



وزارت علوم و تحقیقات و فناوری

دانشگاه علوم و فنون مازندران

پایان نامه

مقطع کارشناسی ارشد

مهندسی سیستم‌های اقتصادی و اجتماعی

موضوع:

طراحی کاتالوگ های مشتری گرا

با استفاده از یک الگوریتم خود تطبیق ژنتیک

استاد راهنما:

دکتر ایرج مهدوی

استاد مشاور:

دکتر نیکبخش جوادیان

نگارش:

مهیار موحد نژاد

تابستان ۱۳۸۷

کتابخانه مرکزی مازندران
تاسیس ۱۳۸۸

۱۳۸۸ / ۲ / ۵۱

۱۱۱۸۰۹

Abstract

Considering the importance of introducing products, in this research, the problem of designing costumer-oriented catalogs has been investigated. After generalizing the mathematical model of customer-oriented catalog problem, we developed a self-adaptive genetic algorithm to solve the problem. It includes clever features to avoid getting trapped in a local optimum. The results of an extensive computational study using real and synthetic data sets show the performance of the proposed algorithm in comparison with Greedy Out algorithm (GO).

چکیده

با توجه به اهمیت روز افزون معرفی مطلوب محصولات به مشتریان، در این تحقیق مسئله "کاتولوگ های مشتری گرا" معرفی و بررسی شده است. پس از گسترش مدل ریاضی طراحی کاتولوگ های مشتری گرا، یک الگوریتم خود تطبیق ژنتیک برای حل مسئله ارائه شده است که شامل ویژگی های لازم برای اجتناب از باقی ماندن جواب در بهینه های محلی می باشد. نتایج محاسبات بر داده های حقیقی و مجازی، کارایی رویکرد معرفی شده را در مقایسه با الگوریتم (GO (Greedy out نشان می دهد.

فهرست :

۳	چکیده
۷	فصل اول
۷	تعریف مسأله
۸	۱-۱- مقدمه
۱۲	۱-۲- بیان مسأله
۱۲	۱-۳- فرضیات مسأله :
۱۴	۱-۴- اهداف تحقیق:
۱۴	۱-۵- خلاصه فصول بعدی این تحقیق
۱۴	۱-۵-۱- خلاصه فصل دوم
۱۴	۱-۵-۲- خلاصه فصل سوم
۱۵	۱-۵-۳- خلاصه فصل چهارم
۱۵	۱-۵-۴- خلاصه فصل پنجم
۱۵	۱-۶- جمع بندی
۱۶	فصل دوم
۱۶	ادبیات موضوع
۱۷	۱-۲- مقدمه
۱۷	۲-۲- بررسی ادبیات موضوعه ی تحقیق
۲۸	فصل سوم
۲۸	زمینه های علمی تحقیق
۲۹	۱-۳- مقدمه
۲۹	۲-۳- مسأله بخشبندی کاتالوگ
۳۲	۳-۳- مسأله خوشه بندی عمومی
۳۲	۳-۳-۱- مسأله بخش بندی برای (D,F) :
۳۳	۳-۳-۲- مسأله بخش بندی برای (D,F) زمانی که k متغیر است

