Engineering Applications of Artificial Intelligence 56 (2016) 30–59

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Evolutionary computing applied to customer relationship management: A survey

Gutha Jaya Krishna [a,b], Vadlamani Ravi [a,*]

[a] Center of Excellence in Analytics, Institute for Development and Research in Banking Technology, Castle Hills Road #1, Masab Tank, Hyderabad 500057, India

[b] School of Computer & Information Sciences, University of Hyderabad, Hyderabad 500046, India

ARTICLE INFO

Article history:
Received 26 January 2016
Received in revised form
19 August 2016
Accepted 24 August 2016
Available online 6 September 2016

Keywords:
Customer relationship management
Data mining
Evolutionary computing
Optimization

ABSTRACT

Customer relationship management (CRM) is a customer-centric business strategy which a company employs to improve customer experience and satisfaction by customizing products and services to customers' needs. This strategy, when implemented in totality eventually increases the revenue of the company. Traditionally, data mining (DM) techniques have been applied to solve various analytical CRM tasks. In turn, optimization techniques have long been used for training some of the DM techniques. However, during the past few years, evolutionary techniques have become so powerful and versatile that they can be deployed as a substitute for some DM techniques. This trend caught the attention of the researchers working in the analytical CRM area as they too started solving the CRM tasks using evolutionary techniques alone. In this context, we present a survey of evolutionary computing techniques applied to CRM tasks. In this paper, we surveyed 78 papers that were published during 1998 and 2015, where the application of evolutionary computing (EC) techniques to analytical CRM tasks is the main focus. The survey includes papers involving evolutionary computing techniques applied to the analytical CRM tasks under single- as well as multi-objective optimization framework. The purpose of the survey is to let the reader realize the versatility and power of EC techniques in solving analytical CRM tasks in the service industry and suggesting future directions.

1. Introduction

"CRM is a strategic process of selecting customers that a firm can most profitably serve and shaping interactions between a company and these customers. The ultimate goal is to optimize the current and future value of customers for the company" (Kumar and Reinartz, 2012). Customer Relationship Management (CRM) advocates a scientific method of identifying new customers, nurturing relationships with them, retaining them by satisfying their financial needs and finally making sure that profitable and loyal customers do not attrite to competition. In Dyche (2002), a business oriented perspective to CRM is very well presented in the form of case studies of some of the well known companies. In Ngai (2005) a literature review and classification of CRM is presented from 1992 to 2002. CRM frequently includes offering new and customized products and services, often in a personalized manner to customers. CRM has primarily three components: (i) Operational CRM (OCRM), (ii) Analytical CRM (ACRM), and (iii) Collaborative CRM (CCRM). The above mentioned CRM components are briefly described with attributes and measures in Iriana and Buttle (2007). Trends, topics and under researched areas are described in Wahlberg et al. (2009). In operational CRM, the service company executes sales and services through various touch points of customers with the help of the knowledge provided by the Analytical CRM component. In a sense, this component subsuming call centers, which are the face of the firm as far as the sales and services are concerned. On the other hand, in analytical CRM, data mining techniques are utilized, subsuming advanced statistical and machine learning techniques, text mining, and web mining techniques, that are employed on the customer data in order to solve various business problems related to customers. ACRM includes customer segmentation, target marketing, product and service recommendation via Market Basket Analysis, credit scoring, default detection, churn detection, customer lifetime value modeling, fraud detection, customer sentiment analysis, customer profitability analysis, etc. In other words, ACRM is the analytical engine which, by way of solving various business problems, achieves customer satisfaction via gaining customer loyalty

* Corresponding author.
E-mail addresses: krishna.gutha@gmail.com (G.J. Krishna), padmarav@gmail.com (V. Ravi).

http://dx.doi.org/10.1016/j.engappai.2016.08.012

also known as customer stickiness. ACRM also helps in acquisition of customer knowledge (Xu and Walton, 2005). Link between analytics and CRM is manifested in the form of ACRM in that analytics is exploited to enhance customer service by discovering business insights which cannot be obtained otherwise through standard business rules (Artun and Levin, 2015; Chorianopoulos, 2015; Provost and Fawcett, 2013). Before the advent of these analytical techniques, the business problems were solved in an ad-hoc manner using business rules, thereby missing the science part of the whole thing. Finally, in collaborative CRM, CRM is performed while servicing the customer requirements in almost real time (Kracklauer et al., 2004). This involves the implementation of newer technologies such as real-time data warehousing, in-memory analytics, and geolocation analytics. The entire fabric of CRM subsuming all the three components is called holistic CRM In Teo et al. (2006), holistic CRM is applied to housing and development board of Singapore for offering better services. Given the activities of three types of CRM, it is evident that evolutionary computation has a significant role to play in the analytical CRM component for the obvious reason that all the tasks in the ACRM or the business problems can be collapsed into data mining tasks, such as classification, clustering, forecasting/regression, association rule mining and outlier/anomaly detection. Rygielski et al. (2002) presented several data mining techniques which can be applied to solve CRM tasks. For the purpose of this survey, henceforth, we use CRM and ACRM synonymously.

1.1. Customer lifecycle

Typically there are three phases of the customer lifecycle as depicted in Fig. 1. Understanding customer lifecycle is critical to the successful employment of EC techniques in solving the business problems in an analytical way.

1.1.1. Acquire new customers

Acquiring new customers no longer follows mass marketing, that is, wherein every customer, irrespective of his/her needs would be contacted while selling/offering a product or a service. This simply backfires as the optimal customer acquisition method would gather all the customer data, identify an appropriate customer segment, identify their needs and accordingly, customize the products or suggest the available products to the group of customers. This is the concept behind target marketing. Thanks to the availability of sophisticated technologies such as social media analytics, geo-location analytics, and faster analytical techniques, by which nowadays even personalized marketing is feasible. This strategy is being implemented in several service industries.

1.1.1.1. I. Customer segmentation. Segments are homogeneous gatherings of comparative buyers with comparative needs and desires (Greenberg, 2009). Customer segmentation is the division of customer base into a small number of homogeneous groups that have comparable qualities or attributes. Customer segmentation is an effective tool for recognizing unfulfilled customer needs. Organizations can identify underserved segments and then beat the competition by serving them suitably and growing remarkably in the whole process. Customer segmentation can be termed successful when an organization caters to those segments that are the most profitable. It helps organization practice target marketing and eventually leading to personalized marketing. Amazon, Google, and other such companies already practice one-to-one or personalized marketing. This prioritization can help organizations evaluating techniques to both high- and low-benefit clients (Greenberg, 2009).

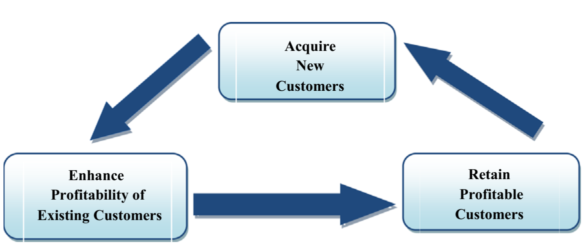


Fig. 1. Customer life cycle.

1.1.1.2. Direct marketing, customer campaigning and service marketing. Direct Marketing is performed by purchasing the database of customer profiles and immediately showcasing the product to a specific group of customers, whereas, customer campaigning is performed by presenting a product to the customers by way of advertisements, rebates and regular mails (Kotler and Keller, 2012).

Service marketing incorporates showcasing of services to the customer, for example, information transfer services, financial services, rental services, insurance services, expert services, etc. (Kotler and Keller, 2012).

1.1.2. Enhance profitability of existing customers

Profitability of existing customers is enhanced in the following CRM tasks:

1.1.2.1. Customer lifetime value (CLV). CLV or CLTV is also called lifetime customer value (LCV), or lifetime value (LTV). It is one of the important CRM tasks. It is an expectation of the net revenue, through tangible or intangible benefit obtained by future association with a customer. It is the present value of money streams associated with the customer by his/her whole relationship with the organization (Greenberg, 2009). This can be estimated by the recency, frequency and monetary (RFM) aspects of his/her association with the company.

1.1.2.2. Customer profiling based on credit scoring. Credit scoring assessment results in a numerical expression taking into account the individual's credit documents that speak financial status of the individual. The rating is arrived at basically from the credit report data normally sourced from credit agencies. It can also be utilized to profile customers (Greenberg, 2009). Credit scoring assesses the credit worthiness of a customer.

1.1.2.3. Market Basket Analysis (MBA). Market Basket Analysis is based on the hypothesis that many customers tend to purchase a certain group of things together (Han et al., 2011). Therefore, after applying association rule mining algorithms on the customers' transactional data, we obtain association rules which together represent the knowledge. Using this knowledge, the customers who purchased a subset of those products, can also be cross-sold other products, which are not yet purchased by them. Cross-sell and up-sell constitute the MBA. While cross-sell involves recommending those products not already held by a customer, up-sell typically involves recommending the upgraded products in the same category.

1.1.2.4. Fraud detection. Fraud is deliberately acting to deny another from claiming his/her worth for one's own specific benefit. Fraud manifests itself in diverse forms. On the other hand, of late, the technological advancements while providing a lot of comfort and convenience to the customer also leave ample scope for perpetrating technology driven fraud. Detecting and identifying fraud becomes very critical to the profitability and reputation of a firm (Palshikar, 2002).

1.1.2.5. Default detection. Credit scoring and Default detection are two sides of the same coin. While credit scoring happens before a loan is granted, default detection happens after a loan is granted. The default is the inability to meet the legitimate commitments. Default happens when the borrower has not made a planned installment of interest or principal, typically in specified period of time. Technically, default occurs when an agreed upon or a negative pledge is violated. Defaults can be predicted based on the customers' transactional and credit related data thereby saving huge losses to the financial firm (Greenberg, 2009).

1.1.3. Retain profitable customers

1.1.3.1. I. Customer churn detection. Customer retention is much easier and cheaper proposition than customer acquisition because we have complete data about the existing customers. If we can analyze that data meticulously and perform cross-sell and up-sell, then they will become more profitable than before. However, we hardly have any information about the prospective or new customers. Despite this fact, service industries do lose profitable and loyal customers due to competition and various other reasons. In order to address this tendency, using predictive analytics, the firms must predict well in advance the set of customers who are going to attrite so that the marketing teams swing into action immediately to have a dialog with them and retain them. In that sense, customer churn detection and customer retention are two sides of the same coin (Van den Poel and Larivière, 2004).

