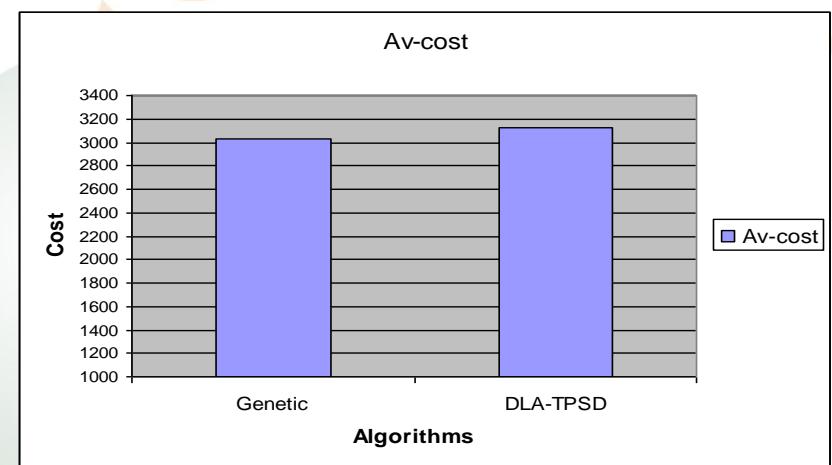
شکل ۴: حاصلضرب احتمالات انتخاب بهینه از بین مشتریها در الگوریتم پیشنهادی، در آزمایش ۴-А и ۴-С

آزمایشات نشان داده است که الگوریتم ذکر شده در مجموعه دادهای 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 تقاضاً و اجرت مشتریان و حتی نوع توزیع از لحاظ پیوسته یا گسسته بودن، نتایج مطلوبی در کل داشته است.

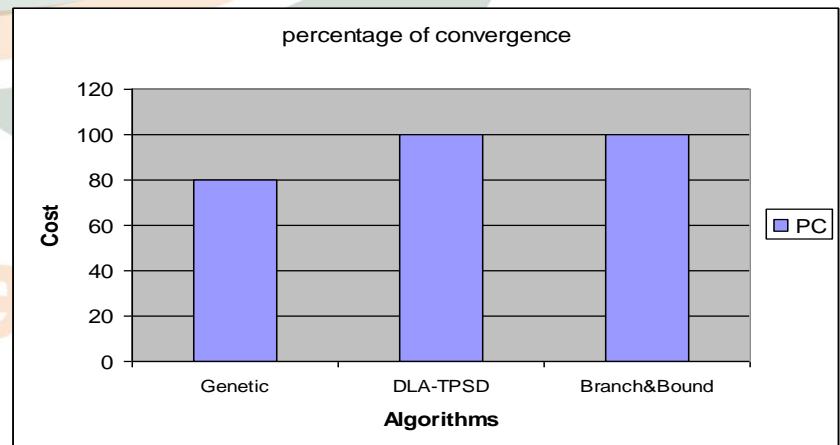
حال جهت بررسی عملکرد الگوریتم پیشنهادی در مقایسه با الگوریتم زنتیک که توسط یوشیتومی و هیروکو ارائه گردیده است [۱۴]، و نیز الگوریتم شاخه‌قید به روش اکتشافی که توسط ماینولیل کن، روی همین مسئله ارائه گردیده است [۱۳]، آزمایشاتی طراحی شده است. در این آزمایشها معیار میانگین ارزش نمونه‌های انتخاب شده توسط الگوریتم و نیز درصد اجرهایی که به پاسخ بهینه خواهند رسید در اجرهای متوالی مورد بررسی قرار گرفته است. در این آزمایشات، در الگوریتم زنتیک سایز 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 زنتیک در هر بار اجرا، یک جواب با بیشترین تکرار در کل نسلها را، به عنوان تنها جواب بهینه معرفی میکند، این الگوریتم بروی دادهای یاد شده، ۱۰ مرتبه اجرا میشود و نتایج حاصل با نتایج الگوریتمهای پیشنهاد شده که آنها نیز بروی دادهای مشابه ۱۰ مرتبه اجرا شده‌اند، مقایسه شده است. نتایج در شکلها ۵، ۶، ۷ و ۸ قابل مشاهده است. در این شکلها، نتایج آزمایشات بروی مجموعه ۱ و ۴ نشان داده شده است. نتایج آزمایشات بروی مجموعه ۲ و ۳ نیز تقریباً مشابه با این نتایج می‌باشد.