۳۳ پیچیدگی مسائل بخشبندی
۳۴ ۵-۳ مسائل پست الکترونیکی
۲۴ ۱-۵-۳-تعریف مساله تک پست الکترونیکی [۱۵]:
۲۵ ۱-۱-۵-۳ مساله ی تک پست الکترونیکی با یک کاتالوگ:
۲۶ ۲-۱-۵-۳ مساله ی تک پستی الکترونیکی با ۲ کاتالوگ:
۲۷ ۲-۵-۳ مساله ی چندین پست الکترونیکی
۲۵ ۶-۳-الگوریتم ژنتیک
۳۵ ۱-۶-۳ مقدمه
۳۶ ۲-۶-۳ الگوریتم ژنتیک
۳۷ ۱-۲-۶-۳ مفاهیم اولیه
۳۹ ۲-۲-۶-۳ عوامل اساسی مؤثر بر الگوریتم های ژنتیک
۳۹ ۱-۲-۲-۶-۳ نحوه نمایش کروموزوم
۴۰ ۲-۲-۲-۶-۳ جمعیت اولیه
۴۰ ۳-۲-۲-۶-۳ استراتژی انتخاب والدین
۴۳ ۴-۲-۲-۶-۳ تابع برازندگی
۴۳ ۵-۲-۲-۶-۳ عملگرهای ژنتیکی
۴۶ ۶-۲-۲-۶-۳ معیار توقف الگوریتم
۴۷ ۷-۲-۲-۶-۳ همگرایی در GA
۴۸ ۳-۶-۳ الگوریتم ژنتیک استاندارد
۴۹ ۷-۳ الگوریتم های خود تطبیق ژنتیک در داده کاوی
۵۱ ۸-۳ خلاصه و نتیجه گیری
۵۲ فصل چهارم
۵۲ یافته های علمی تحقیق
۵۳ ۱-۴ مقدمه
۵۴ ۲-۴ فرمول سازی مسئله
۵۹ ۳-۴ الگوریتم ژنتیک خود تطبیق:
۶۰ ۱-۳-۴ معرفی کروموزوم
۶۱ ۲-۳-۴ جمعیت اولیه:
۶۲ ۳-۳-۴ عملگر جهش

۶۵ عملگر تقاطع: ۴-۳-۴
۶۶ عملگر تولید مجدد: ۵-۳-۴
۶۶ استراتژی انتخاب والدین: ۶-۳-۴
۶۷ ۷ بررسی سناریو ۳-۳-۴
۶۸ یافته های علمی تحقیق ۴-۴
۶۸ ۱۱ آزمون های محاسباتی ۴-۴
۷۱ ۲-۴-۴ آزمون های محاسباتی مقایسه ای:
۷۱ ۱-۲-۴-۴ داده های حقیقی
۷۲ ۲-۲-۴-۴ داده های مصنوعی:
۷۹ ۵-۴ جمع بندی
۸۰ فصل پنجم
۸۰ نتیجه گیری و پیشنهاد
۸۱ ۱-۵ خلاصه و نتیجه گیری کلی
 Error! Bookmark not defined. ۲-۵ پیشنهاد برای تحقیقات آتی
۸۳ منابع و مراجع
۸۸ پیوست: کد برنامه

فصل اول

تعريف مسأله

بسیاری از شرکت ها در فضای رقابتی بازاریابی موجود ، کاتالوگ های معرفی کالاها و خدمات را به عنوان ابزار کارایی به منظور رسیدن به اهداف خود نظیر جذب مشتری^۱ و افزایش رضایت^۲ و وفاداری^۳ مشتریان استفاده می کنند. این شرکتها ، علاقه بسیار به طراحی بهینه کاتالوگ های معرفی محصولات دارند. یک شرکت با تولید کاتالوگ های متفاوت می تواند محصولات را بر اساس نیاز گروه های مختلفی از مشتریان ارائه کند .

با توجه به افزایش تعداد کالا های ارائه شده توسط خرده فروشان سنتی و الکترونیکی ، طراحی کاتالوگی که معرفی کننده تمامی محصولات باشد غیر عملی و از نظر اقتصادی به صرفه نمی باشد . در بسیاری از موارد ، مشتریان تنها به برخی از محصولات شرکت علاقه مندند . مشتریان ، در صورتی که کاتالوگی شامل دسته ای از محصولات مناسب با نیاز خود دریافت کنند ، رفتار های خرید سودآورتری از خود بروز می دهند .

کاتالوگ ها بطور گسترده ای توسط سایت های فروش برای گرفتن سفارش های آنلاین یا تلفن استفاده می شود. به عنوان مثال ، Americanmusical.com و Landsend.com کاتالوگ های خود را با سایت های فروش آنلاین خود هماهنگ کرده اند ، بطوریکه مشتریان تنها با وارد کردن شماره محصول در کاتالوگ ، می توانند سفارش خرید بدهند .