1.1.3.2. Sentiment analysis. It is a social CRM task wherein a customer trend and perception analysis is done by taking data from social networking sites, blogs, compliant portals, etc. (Greenberg, 2009). The sentiments expressed by disgruntled customers via call centers records and social media are essentially unstructured data, which should be analyzed along with the demographic and transactional data of customers in order to come out with higher accurate predictions about who would churn out in the near future compared to the case when unstructured data was not used at all. Sentiment analysis can also be used as an input to a recommender system for effective cross-selling and up-selling.

1.2. Evolutionary computation

Evolutionary computing is the aggregate name for a set of problem-solving strategies inspired by biological evolution; for instance, natural selection and genetic re-engineering. Swarm intelligence that is inspired by the flocking and foraging behavior of insects, birds, and other creatures also forms part of this grand set. These techniques have found a myriad of applications in diverse disciplines. They form a subset of a bigger set called optimization techniques, which include classical and derivative based techniques, point based meta-heuristics such as simulated annealing, tabu search, etc.

Evolutionary algorithms (EA) are, without exception, population-based random search techniques where a population of solutions gets updated iteratively using algorithm-specific heuristics until convergence is achieved. These techniques differ in the way in which the solutions are encoded and the heuristics they adopt when updating the solutions (Back et al., 1997). Genetic Algorithms, Differential Evolution, Particle Swarm Optimization, Ant Colony Optimization, Firefly Algorithm, cuckoo search, artificial bee colony, and the harmony search algorithm are some of the members of this family.

1.3. Optimization techniques and data mining (DM)

Numerous DM techniques require optimization at various stages, such as feature selection, training a classifier/regression model or determining the optimal user-defined parameter combination that would result in accurate results for the technique. Fitting of logistic regression, wrapper-based feature selection techniques, parameter and structure optimization of the neural network, finding the optimal parameter combination for a support vector machine, fine tuning the parameters of a fuzzy inference system, and optimizing the rule base size in a fuzzy rule-based classifier are some of the examples in this category (Padmanabhan and Tuzhilin, 2003). On the other hand, evolutionary techniques can be employed in standalone mode to solve various data mining tasks such as classification, regression, clustering, association rule mining, and outlier detection, along with subtasks such as feature selection. Given this established versatility of EC techniques, one can safely imagine that EC techniques alone can be employed to solve various ACRM tasks. This survey centers around this unique strength of EC techniques and also includes the review of papers involving both standalone and hybrid EC techniques in a single- as well as a multi-objective optimization framework.

The works dealing with soft computing hybrids involving EC on one hand and machine learning techniques on the other are outside the scope of this survey.

1.3.1. Evolutionary computation (EC) based data mining

Having observed that EC techniques can be used as a viable alternative to traditional DM techniques, one has to remember that the performance of EC techniques is sometimes superior to that of traditional DM techniques, given the data dependent nature of all of these techniques encapsulated in the no-free-lunch theorem (Wolpert and Macready, 1997).

CRM tasks can be formulated as data mining (DM) techniques. Data mining and optimization go hand-in-hand when solving CRM tasks. We can apply optimization alone for solving CRM tasks, or we can apply optimization before or after applying DM techniques (Padmanabhan and Tuzhilin, 2003).

Fig. 2 depicts various CRM tasks and the corresponding DM techniques used traditionally to solve CRM tasks.

The rest of the paper is organized as follows: Section 2 describes review methodology. Section 3 presents the distribution of the reviewed articles. In Section 4, review of works in the current study is presented. Section 5 presents the discussion of the review and future directions. Section 6 concludes the paper.

2. Review methodology

2.1. Techniques relevant to the review

In an earlier survey, Ngai et al. (2009) reviewed pure and hybrid data mining (DM) techniques and DM techniques hybridized with optimizations techniques applied to CRM tasks. Therefore, for our review paper to be unique, we considered stand-alone and hybrid, EC techniques applied in single- and multi-objective framework and reviewed those papers involving them. Fig. 3 summarizes the above.

In Appendix A, two tables are presented summarizing the EC techniques utilized in the current review. Table A1 presents a brief overview of EC techniques that are relevant to the current review with their advantages and disadvantages. Table A2 presents the objective functions, their properties and parameters that are set in each research article.

2.2. Flowchart of review process

Online databases are used for searching research articles in the survey. The order of preference for search is journals, book chapters, and conference proceedings. Research articles which have the evolutionary algorithms applied to CRM tasks are searched in the

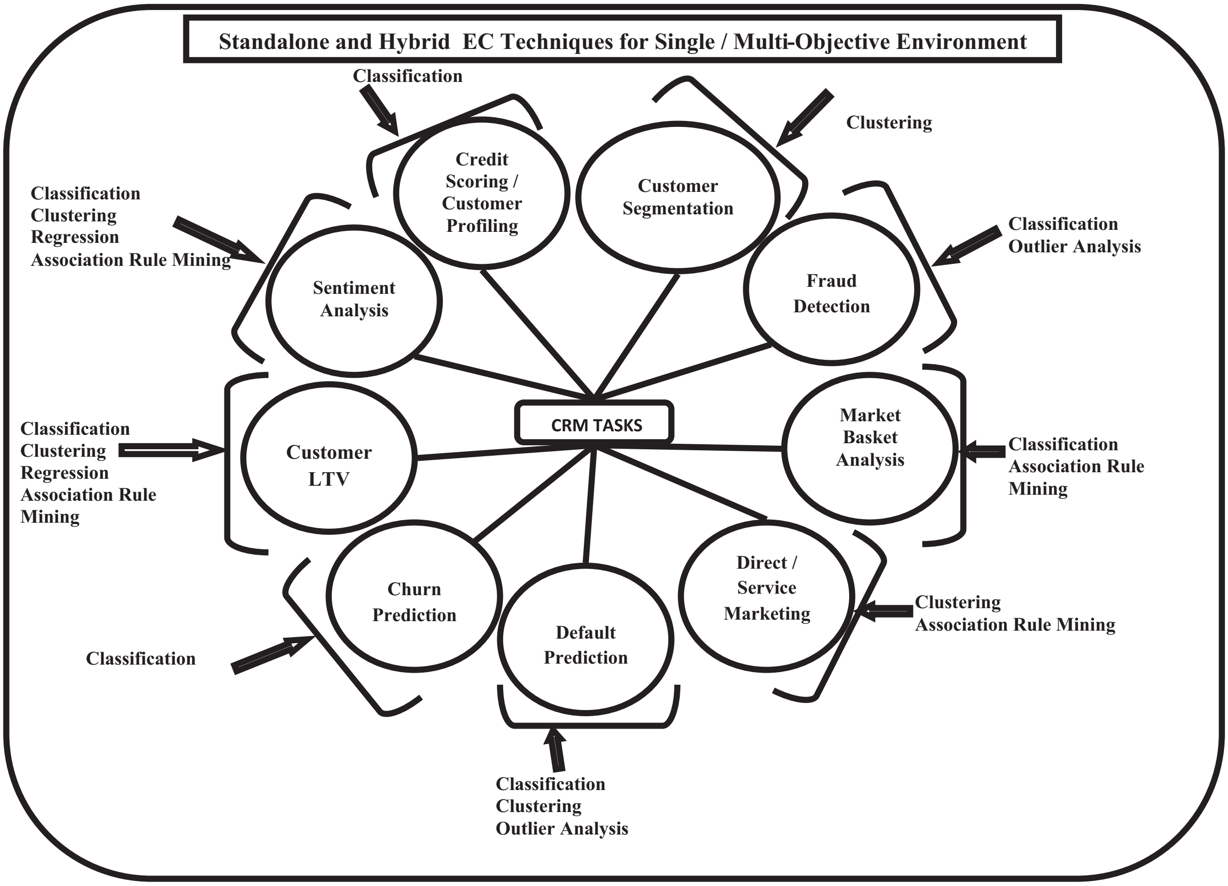


Fig. 2. Relationship between CRM tasks, data mining, and optimization (EC) Techniques.

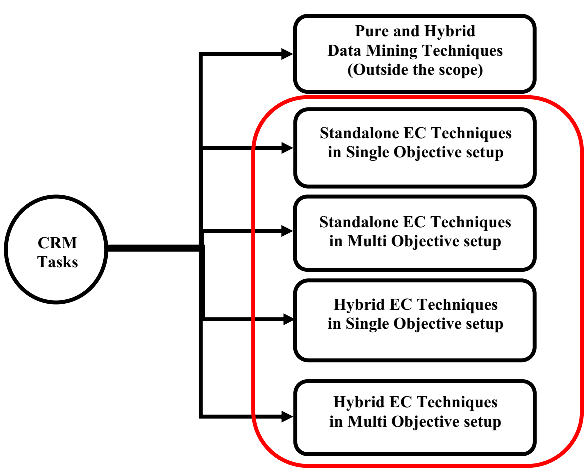


Fig. 3. Structure of considered techniques.

online databases. Then, the data set relevance to the banking, telecom, marketing, retail, and manufacturing sectors is checked in the research articles. The flowchart presented in Fig. 4 depicts the methodology followed in collecting the articles for the review.

3. Distribution of reviewed articles

3.1. Count of research articles reviewed by various journals/books/proceedings

We surveyed around 35 papers that appeared in international journals with relevance to the topic of interest as shown in the flow chart. 43 book chapters and proceedings of conferences were also surveyed. So, a total of 78 research articles were surveyed. Expert Systems with Applications has the highest number of surveyed papers in journals (11), and IEEE Proceedings have the highest number of conferences (28). The summary is presented in Table 1.

3.2. Distribution of research articles by CRM task

This subsection summarizes the count of research articles arranged according to the CRM tasks in the Tables 2 and 3. From Tables 2 and 3, we can infer that customer lifetime value (LTV), default prediction, service marketing, and customer campaigning have the possible scope of further research for applying EC techniques. Tables 2 and 3, Figs. 5 and 6 depict the distribution.