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

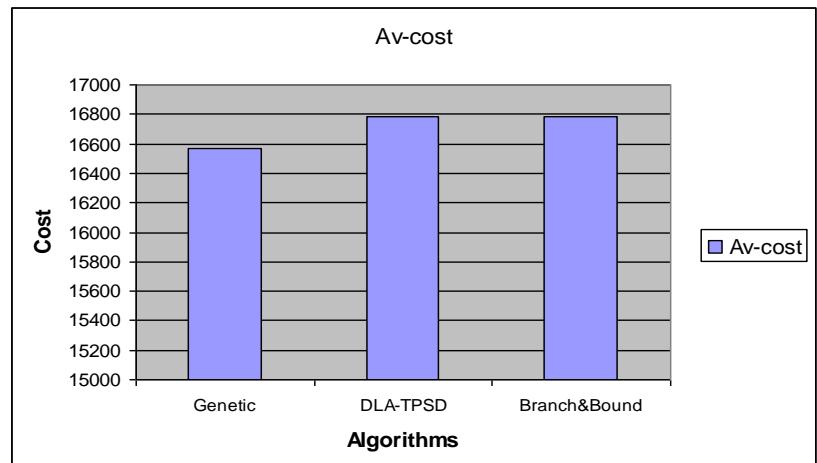


شکل 6: مقایسه الگوریتم پیشنهاد شده با الگوریتم ژنتیک بر اساس میانگین ارزش مشتریهای انتخاب شده در مجموعه 1



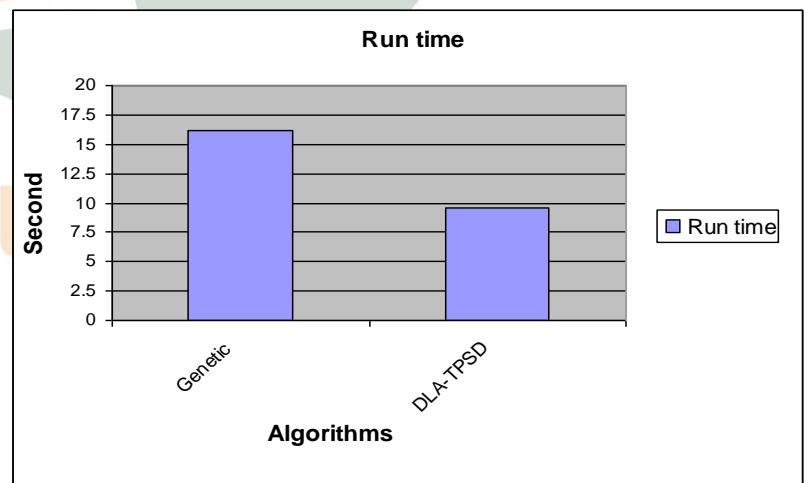
شکل 7: مقایسه الگوریتمها بر اساس درصد تکرارهایی که به جواب بهینه رسیدهاند در مجموعه 4

وی ایران



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

همانطور که مشاهده میشود از لحاظ ارزش نمونهای 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 نمونهای 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 رعف 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 4 از مشتریها، قابل مشاهده است. همان طور که مشاهده می شود زمان اجرای الگوریتم پیشنهاد شده به مرتب از زمان اجرای الگوریتم ژنتیک کمتر میباشد. همانند همین آزمایش بروی مجموعهای 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