یک بررسی که توسط Catalog Age Magazine [1] صورت گرفته ، حاکی از آنست که ۷۰٪ از شرکتهایی که فروش حداقلی ۱۰ میلیون دلار در سایت های فروش خود داشته اند از سیستمهای مدیریتی پیشرفته و پرخرج برای ارائه کاتالوگ های خود استفاده کرده اند این شرکتها

¹ Customer attraction

² Satisfaction

³ Loyalty

از کاتالوگ ها (چه از نوع چاپی ، چه CD و چه آنلاین) به عنوان یک ابزار کار جهت جذب مشتری استفاده کرده اند .

هزینه های بازار یابی ، به عنوان یک چالش عمده در مقابل شرکت های استفاده کننده از کاتالوگ ها قرار گرفته است بررسی دیگری که توسط Catalog Age Magazine [2] صورت گرفته ، حاکی از آنست که شرکت های بررسی شونده بطور متوسط ۲۶/۱٪ از سود خود را صرف بازاریابی می کند . جالب آنست که چاپ و پست معرفی کننده های محصولات نزدیک به نصف این هزینه ها را شامل می شود .

تعداد کاتالوگ هایی که شرکتها به مشتریان بالقوه و بالفعل خود می فرستند ، با سرعت در حال افزایش است به عنوان مثال ، Victoria's Secret حدود ۳۹۵ میلیون کاتالوگ در هر سال پست می کند . بررسی اشاره شده در بالا ، نشان می دهد که از هر ۵ شرکت تقریباً ۴ شرکت (۷۹٪) ، " نیاز به کاهش هزینه های بازاریابی، بدون کاهش خدمات ارائه کالا " را یکی از ۳ چالش عمده رو در روی مدیریت عنوان می کنند .

با توجه به افزایش هزینه های بازاریابی و رقابت در اقتصاد جهانی ، شرکت ها در جستجوی رویکردهای جدید در تولید و ارائه کاراتر از کاتالوگ ها می باشند . یک روش که هم اکنون توسط شرکت های متمول استفاده می شود ، اتوماتیک کردن تولید کاتالوگ ها با استفاده از بسته های نرم افزاری مدرن و بسیار پرهزینه (میلیون ها دلار) که قیمت آنها بسته به ظرفیت نرم افزار ها در پذیرش پیچیدگی و حجم مساله ، تغییر می کند . [3]

تکنولوژی اینترنت، شرکت ها را قادر به پیش بینی تعاملات مشتریان و رفتار خرید آنها می کند. این پدیده منجر به سفارشی کردن کالا، خدمات، ارتباطات و قیمت ها می گردد. بررسی تعاملات

مشتریان در مدیریت الکترونیکی ارتباط مشتریان¹ (e-CRM) می تواند استفاده از داده کاوی²، رویکردهای بهینه سازی³ و یا رویکردهای ترکیبی⁴ صورت پذیرد [4].

استفاده از اینترنت برای ارائه کالاها و خدمات به مشتریان شرکت ها را ملزم به فهم درستی از نیازهای مشتریان می کند. زمانی که مشتریان از وب سایت یک شرکت دیدن می کنند، پیش بینی محصولات مورد توجه و ارائه ی محصولات مرتبط با آنها برای شرکت یک امر حیاتی است. معرفی اطلاعات محصولات، به گونه ای که منجر به ترویج مشتریان به خرید از طریق وب سایت ها گردد، برای شرکت ها اهمیت بسیاری یافته است.