3.3. Distribution of research articles by EC technique

This subsection summarizes the count of research articles arranged by EC techniques in Tables 4–9 and shown in Figs. 7 and 8. From the Tables 4–9, we can infer that GA was used extensively. We can also infer that there are also other EC techniques applied to CRM tasks, which is a direction for further research. Also, multi-objective optimization framework can also be considered for further research apart from single objective optimization framework. Hybrid EC frameworks developed by combining one or more EC frameworks can also be utilized for the CRM tasks as they feed on merits of one and another.

3.4. Datasets used in reviewed articles

Specific companies' datasets were analyzed in 29 papers. Bank datasets were analyzed in ten (10) papers. German Credit Scoring

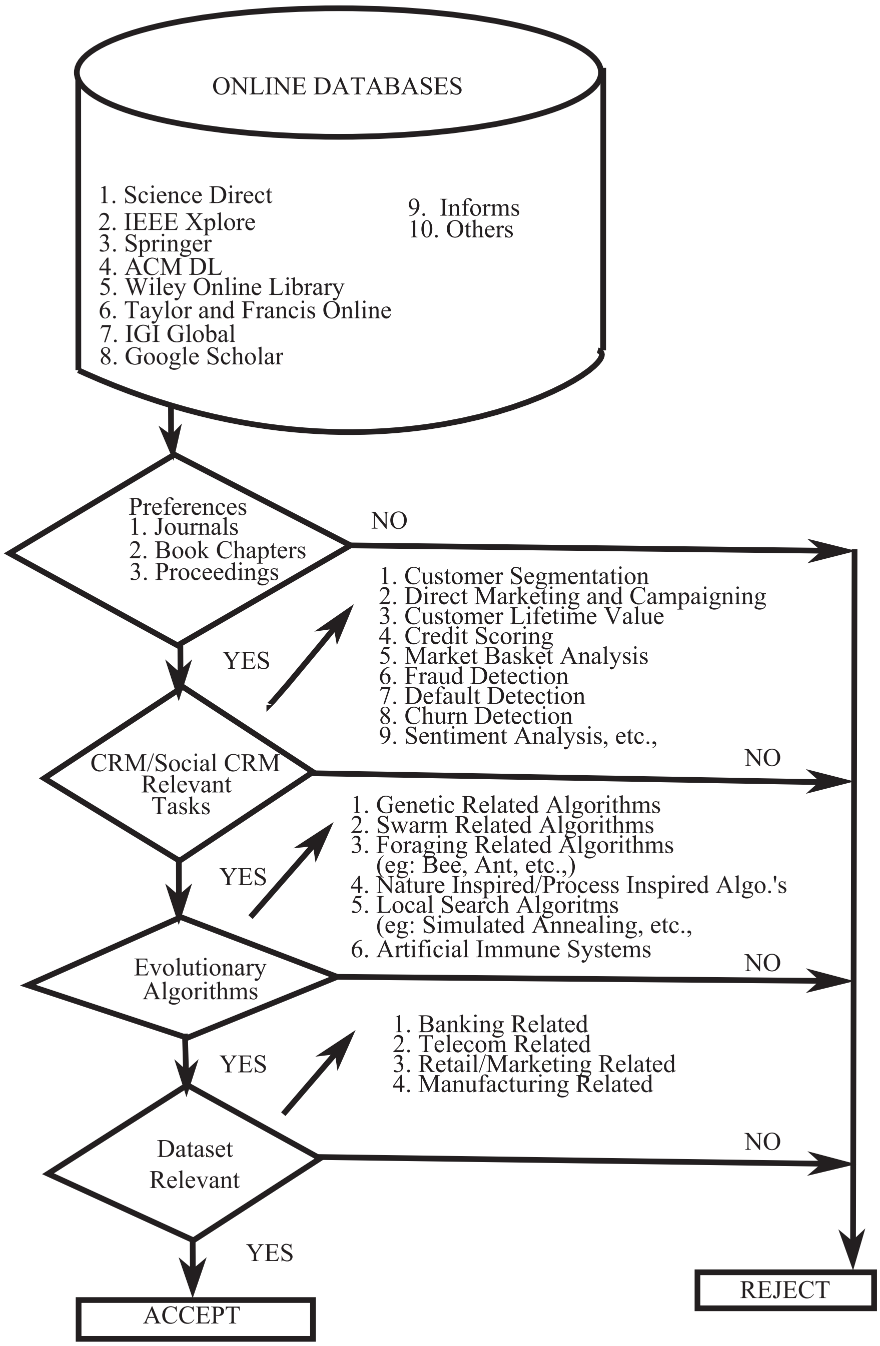


Fig. 4. Flowchart of review methodology.

dataset was analyzed in nine (9) papers. Telecom datasets were analyzed in six (6) papers. Australian Credit Scoring dataset was analyzed in five (5) papers. Credit Card datasets were analyzed in five (5) papers. Government and Stock Market datasets were analyzed in five (5) papers. Retail datasets were analyzed in four (4) papers. Survey datasets were analyzed in three (3) papers. The point of sales data was analyzed in two (2) datasets. All the details are summarized in Table 10.

3.5. Trends

Figs. 9–11 depict the trends of the number of publications by year in journals, conferences and both of them together respectively. These figures indicate that the trend is increasing starting from 1998 to 2015.

4. Review of works in the current study

In what follows, we now present a brief overview of the research studies conducted in each of the CRM tasks with the help of EC techniques during the period of review.

4.1. Credit scoring

Financial service firms face an intriguing choice of whether to grant a loan to an applicant or not. The problem of determining the creditworthiness of a customer based on his/her data is called credit scoring. Credit scoring is a binary classification problem for which many traditional techniques exist, including both statistical and machine learning methods. However, of late, stand-alone

Table 1
Count of surveyed papers.

Articles in journals/books/proceedings	Number
Expert systems with applications	11
Applied soft computing	3
Journal of retailing	2
Knowledge based systems	1
Journal of interactive marketing	1
Information sciences	1
Data and knowledge engineering	1
IEEE transactions on evolutionary computation	1
Computational intelligence	1
INFORMS journal on computing	1
Applied mathematics and computation	1
Cybernetics and systems: an international journal	1
Journal of marketing research	1
Journal of system science and systems	1
International journal of artificial intelligence and applications	1
An international journal: computers and mathematics with applications	1
Journal of research and reviews in information technology	1
Intelligent systems in accounting, finance and management	1
Engineering applications of artificial intelligence	1
Handbook of research on novel soft computing intelligent algorithms: theory and practical applications	1
Soft computing for data mining applications: studies in computational intelligence	1
Information systems journal	1
IEEE proceedings	28
Springer book chapters	8
ACM proceedings	2
Elsevier procedia	1
Taylor and Francis proceedings	1
Others	3
Total	*78*

Table 2
Count based on CRM tasks.

CRM task (using EC techniques)	No. of journals	No. of conferences/book chapters	Total
Customer segmentation	5	2	7
Fraud detection	3	9	12
Market Basket Analysis	9	7	16
Credit scoring	6	7	13
Customer LTV	2	2	4
Customer churn detection	6	5	11
Default detection	1	1	2
Service marketing	–	1	1
Direct marketing	3	2	5
Customer campaigning	–	1	1
Sentiment analysis	1	8	9

evolutionary techniques have also been applied to solve scoring customers.

Credit scoring models built using the Genetic Algorithm (GA), Genetic Programming (GP), Multi-Objective Evolutionary Algorithm (MOEA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO) and Simulated Annealing (SA) are reviewed in this subsection.

GP has attracted the attention for the following reasons (i) removal of irrelevant attributes (ii) determination of correct discriminant function (iii) used for both small and large datasets. So, Ong et al. (2005) developed Credit Scoring models, using GP. Earlier, the same credit scoring models were developed using, Neural Networks, Decision Trees, Rough Sets and Logistic Regression. Australian credit scoring and German credit scoring datasets were used. The results obtained were better than those from earlier models. The GP-based model is better than statistical methods as there was no need for assumptions on the dataset. Also, the GP-based model is flexible and accurate.

According to Ong et al. (2005), GP is better than the remaining models, and there is a need for intelligence rules for the decision-maker to understand the meaning of the dataset. For this purpose, Huang et al. (2006), proposed the two-stage GP for building a credit scoring model. The first stage of GP was employed to derive the IF-THEN rules for the decision maker. In the second phase of GP, the reduced dataset is applied in building the discriminant function for providing the capability of forecasting. The data sets used are the German Credit Scoring Data Set and the Australian Credit Scoring Data. The results obtained are better than those reported by Ong et al. (2005). Apart from finding discriminant function, IF-THEN rules have also been incorporated for even better accuracy compared to that of Ong et al. (2005).

Behavior Scoring is an important task in risk management. Chen et al. (2007) introduced Modified Genetic Programming (MGP) to solve behavior scoring problems on Credit Card Data provided by the Chinese Commercial Bank. MGP showed better accuracy for behavior scoring. MGP is compared with GP with normalized inputs (NGP) and back-propagation neural network (BPN). MGP performs well for small datasets. In future, MGP can be combined with other artificial intelligence techniques in order to classify based on rules generated. The behavior of the customers who had defaulted, the timing of the default and feature selection can also be investigated.

Derived Characteristics are usually regarded as the important index in credit scoring, and the selection of derived characteristics is formulated as an optimization problem. GP and human intervention are applied to obtain derived characteristics. Liu et al. (2008) applied GP for selecting derived characteristics that are important in credit scoring. GP utilized 47,000 samples and 76 original characteristics of Finance Enterprise, and the outcomes are satisfactory. Later, discrimination analysis is used on the set of valuable derived characteristics, and the results turned out to be good.

A comparison, between conventional methods and non-conventional ones, has also been carried out in Abdou (2009) where GP was implemented for Credit Scoring with a case of Egyptian Public Sector Banks. The GP model was compared with the weight of evidence measure and the probit models which are the conventional methods for credit scoring. The average correct classification (ACC) was high and estimated misclassification cost (ECM) was lower than those of other models. This study can further include mortgages, house loans, and corporate loans. Also, employing other evaluation criteria, such as GINI coefficient and ROC curve, would be useful.