## ۵-نتیجه گیری

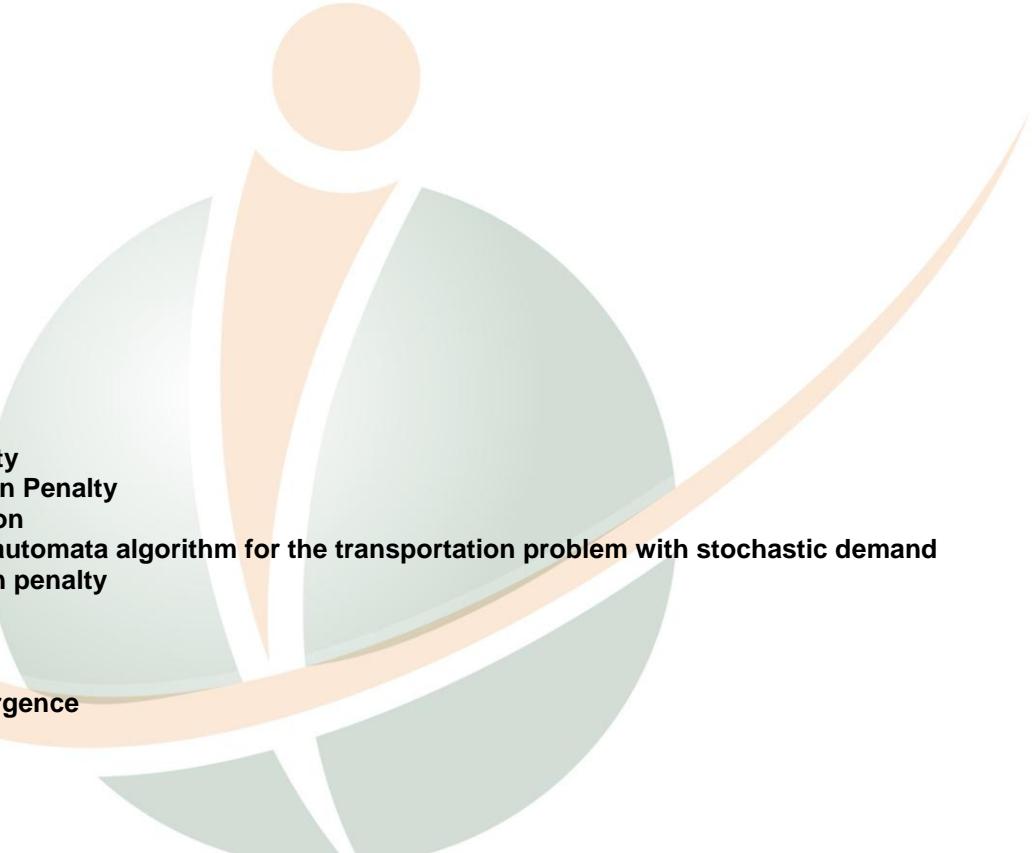
در این مقاله یک الگوریتم جدید مبتنی بر اتوماتاهای یادگیر برای حل مسئله حمل و نقل و فروش کالا پیشنهاد گردید. در این الگوریتم شبکه‌ای از آتوماتاهای یادگیر برای حل مسئله همکاری می‌نمایند. در این شبکه در هر مرحله یک آتوماتی یادگیر فعال شده و سپس یکی از اعمال خود را انتخاب شده در محیط اجرا و با توجه به نتیجه عمل، به آنها پاداش و یا جریمه داده می‌شود. الگوریتم پیشنهاد شده سعی می‌کند که با حداقل تعداد نمونه‌گیری از آیتمها 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. Cervi 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.

# کنفرانس داده کاوی ایران

- 
- ۱ Vehicle Routing Problem with Stochastic Demand
  - ۲ Tillman
  - ۳ Stewart
  - ۴ Golden
  - ۵ Bertsimas
  - ۶ Simchi-Levi
  - ۷ Re-optimization
  - ۸ Secomandi
  - ۹ Iterated Local Search
  - ۱۰ Cyclic Heuristic
  - ۱۱ Yang
  - ۱۲ Mainville Cohn
  - ۱۳ Iterative
  - ۱۴ Yoshitomi
  - ۱۵ Stationary
  - ۱۶ Non-Stationary
  - ۱۷ Linear Reward-Penalty
  - ۱۸ Linear Reward epsilon Penalty
  - ۱۹ Linear Reward Inaction
  - ۲۰ Distributed learning automata algorithm for the transportation problem with stochastic demand
  - ۲۱ Linear reward epsilon penalty
  - ۲۲ Average cost
  - ۲۳ Yoshitomi
  - ۲۴ Hiroko
  - ۲۵ Percentage of convergence
  - ۲۶ Average of iteration



کنفرانس داده کاوی ایران