یک نکته مهم در بررسی رفتارهای خرید مشتریان از طریق وب سایت ها، خوشه بندی⁵ و طبقه بندی⁶ آنها بر پایه ی رفتارهای خرید قبلی مشتریان یا علائق آینده آنها (که از طریق آنالیز محیطی بازار به دست می آید) است. پایگاه داده ی⁷ علاقه مندی های مشتریان، شامل مجموعه ای از محصولات مورد علاقه ی مشتریان می باشد. محصولات مورد علاقه ی یک مشتری، با استفاده از جمع آوری تعاملات قبلی مشتری یا با استفاده از شناسایی اولویت های مشتریان برای مجموعه ای از کالاها به دست می آید. [5] چنین پایگاه داده ای معمولاً از طریق جمع آوری تعاملات قبلی مشتریان صورت می گیرد. سازمان های موفق، به چنین پایگاه داده ای به عنوان زیر بنای بازاریابی خود می نگرند. آنها علاقه مند به ایجاد کردن یک فرآیند بازاریابی با استفاده از مدیریت پایگاه های داده ی خود هستند، به طوریکه سازمان ها را قادر به اعمال استراتژی ها و برنامه های بازاریابی⁸ مشتری پسند گرداند.

¹ Electronic customer relationship management

² Data mining

³ Optimization method

⁴ Combined approaches

⁵ Clustering

⁶ Classification

⁷ Data base

⁸ Marketing

در سال های اخیر، افزایش چشمگیری در گردآوری و ذخیره سازی اطلاعات مرتبط با مشتریان از طرف بنگاه های تجاری صورت گرفته است. مقادیر حجیم از تعاملات مشتریان، منجر به علاقه ی روز افزونی به تحقیقات برای استخراج اطلاعات مفید برای تصمیم گیری از این پایگاه ها شده است. با اینکه اطلاعات محصولات در برخی از مواقع به طور کارایی می تواند مختص به تک تک مشتریان گردد، این چنین رویکردی در همه ی موقعیت ها عملی نیست. مانند کاتالوک های کوچکی که توسط پست الکترونیکی به گروهی از مشتریان در سال ارسال می گردد. همین طور، وب سایت های فروش شرکت ها یا نوارهای تبلیغاتی¹ بر وب سایت های خارج از تحت تملک شرکت می باشد. شرکت ها باید در مورد مجموعه ی خدماتی که توسط این گونه موارد با هدف افزایش اشتیاق خرید مشتریان رضایت وفاداری آنها، تصمیم گیری کند.

در نتیجه این تحقیق، با استفاده از اطلاعات استخراج شده در مورد رفتارهای خرید مشتریان و تطبیق آنها با این نگرش که اغلب استفاده ی سودآور از این اطلاعات در معرفی محصولات از طریق کاتالوک ها صورت می گیرد، به معرفی کاتالوک ها به گروه های مختلف مشتریان می پردازد به این صورت که کاتالوک ها مختص به تک تک مشتریان نمی باشد و برای مشتریان با رفتارهای خرید شبیه به هم ارائه می شود.

سپس، الگوریتمی با هدف ارائه ی کاتالوک های ترویج کننده² در محیط (e-CRM) ارائه می شود. این الگوریتم، مجموعه ای از محصولات موجود را به مشتریان از طریق وب سایت با هدف تشویق مشتری به خرید ارائه می کند. هر کاتالوک زیر مجموعه ای از مجموعه ی کل محصولات می باشد و کاتالوک های مختلف ممکن است شامل برخی از محصولات کاتالوک های دیگر باشند بدین گونه که، کاتالوک ها ممکن است دارای اشتراک تهی نباشند. با این هدف ما به مسئله ی

¹ Banner ads

² Promotional catalog

یافتن ارتباطات، میان پایگاه های داده ی عظیم برای مشخص کردن استراتژی های فروش می پردازیم.