Scoring models have been developed to model/predict credit risk. These scoring models are categorized based on statistical methods like logistic regression, classification trees, nonlinear regression, nearest neighbor approach, etc., and non-statistical methods like GA, neural networks, etc. Cai et al. (2009) solved the personal credit scoring problem with GA using the German Credit Dataset (http://archive.ics.uci.edu/ml/), and their model proved to be of high efficiency when evaluating the significance values. But, GA is more complicated than other methods, and its computing time is also high. Nevertheless, it can be optimized further to reduce computational time and increase efficiency.

SA is a local search meta-heuristic that can be used for feature selection, classification, and clustering. Dong et al. (2009) presented the SA-based rule extraction algorithm for credit scoring and performed experiments on German credit dataset. The SA-based model outperformed statistical techniques like Decision Trees (DT), K-nearest neighbor, linear discriminant analysis (LDA), and quadratic discriminant analysis (QDA). The benefits of the proposed model are that it can extract rules as well as find attribute correlations. Its limitations are that the parameters have to be fine-tuned and training time is high.

Table 3
References based on CRM tasks.

CRM task (using EC techniques)	Journal references	Conferences/book chapters references
Customer segmentation	(Brusco et al., 2002; Chan, 2008; Chiu and Kuo, 2010; Chiu et al., 2009; Mahdavi et al., 2011)	(Mzoughia and Limam, 2014; Zhang, 2008)
Fraud detection	(Hoogs et al., 2007; Soltani Halvaiee and Akbari, 2014; Wong et al., 2012)	(Assis et al., 2014; Brabazon et al., 2010; Duman and Elikucuk, 2013; Gadi et al., 2008; R. Huang et al., 2010; Jungwon et al., 2003; Brun et al., 2009; Ozçelik et al., 2010; Soltani et al., 2012)
Market Basket Analysis	(Chien and Chen, 2010; Ghosh and Nath, 2004; Hansen et al., 2010; Kuo and Shih, 2007; Kuo et al., 2011; Luna et al., 2013; Sarath and Ravi, 2013; Shenoy et al., 2005; Yang et al., 2011)	(Bhugra et al., 2013; Birtolo et al., 2013; Cheng, 2005; Christian and Martin, 2010; Cunha and Castro, 2013; Ganghishetti and Vadlamani, 2014; Khademolghorani, 2011)
Credit scoring	(Abdou, 2009; Fogarty, 2012; Huang et al., 2006; Kožený, 2015; Ong et al., 2005; Wang and Huang, 2009)	(Aliehyaei and Khan, 2014; Cai et al., 2009; Corazza et al., 2014; Dong et al., 2009; Hochreiter, 2015; Liu et al., 2008; Waad et al., 2014)
Customer LTV	(Chan, 2008; George et al., 2013)	(Mzoughia and Limam, 2014; Zhang, 2008)
Customer churn detection	(Au et al., 2003; Bandaru et al., 2015; Gao et al., 2014; B. Huang et al., 2010, 2012; Subramaniam and Thangavelu, 2011)	(Basiri and Taghiyareh, 2012; Eiben et al., 1998; Faris et al., 2014; Qiwan and Min, 2008; Sotoodeh, 2012)
Default detection	(Gordini, 2013)	(Bozsik, 2010)
Service marketing	–	(Qiwan and Min, 2008)
Direct marketing	(Bhattacharyya, 1999; Cui et al., 2015; Jonker et al., 2004)	(Bhattacharyya, 2000; Cui et al., 2010)
Customer campaigning	–	(Dehuri et al., 2008)
Sentiment analysis	(Guo et al., 2013)	(Banati and Bajaj, 2012; Carvalho et al., 2014; Gupta et al., 2015; Hochreiter, 2015; Huang et al., 2011; Kaiser et al., 2010; Yamada and Terano, 2006; Yamada and Ueda, 2005)

Fig. 5. CRM tasks covered in literature using EC.

Credit Scoring can also be performed applying MOEA by proposing new criteria for feature selection. Wang and Huang (2009) employed feature selection techniques along with MOEA feature selection with relative correlations. Then, various classifiers such as SVM, multi-layer perceptron are used on these features. German and Australian credit datasets have been analyzed. MOEA yielded superior AUC than other feature selection methods. This work can be extended by applying data pre-processing and feature selection for various classifiers on large-scale problems and also new multi-objective algorithms incorporating various performance metrics can also be considered.

In large organizations, it is very difficult to update continuously complex credit scoring systems and it is better to automate the scoring maintenance functions within firms. To this end, Fogarty (2012) developed maintenance functions based on the credit scoring system using GA. The scorecard prepared using GA performed as well as or, if not better than the traditional methods. GA is used to refine the score created by the initial model driven by traditional techniques. Data from Equifax, Experion and Trans Union Credit Bureaus is collected and used. Marketing models have the highest improvement on an auto loan in the Pacific Rim region and risk models also have an improvement for the credit cards of the Pacific Rim region. Limitations are that GA needs to have a lot of computational power, data, and clear definition of objective function for the business. Also, there is no generic model for all problems and data. Further merit of GA over other search techniques has to be investigated.

Here, we discuss a hybrid, GP+ACO, for credit scoring. Aliehyaei and Khan (2014) investigated GP, the modified ACO, and hybrid GP-ACO for credit scoring. German and Australian credit datasets were analyzed, and the hybrid approach was shown to perform the best of the other three methods. In this hybrid, the insight of ACO is input to GP for a head start. Further investigation into the better performance of ACO and impact of improving GP were to be carried out.

Preference Disaggregation is a concept based on simple findings that, in general, real-world decision makers are unable or unwilling to provide. Corazza et al. (2014) proposed a PSO based preference disaggregation in the multi-criteria credit worthiness of a set of firms. The data set was provided by the major bank in North Eastern Italy and the results obtained were highly consistent. The comparison was made with actual parameters of MUlticriteria RAnking MEthod (MURAME) and parameters obtained by PSO for ordering of the alternatives. Further, research using realistic computational time and finding the values of parameters of the MURAME model that make them consistent with the classification provided by the banks for credit scoring have to be performed.

Waad et al. (2014) employed the GA-based feature selection technique to the credit scoring data of German and Australian Credit Datasets (http://archive.ics.uci.edu/ml/), Tunisian Bank Dataset, and the HMEQ Dataset. They concluded that the proposed technique is robust and accurate, on performance metrics like precision, recall, and F-measure.

Parameter range estimation, polynomial fitness functions, and an alternate fitness function called weighted bitmask were applied to GA for building credit scoring models (Kožený, 2015). A sample of 489 credit card customers' data from Venezuela was used. The sensitivity, specificity, and accuracy of the model with three fitness functions were compared. The weighted bitmask fitness function

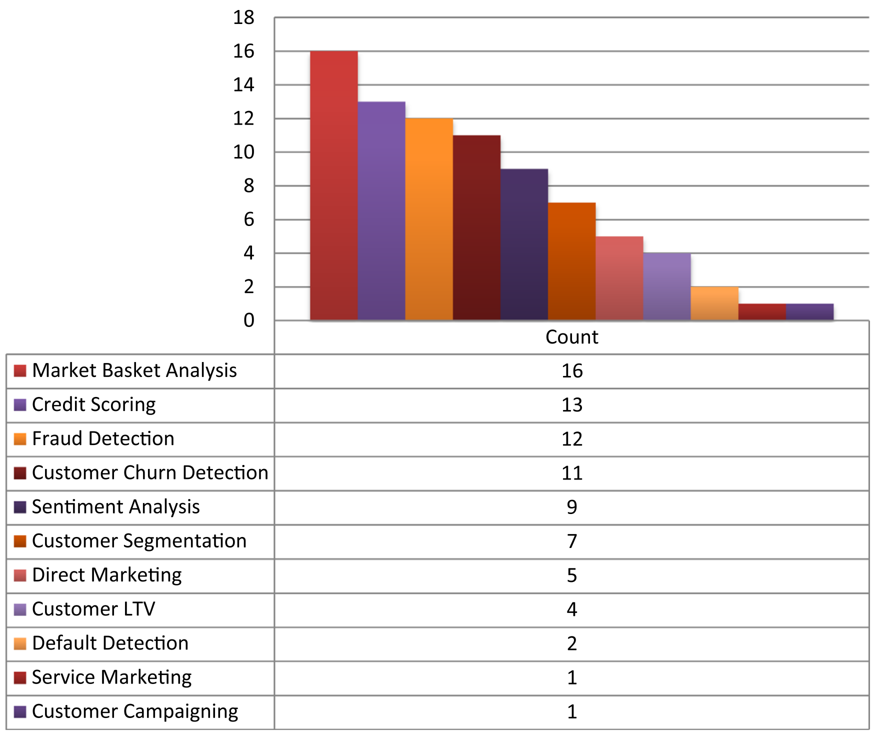


Fig. 6. Column chart based on count of CRM tasks.

Table 4
Count based on single objective EC techniques.

Evolutionary algorithms (used for CRM task)	No. of papers in journals	No of papers in conferences/book chapters	Total
Genetic Algorithms (GA)	15	13	28
Genetic Programming (GP)	3	6	8
Interactive Genetic Algorithm (IGA)	–	1	1
Genetic Network Programming (GNP)	1	–	1
Grammar Guided Genetic Programming(G3P)	1	–	1
Particle Swarm Optimization (PSA)	4	3	7
Bee hive	1	–	1
Honey bee	–	1	1
Ant Colony Optimization (ACO)	1	5	6
Ant Cluster Algorithm (ACA)	1	1	2
Honey Bee Mating (HBM)	1	–	1
Migrating Birds Optimization (MBO)	–	1	1
Simulated Annealing (SA)	1	1	2
Imperialist Competitive Algorithm (ICA)	–	2	2
Biogeography Based Optimization (BBO)	–	1	1
Firefly Algorithm (FFA)	–	1	1
Differential Evolution (DE)	–	1	1
DMEL[a]	2	–	2
Artificial Immune System (AIS)	3	6	9

[a] DMEL = Data mining by evolutionary learning.

yielded better results than others as indicated by a statistical test. The major limitations are the availability of a quality database and the computational complexity of the maximum bitmask size. Different encoding schemes can also be used to reduce computational complexity.

Table 5
Count based on multi-objective EC techniques.

Multi-Objective Evolutionary Algorithms (used for CRM task)	No. of journals	No of conferences/ book chapters	Total
Multi-Objective Genetic Algorithm	1	2	3
NSGA-II*	4	–	4
SPEA-2*	1	–	1
MO-BPSO*	–	1	1
MO-BPSO-TA*	–	1	1
MO-BFFO-TA*	–	1	1

*NSGA-II = Non-dominated Sort Genetic Algorithm-II.
*SPEA-2 = Strength Pareto Evolutionary Algorithm-2.
*MO-BPSO = Multi-Objective Binary Particle Swarm Optimization.
*MO-BPSO-TA = Multi-Objective-Binary Particle Swarm Optimization-Threshold Accepting.
*MO-BFFO-TA = Multi-Objective-Binary Firefly Optimization-Threshold Accepting.

Table 6
Count based on hybrid EC Ttechniques.

Hybrid evolutionary algorithms (used for CRM task)	No. of journals	No of conferences/book chapters	Total
GP-ACO*	–	1	1
IACA*	1	–	1
PSHBMO*	1	–	1
BPSO-TA*		1	1
BFFO-TA*		1	1

*GP-ACO = Genetic Programming-Ant Colony Optimization.
*IACA = Immunity-based Ant Clustering Algorithm.
*PSHBMO = Particle Swarm Honey Bee Mating Optimization.
*BPSO-TA = Binary Particle Swarm Optimization-Threshold Accepting.
*BFFO-TA = Binary Firefly Optimization-Threshold Accepting.

4.2. Customer segmentation

Customer Segmentation is a major strategy that entails the grouping of customers into homogeneous clusters so that products

Table 7
Use of EC techniques in single objective framework.

Evolutionary algorithms (used for CRM task)	No. of journals	No of conferences/book chapters
Genetic Algorithm	(Bhattacharyya, 1999; Chan, 2008; Chien and Chen, 2010; Cui et al., 2015; Fogarty, 2012; George et al., 2013; Ghosh and Nath, 2004; Gordini, 2013; Guo et al., 2013; Hansen et al., 2010; Hoogs et al., 2007; Jonker et al., 2004; Kožený, 2015; Mahdavi et al., 2011; Shenoy et al., 2005)	(Birtolo et al., 2013; Bozsik, 2010; Cai et al., 2009; Carvalho et al., 2014; Cheng, 2005; Christian and Martin, 2010; Cui et al., 2010; Huang et al., 2011; Ozçelik et al., 2010; Qiwan and Min, 2008; Waad et al., 2014; Yamada and Terano, 2006; Yamada and Ueda, 2005)
Genetic Programming	(Abdou, 2009; Huang et al., 2006; Ong et al., 2005)	(Aliehyaei and Khan, 2014; Assis et al., 2014; Eiben et al., 1998; Faris et al., 2014; Hochreiter, 2015; Liu et al., 2008)
Interactive Genetic Algorithm	–	(Qiwan and Min, 2008)
Genetic Network Programming	(Yang et al., 2011)	–
Grammar Guided Genetic Programming (G3P)	(Luna et al., 2013)	–
Particle Swarm Optimization	(Chiu and Kuo, 2010; Gao et al., 2014; Kuo et al., 2011; Sarath and Ravi, 2013)	(Corazza et al., 2014; Ganghishetti and Vadlamani, 2014; Gupta et al., 2015)
Bee hive	(Subramaniam and Thangavelu, 2011)	–
Honey bee	–	(Dehuri et al., 2008)
Ant Colony Optimization	(Kuo and Shih, 2007)	(Aliehyaei and Khan, 2014; Jin et al., 2006; Kaiser et al., 2010; Sotoodeh, 2012; Zhang, 2008)
Ant Cluster Algorithm	(Chiu et al., 2009)	(Jin et al., 2006)
Honey Bee Mating	(Chiu and Kuo, 2010)	–
Migrating Birds Optimization	–	(Duman and Elikucuk, 2013)
Simulated Annealing	(Brusco et al., 2002)	(Dong et al., 2009)
Imperialist Competitive Algorithm	–	(Basiri and Taghiyareh, 2012; Khademolghorani, 2011)
Biogeography-Based Optimization	–	(Bhugra et al., 2013)
Firefly Algorithm	–	(Banati and Bajaj, 2012)
Differential Evolution	–	(Brun et al., 2009)
DMEL	(Au et al., 2003; Huang et al., 2012)	–
Artificial Immune System	(Chiu et al., 2009; Soltani Halvaiee and Akbari, 2014; Wong et al., 2012)	(Brabazon et al., 2010; Cunha and Castro, 2013; Gadi et al., 2008; Huang et al., 2010; Jungwon et al., 2003; Soltani et al., 2012)

and services can be customized or tailor-made for a given cluster of customers, thereby reducing huge operational costs. Here the assumption is one size does not fit all, which is very well proven in the real world applications. Hence, the need of clustering. In this sub-section, we will review the customer segmentation models built using GA, the Multi-Objective Genetic Algorithm (MOGA), PSO, Honey Bee Mating (a variant of Bee Swarm), ACO, and SA.

SA has been applied to market segmentation. Brusco et al. (2002) carried out the segmentation of the Bi-criterion (identifiably, responsiveness) Partitioning problem using SA. Also, Simulated Annealing Heuristic (SAH) was compared with EXCLU (Exogenous variable-related Clustering), and in turn, SAH used EXCLU and EXCLU-R (Reverse Version) in developing a better model. The national technology survey data was used. The results indicated that the SAH outperformed k-means and EXCLU. The limitations are that SAH does not guarantee the identification of global optimum, and also, parameter tuning is hard. The following extensions are possible: 1) objective function is not sensitive to changes in functional form of the objective and can be modified, 2) three (explanation of variance for two responses variables is of separate interest) or more criteria can also be utilized, 3) adaptation to the problem of overlapping clusters is possible, where fuzzy clustering can also be considered.

Table 9
Use of hybrid EC techniques.

Name of the hybrid EC Techniques (used for CRM task)	References of journals	References of conferences/book chapters
GP-ACO	–	(Aliehyaei and Khan, 2014)
IACA	(Chiu et al., 2009)	–
PSHBMO	(Chiu and Kuo, 2010)	–
BPSO-TA (association rule mining, MBA)	–	(Ganghishetti and Vadlamani, 2014)
BFFO-TA (association rule mining, MBA)	–	(Ganghishetti and Vadlamani, 2014)

The following study combines the customer targeting and customer segmentation for campaigning. Chan (2008) carried out a marketing campaign by evaluating CLV from customer segments

Table 8
Use of multi-objective EC techniques.

Multi objective evolutionary algorithms (used for CRM task)	Journal references	Conferences/book chapters references
Multi-Objective Genetic Algorithm	(Ghosh and Nath, 2004)	(Bhattacharyya, 2000; Mzoughia and Limam, 2014)
NSGA-II	(Bandaru et al., 2015; Huang et al., 2010; Luna et al., 2013; Wang and Huang, 2009)	–
SPEA-2	(Luna et al., 2013)	–
MO-BPSO (association rule mining, MBA)	–	(Ganghishetti and Vadlamani, 2014)
MO-BPSO-TA (association rule mining, MBA)	–	(Ganghishetti and Vadlamani, 2014)
MO-BFFO-TA (association rule mining, MBA)	–	(Ganghishetti and Vadlamani, 2014)

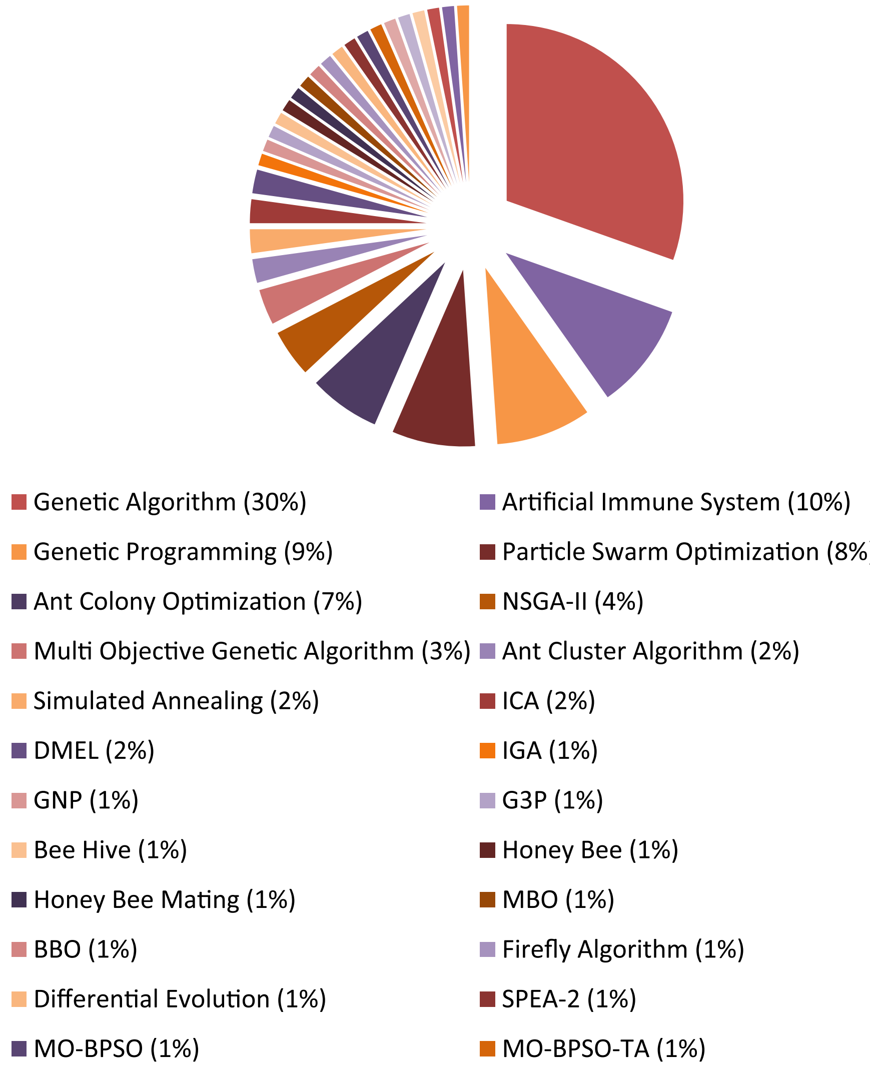


Fig. 7. Use of all EC techniques.