1-2- بیان مسأله

طراحی کاتالوگ های مشتری گرا در e-CRM (e-COC)¹ شامل طراحی n کاتالوگ (صفحه) می باشد، به طوریکه کاتالوگ هایی با ظرفیت ارائه ی s محصول با هدف بیشینه کردن تعداد مشتریان پوشش داده شده تولید کند. با توجه به اینکه همواره الویت به صفحه ی اول داده می شود و صفحات بعدی با یک الویت نزولی از اهمیت در نظر گرفته می شوند. بنابراین هدف ما، بشنیه کردن تعداد مشتریان پوشش² داده شده با توجه به ترتیب صفحات کاتالوگ یا ترتیب صفحات وب سایت می باشد. یک مشتری زمانی پوشش داده می شود که، او به حداقل تعداد مشخصی از کالاهایی که در یک صفحه ارائه می شود علاقه مند³ بوده باشد.

ما مسئله ی طراحی کاتالوگ های مشتری گرا در محیط e-CRM را (e-COC) با توجه به افزایش چشمگیر تقاضا در این محیط ارائه می کنیم. این مسئله یک تعمیم از مسئله ی بخش بندی کاتالوگ مشتری گرا⁴ می باشد که ابتدا توسط Ester et al [6] که به خودی خود یک تعمیم از مسئله ی کلاسیک و NP-complete بخش بندی کاتالوگ می باشد، ارائه شده است.

مسئله ی e-COC، مشتریان را به n بخش خوشه بندی می کند. هر خوشه⁵ می بایست شامل s محصول و افراز مشتریان مرتبط باشد. هر خوشه، زمانی یک مشتری را به خود تخصیص می دهد که حداقل به t محصول تحت شرط ذیل علاقه مند باشد:

¹ Electronic Customer-Oriented Catalogs

² Cover

³ Interest

⁴ Customer-oriented catalog segmentation problem

⁵ Cluster

*هر مشتری پوشش داده شده فقط و فقط به یک خوشه تخصیص داده می شود. یعنی اینکه، اشتراک میان خوشه های مشتریان تهی می باشد.

هدف پیشینه ساختن تعداد مشتریان تخصیص داده شده با توجه به ترتیب خوشه هاست.

در مسئله ی e-COC از آنجایی که صفحه ی اول دارای اولویت اول است مشتریان می بایست تا آنجا که ممکن است در صفحه ی اول پوشش داده شوند. در صفحه ی دوم همین سیاست برای باقی مانده ی مشتریان در نظر گرفته می شود. این فرآیند برای باقی مانده ی صفحات اعمال می شود. بدین علت که مشتریان به خود زحمت دیدن و بررسی تعداد زیادی از صفحات را نمی دهند، وزنی (W) به هر صفحه تخصیص داده و محدودیتی برای تعداد کل صفحات در نظر گرفته می شود.

با توجه به عدم کارایی رویکردهای موجود در حل مسائل اینچینی، محققان ناچار به ساده سازی مدل ها و اضافه سازی محدودیت ها می گشتند که این امر منجر به دور گشتن مدل‌های موجود در ادبیات موضوعه از نیاز حقیقی بازار می گشته. این تحقیق با ارائه رویکردی جدید قادر به معرفی نزدیک ترین مدل به نیاز شرکتها شده است.

۱-۳ فرضیات مساله :

- پایگاه داده مشتریان که شامل کالاهای مورد علاقه هر مشتری است مشخص می باشد. این پایگاه داده یا از طریق ضبط سابقه خرید حقیقی هر مشتری ویا از طریق روشهای آماری تولید می گردد.
- تعداد کاتالوگها با در نظر گرفتن رابطه سود و هزینه از طرف شرکت مشخص میگردد.
- تعداد محصولات هر کاتالوگ مشخص می باشد.
- هر مشتری تنها یک بار پوشش داده می شود.
- تعداد مشتریان پوشش داده شده در صفحات اول از اهمیت بیشتری برخوردار می باشند.