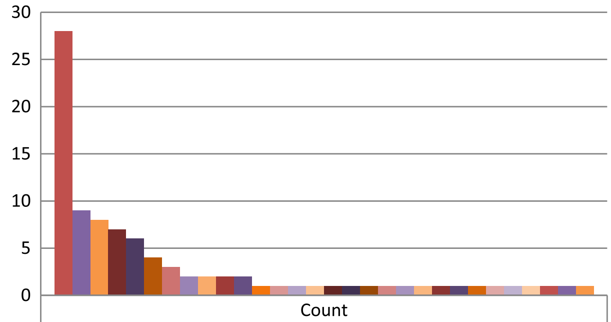
formed by using the RFM model. GA was used to select appropriate customers for the campaign strategy. In this paper, Nissan's historical customer data of nine cars, collected by Empower Corporation, was used. The final results show that the proposed approach can target potential customers. However, there are a few limitations: 1) it requires huge amount of customer data; 2) only one break point was considered; several break points should be considered in the future, 3) too many methods were used for segmentation, and it is difficult to compare them all; and finally, 4) it only considers existing customers; new potential customers should also be targeted.

Changes in customer segments are tracked based on ant colony clustering. Zhang (2008) efficiently performed customer segmentation by ant colony clustering on the Shoes Retailer Dataset and this study can help enterprises to check their strategies for discovering customers. Directions for future research are as follows: 1) Future patterns depend on the past and current patterns, so the customer lifetime value should be considered, 2) The clustering threshold adjustments need to be studied, and 3) Alternative methods can be used for future segmentation.

A hybrid of Artificial Immune System (AIS) and ant algorithm (IACA) has been developed for market segmentation. Ant algorithm is utilized to generate good solutions for clustering while AIS is used for the optimization of clusters. Chiu et al. (2009) generated segments using the immunity-based ant clustering algorithm which is an integration of an ant algorithm and artificial immune system. Data from a large 3C appliance chain store was used. The proposed method outperformed the self-organizing maps and ant clustering algorithm.

A hybrid that integrates swarm optimization and honeybee mating optimization (PSHBMO), for market segmentation, is considered. Chiu and Kuo (2010) performed segmentation using the hybrid PSHBMO. Podak Co.'s Data (an Agent of Panasonic) from Jan 1, 2003 to Jun 15, 2006, was used for RFM computations. The PSHBMO produced less error than PSO+K means and SOM+K means. The model helps to meet the needs of more customers, to increase customer loyalty, and to increase profits. Limitations are that transactions must be completely recorded, and data must be precise. Further research focuses on applying complex market segments and improving the efficiency of clusters.

Designing customer-oriented catalogs in e-CRM aim at maximizing the number of covered customers. Mahdavi et al. (2011) applied a self-adaptive GA (*s*-AGA) for customer catalog segmentation. Real and Synthetic datasets were used. Real data from Belgian retail store was available at FIMI repository, and Synthetic datasets were generated from state-of-the-art IBM data generators. The proposed s-AGA algorithm performs better and consumes reasonable computational time with real data because the parameters are adjusted based on observed performance. It performs better on synthetic data too. Because of s-AGA, the model is



	Count
Genetic Algorithm	28
Artificial Immune System	9
Genetic Programming	8
Particle Swarm Optimization	7
Ant Colony Optimization	6
NSGA-II	4
Multi Objective Genetic Algorithm	3
Ant Cluster Algorithm	2
Simulated Annealing	2
ICA	2
DMEL	2
IGA	1
GNP	1
G3P	1
Bee Hive	1
Honey Bee	1
Honey Bee Mating	1
MBO	1
BBO	1
Firefly Algorithm	1
Differential Evolution	1
SPEA-2	1
MO-BPSO	1
MO-BPSO-TA	1
MO-BFFO-TA	1
GP-ACO	1
IACA	1
PSHBMO	1
BPSO-TA	1
BFFO-TA	1

Fig. 8. Column chart based on count of all EC techniques.

rapid in convergence. It can be made more realistic by replacing a crisp threshold of customer catalog attraction with a probabilistic threshold.

Several studies demonstrate that segmentation based on customer lifetime value (CLV), rather than descriptive variables alone, presents interesting features for marketing. Mzoughia and Limam (2014) proposed a multi-objective GA considering descriptive variables and CLV for segmentation, using the data from credit card transactions of a Retail Bank in North Africa. The proposed method characterized segments well. Finding the optimal number of clusters is difficult and needs investigation. Also, the influence of the multi-objective optimization algorithm is to be investigated.

4.3. Fraud detection

Fraud is characterized as an activity that is undesired, not allowed by law, and illegal. Only a minuscule percent of the activities/transactions are illegal but can have a catastrophic impact. Fraud detection typically is a data mining problem. Data imbalance is the issue in fraud detection, where genuine transactions disproportionately outnumber the fraudulent one, posing a problem for almost all machine learning techniques.

In this subsection, fraud detection models built using GA, GP, Migrating Birds Optimization (MBO), Artificial Immune System (AIS) Differential Evolution (DE) are reviewed.

Work has also been reported in detecting fraud in the retail sector, another sector that does not possess enough knowledge about potential or actual frauds. Jungwon et al. (2003) detected anomalies for combating financial fraud in the retail sector using an AIS called "Computer Immune System for Fraud Detection" (CIFD). Details of the dataset were not provided, but scalability and adaptability of fraud detection on transactions were improved using CIFD. The benefits of CIFD are its adaptability and ability to learn. Further work includes improving CIFD in reducing false positives.

GAs are evolutionary algorithms that aim to obtain better solutions and can also be used in detecting fraud. Hoogs et al. (2007) proposed a GA approach to detect financial statement fraud. The

Table 10
Datasets analyzed in articles.

Reference	*Dataset description*
(Ong et al., 2005)	German and Australian Credit Datasets
(Huang et al., 2006)	German and Australian Credit Datasets
(Kožený, 2015)	Venezuelan Bank Data
(Fogarty, 2012)	Data from Equifax, Experion, and Trans Union Credit Bureaus Data
(Abdou, 2009)	Egyptian Public Sector Banks Data
(Chiu and Kuo, 2010)	Podak Co.'s Data (Agent of Panasonic)
(Brusco et al., 2002)	National Technology Survey Data
(Mahdavi et al., 2011)	Belgian Retail Store Data (FIMI Repository)
(Gordini, 2013)	Data of Defaulting 6200 firms of Italy by the end of 2009
(Subramaniam and Thangavelu, 2011)	Questionnaire Survey from Selling Departments for period of Four Months (Marketing Firms)
(Hoogs et al., 2007)	SEC's (Securities Exchange Commission) Data from May 1998 to March 2004
(Kuo et al., 2011)	FoodMart2000 (Sales_Fact_1999 Data Table)
(Kuo and Shih, 2007)	Data from National Health Insurance Research Database provided by National Health Insurance, Department of Health and National Health Research Institutes
(Sarath and Ravi, 2013)	XYZ Bank Dataset
(Shenoy et al., 2005)	IBM_POS Dataset (Point Of Sale)
(Huang et al., 2010)	Ireland Telecom Data
(Wang and Huang, 2009)	German and Australian Credit Datasets
(Chan, 2008)	Empower Corporation Data(Historical Customer Data of Nine Cars)
(Chiu et al., 2009)	Large 3C Appliance Chain Store Data
(Jonker et al., 2004)	Dutch Charitable Organization Data
(Au et al., 2003)	Credit Card Database/Social Database/PBX Database
(Huang et al., 2012)	Ireland Telecom Data
(Gao et al., 2014)	Company B(Large Telecommunication Company in Australia)
(Cui et al., 2015)	Large Direct Marketing Dataset from U.S. Catalog Company provided by Direct Marketing Educational Foundation
(George et al., 2013)	Large U.S. Catalog Retailer(January 1997 to August 2004)
(Soltani Halvaiee and Akbari, 2014)	Brazilian Bank Data (Godi et. al.)
(Ghosh and Nath, 2004)	Market Basket Type Database
(Luna et al., 2013)	German Credit Dataset
(Yang et al., 2011)	German Credit Dataset
(Bhattacharyya, 1999)	Real Life Data
(Bandaru et al., 2015)	Vehicle Sales and Vehicle Service Data
(Guo et al., 2013)	Data Collected from Amazon.com and Cnet.com (Apex, Canon, Creative, Nikon, Nokia)
(Hansen et al., 2010)	Randomly Generated Datasets
(Corazza et al., 2014)	Major Bank in North Eastern Italy, the Bana Popolare di Vicenza
(Liu et al., 2008)	47,000 samples and 76 original Characteristics of Finance Enterprise
(Cai et al., 2009)	German Credit Dataset
(Dong et al., 2009)	German Credit Dataset
(Aliehyaei and Khan, 2014)	German and Australian Credit Datasets
(Waad et al., 2014)	German and Australian Credit Dataset/Tunisian Bank Data/ HMEQ Datasets
(Zhang, 2008)	Shoes Retailer Data
(Eiben et al., 1998)	Invertors Dataset
(Bozsik, 2010)	200 Company Data
(Ozçelik et al., 2010)	Real Life Data
(Assis et al., 2014)	Brazilian Electronic Payment Company Data
(Yamada and Terano, 2006)	One term and eight-term Stock Data
(Yamada and Ueda, 2005)	S&P 100, Nikkei225, Gold Silver Index, USD/JPY, GBP/JPY(1971–2003)/USD/JPY-1998,1999
(Huang et al., 2011)	Taiwan Economic Journal Co. Ltd. Database for period of 24 months from 2009 to 2010 (Taiwan Stock Exchange) Data
(Carvalho et al., 2014)	Stanford (STD)-based on Twitter tweets/Health Care Reform Data
(Hochreiter, 2015)	Dow Jones Industrial Average (DJIA)-2010 to end of 2013
(Chen et al., 2007)	Credit Card Data provided by Chinese Commercial Bank
(Jin et al., 2006)	Telecom Company Data
(Sotoodeh, 2012)	Autonomous System (AS) Dataset collected from 2004 to 2007
(Qiwan and Min, 2008)	Jiangsu VV Group Survey Data
(Basiri and Taghiyareh, 2012)	Data from Teradata Center at Duke University
(Soltani et al., 2012)	POS Registered Transactional Records
(Jungwon et al., 2003)	-Details Not Provided-
(Bhugra et al., 2013)	Market Basket Data (1000 records and 11 attributes)
(Khademolghorani, 2011)	Market Basket Data of a Super Market
(Birtolo et al., 2013)	Contoso BI Demo Dataset (Microsoft) and e-Commerce Platform of Poste Italiane Data
(Christian and Martin, 2010)	Automotive Dataset of UCI ml Repository. (The automotive dataset was first presented in research paper by Kibler et al. (1989)).
(Cui et al., 2010)	U.S. based Catalog Company Recent Promotion as well as Purchase History over 12 years
(Bhattacharyya, 2000)	Cellular Phone Data(Real World Data)
(Gupta et al., 2015)	Data from Domains of Restaurants and Laptops
(Kaiser et al., 2010)	German Online Community Gamestar.de Data
(Dehuri et al., 2008)	Data Randomly Generated and Compared with e-mail marketing company Optus Inc.
(Cunha and Castro, 2013)	Real World Commerce Database
(Mzoughia and Limam, 2014)	Data from Credit Card Transactions of Retail Bank in North Africa
(Cheng, 2005)	Data From Transaction Records of certain Chain Convenience Store in Shanghai(100 kinds of items in 15 days)
(Duman and Elikucuk, 2013)	Original Credit Card Fraud Data (22 Million records with 978 as Fraudulent)
(Faris et al., 2014)	Major Cellular Telecommunications Company in Jordan Dataset
(Wong et al., 2012)	Major Australian Bank Data
(Brabazon et al., 2010)	Data Provided by WebBiz
(Gadi et al., 2008)	Brazilian Card Issuer Data
(Huang et al., 2010)	Synthetic Data

Table 10 (*continued*)

Reference	*Dataset description*
(Ganghishetti and Vadlamani, 2014)	XYZ Bank Dataset
(Brun et al., 2009)	Real data from local electrical energy concessionaire
(Banati and Bajaj, 2012)	Reviews of 10 users of a digital camera Data
(Chien and Chen, 2010)	Data from Yahoo! Finance (10 companies of S&P 500 (2008))

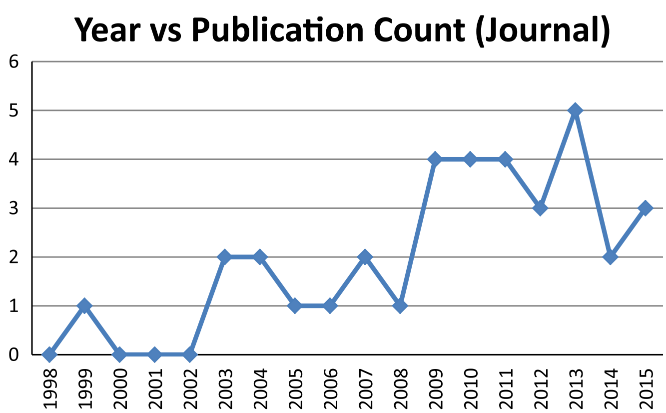


Fig. 9. Number of papers published in journals.

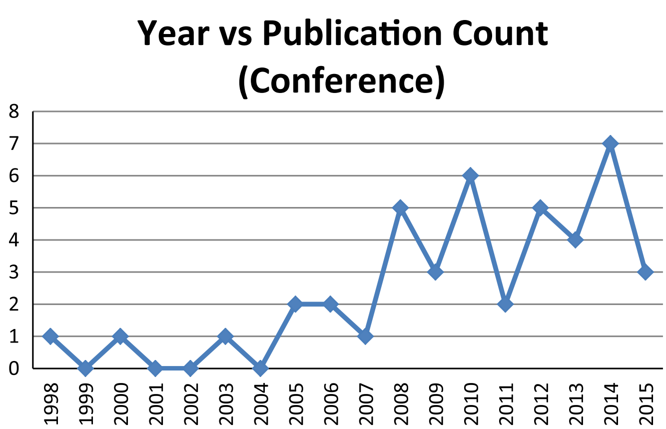


Fig. 10. Number of papers published in conferences.

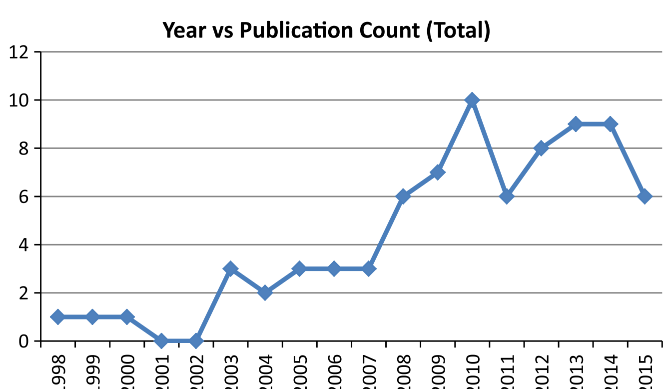


Fig. 11. Number of papers published in both journals and conferences.

study comprised the target sample (fraudulent companies) of 51 companies accused by Security and Exchange Commission of improperly recognizing revenue and peer class (non-fraudulent companies) of 339 companies. GA accurately classified 63% of target class companies and 95% of peer class companies. Research can be extended to other financial statement frauds. The techniques presented here can also be extended to incorporate other qualitative indicators and events that have additional discriminative power. Further, research could also test the stability of learned patterns on a selected sample of test patterns at a later time frame than the training sample. Moreover, prediction accuracy for fraudulent companies can be increased by resorting to multi-objective and hybrid EC techniques.

Fraud detection costs are minimized if the parameters are tuned with an EC technique before applying AIS. Gadi et al. (2008) applied AIS with parameters optimized using GA for credit card fraud detection and compared its performance with that of other DM techniques. Credit card data was provided by a Brazilian card issuer, and the results were consistent with earlier applied DM techniques. Regarding future directions, we can analyze the details of optimized parameters and choose a better metaheuristic for parameter tuning.

Electrical energy fraud can also be detected with the EC techniques. Brun et al. (2009) proposed the use of the DE-based algorithm to determine parameters to pre-select suspected fraudulent consumers. Real data from a local electrical energy concessionaire with a total of 14,273 consumers, of which 331 were fraudsters, was analyzed and the results obtained with the proposed approach were good. The system could be used as a control tool for inspections. However, success rate and coverage could be further improved.

A particular point to note is that problem of fraud is highly dynamic as fraudsters continually adapt. So, in some studies, AIS is preferred over other techniques. Brabazon et al. (2010) suggested the use of AIS for online credit card fraud detection. The data was provided by WebBiz, and the results suggest that AIS is good for fraud detection. The system obtained high accuracy but with high false negatives. Also, parameters can be tuned using GA, and a hybrid or multi-stage AIS detection (Negative Selection, Modified Negative Selection, and Clonal Selection) can also be considered.

The comparison of standard AIS techniques with a hybrid also needs to be performed. Huang et al. (2010) applied models of Negative Selection (NS), Danger theory (DT) and a hybrid, viz., NS and DT on CART for the detection of online break-in fraud in E-commerce. Using synthetic data hybrid performed well, and the DT model also gave good results. Further, portable and extensible fraud detection has to be developed.

Nowadays fraud analytical solutions that analyze customer behavior patterns (apart from technical solutions like cyber-security driven ones) are utilized for detecting fraud. Ozçelik et al. (2010) applied GA to improve credit card fraud detection utilizing real life data. The results are better than standard DM techniques with appropriately chosen parameter values.

Fraud detection methods are based on two approaches: 1) User behavior analysis and 2) fraud analysis. Soltani et al. (2012) presented AIS based credit card fraud detection called Artificial Immune Recognition System that considers user behavior. Data from POS registered transactional records was analyzed, and it was observed that the false positive rate decreased.

Wong et al. (2012) discovered credit card fraud by the AIS. Desensitised credit card transaction data obtained from a major Australian bank consisting of a total of 640,361 transactions for 21,746 credit cards. The proposed system achieved a detection rate of 71.3% on real transaction data. Many avenues of research are possible, such as the distribution of components of AIS, improvements to AIS, more sophisticated match rules, integration of AIS to transaction processing and even trying out other evolutionary techniques in order to improve the detection rate.

Research is underway for applying new meta-heuristics other than GA and GP for fraud detection. Duman and Elikucuk (2013) proposed an MBO algorithm to credit card fraud detection. MBO analyzed original credit card fraud data (22 million records with 978 as fraudulent ones) and yielded good performance compared to GA and standard DM techniques. As an extension, MBO can be further improved by testing alternative neighborhood functions and alternative benefits mechanisms.

In online payments, if a transaction is fraudulent, then the company absorbs the loss and consequently refunds the lost amount to the customer concerned, provided the customer is not the perpetrator of the fraud. Assis et al. (2014) identified fraud in online credit card transactions using GP on the Brazilian Electronic Payment Company Data. The algorithm achieved good performance in fraud detection. In future studies, multiple objectives should be considered, including rule complexity, higher accuracy, and additional datasets.

Fraud detection should have a high detection rate, low false alarm rate, and real-time responses. So, Soltani Halvaiee and Akbari (2014) detected credit card fraud using an artificial immune-based recognition system (AIRS). The Brazilian Bank Data was used for this paper. It had 3.74% fraudulent data of lost cards, stolen cards, skimming, mail orders, telephone orders, and account takeover. The accuracy increased up to 25% while cost was reduced up to 85% and response time decreased by 40%. Also, improvements can be made like weighting the features using feature selection methods in distance function, considering artificial immune networks for fraud detection (AFDM) using cloud computing.

4.4. Market Basket Analysis

Market Basket Analysis (MBA) is the analysis of items that are bought (or present) together by many customers in single, multiple or sequential transactions. Understanding the connections and the quality of those connections provides important knowledge that can be utilized to make suggestions, cross-sell, up-sell, offer coupons, and so forth.