۴-۱ اهداف تحقیق:

کاربردهای فراوان مساله در جلوگیری از اتلاف منابع شرکتها و مشتری گرا کردن ارائه و فروش محصولات که منجر به بالا بردن چشمگیر رضایت و وفاداری مشتریان می گردد. این تحقیق پس از ارائه مدل تصمیم برای طراحی سایت های فروش، رویکردی تطبیق پذیر برای حل مدل تصمیم در جهت رسیدن به اهداف زیر ارائه می کند:

۱. مشخص کردن محتویات صفحات سایت^۱

۲. خوشه بندی کردن مشتریان^۲

۵-۱ خلاصه فصول بعدی این تحقیق

۵-۱-۱ خلاصه فصل دوم

در این فصل ادبیات موضوعه ی مساله بخش بندی کاتالوگ به طور مفصل بررسی می شود. با توجه به اینکه در الگوریتم پیشنهادی برای حل مساله ی طراحی کاتالوگ های مشتری گرا در e-CRM (e-COC)، از یک الگوریتم خود تطبیق ژنتیک^۳ استفاده شده است ادبیات موضوعه ی الگوریتم های خود تطبیق ژنتیک نیز بررسی گشته است.

۵-۱-۲ خلاصه فصل سوم

در این فصل مسائل بخش بندی کاتالوگ، مساله ی خوشه بندی^۴، مساله ی تک پست الکترونیکی^۵، مساله ی چندین پست الکترونیکی^۶ به تفصیل بررسی می شود. سپس روش فرا ابتکاری^۷ الگوریتم های ژنتیک که از روش های معروف و پرکاربرد در ادبیات موضوعی مسائل

¹ Page contents

² Customer clusters

³ Self-adaptive genetic algorithm

⁴ Clustering problem

⁵ Single-mailing problem

⁶ Multiple-mailing problem

⁷ Meta-heuristic

بهینه سازی هستند پرداخته و عوامل و پارامتر های مربوط به تفصیل مورد بحث قرار گرفت و در پایان اشاره ی کوچکی به روش الگوریتم های خود تطبیق ژنتیک می شود.

۱-۵-۳- خلاصه فصل چهارم

در این فصل پس از معرفی و ارائه مدل تصمیم مساله ی e-COC به ارائه ی الگوریتم خود تطبیق ژنتیک پیشنهادی برای حل مساله پرداخته می شود. سپس، پس از نشان دادن کارایی الگوریتم در پایگاه داده ای حقیقی^۱، به مقایسه ی روش پیشنهادی با کاراترین روش حل مرتبجا در ادبیات موضوعه ی مساله ی بخش بندی کاتالوگ(GO)^۲ پرداخته می شود. این مقایسه، شامل آزمون های مقایسه ای زیادی است که نه تنها بر روی داده های حقیقی بلکه بر روی داده های مصنوعی که با استفاده از IBM data generator [7] به وجود آمده است، اعمال شده است. سپس به آزمون حساسیت پارامترهای مساله پرداخته شده است.

۱-۵-۴- خلاصه فصل پنجم

در این فصل نتیجه گیری کلی و پیشنهادات برای تحقیقات آینده مطرح می گردد.

۱-۶- جمع بندی

در این فصل پس از ذکر اهمیت طراحی و ارائه ی کاتالوگ های کارا، و به خصوص کاتالوگ های آنلاین و صفحات آنلاین فروش، برای بنگاه های تجاری به بیان مساله ی e-COC، فرضیات مساله، اهداف تحقیق و خلاصه فصول بعدی پرداخته شده است.

¹ Real data base

² Greedy out algorithm

فصل دوم

ادبیات موضوع

۲-۱- مقدمه

شالوده ی ذهنی مساله بخش بندی کاتالوگ، داده کاوی در ساختاری اقتصادی^۱ است که توسط کلینبرگ و همکاران [9] معرفی شده است. در این ساختار، یک شرکت با تصمیمات متفاوتی در مورد مشتریان رودررو است. بسته به تصمیمی که انتخاب می شود سود شرکت متغیر است. یعنی، تصمیم انتخابی، تاثیر مستقیمی در سود کل شرکت دارد [6].

۲-۲- بررسی ادبیات موضوعه ی تحقیق

در مساله ی بخش بندی کاتالوگ، هدف پیدا کردن تصمیم بهینه برای هر خوشه از مشتریان است. البته، مسلماً، اعمال کردن یک تصمیم بهینه برای هر مشتری سود بیشتری را نصیب شرکت می کند، ولی به تحقق رساندن این نگرش عملاً به علت هزینه های بالا امکان پذیر نمی باشد.

مطلوبیت هر تصمیم (در اینجا کاتالوگ) به وسیله ی تعداد مشتریانی که حداقل به تعداد معینی از محصولات معرفی شده در همان کاتالوگ علاقه منداند، اندازه گیری می شود.

مساله کلاسیک بخش بندی کاتالوگ در کلینبرگ و همکاران [9] معرفی شده است. در این مساله، هدف در بیشینه کردن تعداد کاتالوگ های مورد علاقه مشتریان بدون توجه به مطلوبیت هر مشتری برای شرکت است.

مطالعات چاپ شده در حوزه بخش بندی کاتالوگ، به جز [3,6] بر پایه مسائل مطرح شده در [8,9] می باشد. این مطالعات، عموماً، به ارائه الگوریتم هایی بر پایه شمارش و محاسبه تمامی حالت های امکان پذیر برای یک نمونه از فضای جواب می پردازند.

¹ Microeconomic framework for data mining

همانطور که در [10] اشاره شده، "این الگوریتم ها تنها زمانی قابل اجرا هستند که طیف گسترده ای از مفروضات بر نمونه ها اعمال شود".

ژو و همکاران [11] مساله بخش بندی ۲- کاتالوگ^۱ را معرفی و مطالعه کرده اند که تنها به طراحی ۲ کاتالوگ می پردازد. آنها یک الگوریتم تخمینی^۲، بر پایه ی semi-definite programming طراحی کرده اند که با گارانتی اجرای مناسب ۱/۲ برای هر اندازه ی (r) از کاتالوگ و با گارانتی اجرای مناسب بالاتر از ۱/۲ زمانی که اندازه کاتالوگ حداقل $m/3$ باشد (m تعداد کالاها) قابل اعمال می باشد.

برای غلبه بر نا کارآمدی الگوریتم هایی که بر پایه ی نمونه برداری از فضای جواب اجرا می شوند، استینیچ و همکاران [12] مساله را زمانی که یک مشتری یک کاتالوگ را در مرحله اول و چندین کاتالوگ را در دوره ای از زمان، در مرحله دوم دریافت می کند بررسی کرده اند. آنها ۳ الگوریتم را بر پایه ی خوشه بندی k-means طراحی کرده اند و سپس به مقایسه ی رقابتی این ۳ الگوریتم پرداخته اند.

در استر و همکاران [6] و امیری [3]، مساله بخش بندی کاتالوگ مشتری گرا مطرح شده است. اثر این مساله مطلوبیت هر کاتالوگ با تعداد مشتریانی که حداقل به تعداد معینی از محصولات درون کاتالوگ علاقه مندند، اندازه گیری می شود. این مفهوم در پوشش دادن مشتریان در مساله بخش بندی کاتالوگ، شبیه مفهوم کمینه سازی پشتیبانی در کاوش قوانین ارتباطات^۳ است [13,14].

نگرش پوشش مشتریان چند هدف را دنبال می کند:

¹ Two-catalog segmentation problem

² Approximation algorithm

³ Association rule mining

i. کاتالوگ ها شخصی سازی نمی شوند، بلکه برای گروهی از مشتریان شبیه رفتار با علائق

مشابه طراحی می شود [15].

ii. کاتالوگی که دارای تعداد کمی از محصولات مورد علاقه مشتری می باشد، عملاً ناکارآمد

است.

iii. طراحی کاتالوگ هایی که منجر به جذب بیشتر مشتری و در نتیجه کاهش هزینه های

بازاریابی شوند، مطلوب است.

امیری [3] پس از بحث بر پیچیدگی مساله، دو الگوریتم برای هر ارائه داده است. الگوریتم GO کاتالوگ هایی را با استفاده از الگوریتمی طماعانه¹ بدین گونه طراحی می کند که هر کاتالوگ در ابتدا شامل تمامی محصولات است و محصولات به طوری از کاتالوگ ها حذف می شوند که تعداد مشتریان پوشش داده نشده را کمینه کند. این فرآیند تا زمانی ادامه پیدا می کند که مشتریان درون کاتالوگ به تعداد تعیین شده ی τ برسند.

دومین الگوریتم Association rule mining در حوزه داده کاوی است که کل کاتالوگ را در یک مرحله تولید می کند. محصولات با هدف بیشینه کردن روابط میان محصولی با توجه به روابط بین علائق مشتریان گروه بندی می شوند. سپس با استفاده از یک مطالعه ی محاسباتی برتری الگوریتم GO بر الگوریتم های پیشنهادی استر و همکاران [6] نشان داده می شود.

با توجه به پیچیدگی (NP-complete) مسائل بخش بندی کاتالوگ و گستردگی بسیار زیاد فضای جواب، این مسائل با روش های دقیق قابل حل نمی باشند. این تحقیق با ارائه یک الگوریتم خود تطبیق ژنتیک به حل مساله معرفی شده (e-COC) پرداخته است.

¹ Greedy

گونه ی جدیدی از الگوریتم های ژنتیک که الگوریتم های خود تطبیق ژنتیک نام گرفته اند [16,17,18]، نرخ پارامترهای GA را به طور پویا با توجه به جواب های بدست آمده در حین اجرای الگوریتم، بر اساس مقدار برازش¹ جواب ها در هر نسل منطبق می کنند. علاوه بر خود پارامترها، پارامترهای کنترل نیز می توانند بطور پویا تطبیق پیدا کنند.

استفاده سیستم های موازی در اجرای الگوریتم های ژنتیک، منجر به بوجود آمدن الگوریتم های ژنتیک موازی² گردیده است. در مدل های مهاجرتی الگوریتم های ژنتیک³ [19]، به جای یک جمعیت⁴، مجموعه ای از جمعیت ها دخیل هستند. در این مدل ها، در هر نسل تمامی جمعیت ها بطور مستقل فعال هستند. برخی قوانین، نوعی از مهاجرت داخلی را بین جمعیت ها کنترل می کنند و اجازه جابجایی تعداد کمی از جوابها را بین جمعیت ها می دهند.

لوبو و همکاران [20]، الگوریتم های تطبیقی ژنتیکی طراحی کرده اند که بدون داشتن مقدار بهینه جمعیت اجرا می شود. در این روش، گروههای جستجوی موازی با تعداد مختلف جواب اعمال می شوند تا اینکه یکی از جمعیت ها که به جواب های مناسب دست پیدا کرده است انتخاب شود. به هر حال، این روش بطور کلی تطبیقی نیست. چرا که مقدار مناسب جمعیت تنها بر اساس سعی و خطا بدست می آید. همچنین، این روش زمانی که گروه های جستجوی بزرگ موازی نیاز باشند، به علت زمان طولانی اجرا عملی نیستند. [21]

به گونه ای مشابه، برخی از کاربردهای چنین الگوریتم هایی [22] در جستجوی قوانین چند هدفه در [23] مطالعه شده اند.

یک الگوریتم تطبیقی ژنتیک که³ الگوریتم را همزمان اجرا می کند در [24] ارائه شده است. در هر دوره ی این فرآیند، مقدار برازش نخبگان جمعیت ها مقایسه، و تعداد جوابها در جمعیت بر

¹ Fitness

² Parallel genetic algorithm

³ Migration model of genetic algorithm

⁴ Population