MBA is accomplished by association rule mining in data mining, and in this subsection we will review the applications of GA, MOGA, Grammar Guided Genetic Programming (G3P), Genetic Network Programming (GNP), ACO, PSO, biogeography-based optimization (BBO), Imperialist Competitive Algorithm (ICA), Multi-Objective Binary Particle Swarm Optimization (MO-BPSO), Multi-Objective Binary Particle Swarm Optimization and Threshold Accepting (MO-BPSO-TA) and Multi-Objective Binary Firefly Optimization and Threshold Accepting (MO-BFFO-TA) and AIS applied to MBA.

Association rule mining can also be formulated as a multi-objective problem rather than the single-objective problem. Simultaneously considering more than one quality measure makes the association rule mining multi-objective problem. Ghosh and Nath (2004) considered measures of support count, comprehensibility, and interestingness used for evaluating a rule as objectives for Pareto-based MOGA. These are useful and interesting rules extracted from any market basket type database. Upon experimentation, the algorithm has been found to be suitable for large databases. A random sampling method has been applied, though; other sampling methods can also be applied. Apart from testing on numerical attributes, it can also be tested for categorical attributes.

Selecting items for cross-selling in chain convenience stores is considered NP-hard. Cheng (2005) applied GA for selecting items from transaction records of a certain chain convenience store in Shanghai, and the results obtained were optimal and parameter values adjusted according to convenience. Other meta-heuristics could be considered in place of GA, and the objective function can include support, confidence, lift, etc.

Dynamic datasets have immense potential for reflecting changes in customer behavior patterns. So Shenoy et al. (2005) analyzed a dynamic transaction database using GA and named the method Dynamic Mining of Association Rules using Genetic Algorithm (DMARG) to generate large item sets. Intra, Inter, and Distributed Transactions were considered. The IBM point of sale dataset was used. DMARG outperformed Fast Update (FUP) and E-Apriori. Computational costs are low because of single scan feature. DMARG can handle identical and skewed databases.

Constraint-Based Mining (CBM) concentrates on mining item sets that are important. Kuo and Shih (2007) employed ACO to mine an extensive database to find the association rules efficiently. Considering multi-dimensional constraints was observed to be more efficient. The National Health Insurance Research Database was considered, and ACO yielded better results than Apriori and consumed less computational time. Performing feature selection improves the quality of the rules.

Associative Classifiers are classifiers based on associative classification rules and are more accurate than traditional classification (Thabtah, 2007). However, they cannot handle numerical data. Chien and Chen (2010) built an associative classifier to discover trading rules from a GA-based algorithm on numerical data. Data from ten companies of the S&P 500 from Yahoo! Finance from the year 2008 was analyzed. The proposed model has high prediction accuracy and is highly competitive compared to the data distribution method. Further, the proposed model (GA-ACR) could be extended to more relations between numerical data and can be utilized to deal with portfolio investment problems.

Christian and Martin (2010) generated association rules using Apriori, on the reduced search space generated by GA on the Automotive Dataset of UCI Machine Learning Repository. The execution time decreased as GA was used to reduce the search space. Further work could be done by employing ACO, SA, etc. Also, the consideration of negative associations as well as the development of a distributed version of proposed work could also be incorporated.

Retail shelf-space management is one of the most difficult aspects of retailing decisions as shelf-space is limited. Hansen et al. (2010) considered the retail shelf allocation problem with nonlinear profit functions, vertical and horizontal location effects and product cross elasticity. Heuristic, modified heuristic and GA models were applied. Randomly generated small and large datasets were considered. The meta-heuristic approach yielded best results. Apart from the three configurations, other configurations can also be considered. Other advanced evolutionary algorithms can also be employed in future.

Mining interesting and understandable rules without specifying minimum support and confidence and without frequent item sets is hard. Khademolghorani (2011) proposed ICA for mining interesting and comprehensible association rules on Market Basket Data of a Super Market. Good results were obtained. The generalization of this method to both numerical and categorical rules considering both positive and negative rules is a future research direction.

Determining appropriate thresholds for support and confidence is a research topic in itself. For this purpose, any EC technique can be used. Kuo et al. (2011) proposed an algorithm that mines association rules with better computational efficiency and determines threshold values. PSO finds optimum values of fitness for each particle and finds minimum limits of support and confidence converting data into binary values. The proposed method applied to the FoodMart2000 database yielded good results. The results provided a good reference for marketing strategy. In this case, variants of PSO can be applied to yield possibly better results to the above problem.

Ranking of rules and classifying rules as good or bad is important after generating association rules. Yang et al. (2011) suggested the evolutionary associative classification method for both adjustments of the order of rules as well as the refinement of each single rule. The association rules are re-ranked with respect to the equations which combine support and confidence values. GNP is used to search the equation space for prior knowledge. Apart from global ranking, the local adjustment is made by giving some rewards to the right rules and penalty to bad rules. The German Credit dataset was analyzed and the proposed method improves classification accuracies.

BBO has a way of sharing information between solutions to get more accurate results. Bhugra et al. (2013) generated association rules and optimized rules using BBO on Market Basket Data. The results obtained by the proposed model are accurate. Also, the model can attain a profitable and optimized performance with minimum support and number of items.

Managing product bundles is crucial as it satisfies users, minimizes dead stocks and maximizes net income. Birtolo et al. (2013) searched for product bundles that are optimal using GA on Contoso BI Demo Dataset (Microsoft) and e-Commerce Platform of Poste Italiane data. Experimental results indicated that GA was able to find optimal product bundles. Investigation of the power of other evolutionary algorithms is needed.

Artificial Immune-related algorithms can also be applied for association rule mining for getting diverse solutions. Cunha and Castro (2013) built association rules by evolutionary and AIS algorithms on Real World Commerce Database and presented good association rules in terms of computational time. Elaborate experimentation on feature selection is yet to be carried out. Further, its impact on the performance of the model is yet to be studied.

Luna et al. (2013) extracted association rules using G3P models that enabled extraction of both numerical and nominal association rules in a single step. Proposals combine G3P with Non-dominated Sort Genetic Algorithm (NSGA-II) and Strength Pareto Evolutionary Algorithm (SPEA-2). Credit dataset was used, and a good trade-off between support and lift was obtained by the proposed multi-objective methods. Proposed methods work on any domain whether categorical or nominal. They also restrict the search space using grammar.

Sarath and Ravi (2013) applied PSO to mine M best association rules without having to specify thresholds on support and confidence. Unlike the Apriori and FP-growth algorithm, this algorithm generates the best M non-redundant rules from a given bank dataset. The quality of rules is assessed by a fitness function viz., the product of support and confidence. This algorithm can be used as an alternative to the Apriori and FP-growth algorithm. Future strategies include the usage of better fitness function.

Hybrid Multi-Objective techniques can also be applied to rule generation. Ganghishetti and Vadlamani (2014) developed three multi-objective evolutionary association rule miners, namely MO-BPSO, MO-BPSO-TA, and MO-BFFO-TA on XYZ Bank datasets and concluded that MO-BPSO-TA outperformed other models proposed. All three consumed less time compared to standard association rule mining algorithms and do not generate redundant rules.

4.5. Customer Life Time Value

Customer Lifetime Value (CLV) is a forecast of all the worth a business will get from their whole association with a customer right from the time he/she started doing business until the time he/she attrite. Since we do not know exactly the extent of each relationship with any customer, we would like to make a decent estimate and state CLV by a periodic estimate. Not much work is reported in this important aspect of CRM, which has tremendous significance in growing the business.

References Chan (2008), Mzoughia and Limam (2014) and Zhang (2008) also discuss the estimation of CLV along with other CRM tasks with the help of various EC techniques and these are reviewed in other subsections. Hence, in this subsection, we review the work where only CLV is estimated with an EC technique, i.e., GA.

It is common for retailers to mail multiple catalogs promoting different product categories. George et al. (2013) addressed the problem of multi-category catalog mailing by a retailer. The multivariate proportional hazard model (MVPHM), regression-based purchase amount type in a hierarchical Bayesian framework and control function (CF) approach with GA have been employed to suggest the catalog mailing policy. Large U.S. Catalog retailer data from January 1997 to August 2004 was used, and the proposed methodology was able to generate 38.4% more CLV. Future research directions include the estimation of CLV based on the data of recency, frequency and monetary aspects (RFM) of a customer relation by suitably employing an EC technique.

4.6. Customer churn prediction

The ability to foresee that a specific customer is about to churn, while there is still time to retain him/her results in a potential extra income for every business. Other than the immediate loss of income that results from a customer defecting to the competition, the customer acquisition cost is a lost investment. Further, it is consistently more expensive to acquire a customer than to retain an existing one. Therefore, customer churn detection becomes a quite an essential task in CRM with significant revenue ramifications. Using data and text mining techniques, a service firm can predict which set of customers has a high probability to churn. The activity of customer retention involves marketing teams reaching out to the list of identified dissatisfied customers (obtained through churn models) in order to retain them.

In this subsection, we review the works involving GP, interactive GA, ACO, PSO, Beehive (a variant of Bee Swarm), NSGA-II, data mining by evolutionary learning (DMEL) and ICA as applied to Churn detection.

Eiben et al. (1998) employed and evaluated the GP model on the problem of customer retention modeling using the investors' dataset of a financial company. Other data analysis techniques like the rough set, CHAID, and logistic regression were also applied independently, but GP outperformed all of the others. EC techniques other than GP are yet to be studied. Developing models for other time periods is an ongoing research.

Au et al. (2003) proposed a novel evolutionary data mining algorithm called DMEL for churn prediction. Some databases like Credit Card Database, Social Database, and PBX Database were used. Experiments with different data sets revealed that DMEL effectively discovered interesting and robust classification rules. Feature selection can be considered as an important future research direction here.

Service marketing is related to retaining loyal customers and preventing customer churn (Qiwan and Min, 2008). They applied IGA to solve a dual optimization model using service marketing and customer churn on Jiangsu VV Group Survey Data. They claimed that, compared to NSGA-II, IGA was more suitable for practical problems. Adding influence coefficients of marketing like people, process, and physical evidence on customer churn is an issue in future research.

Customer churn is also a considerable problem in the telecommunication industry. Huang et al. (2010) proposed a multi-objective feature selection approach for churn prediction in the telecommunication service field using NSGA-II. Ireland Telecom data was used for this proposed model of churn prediction, and the experimental results of feature selection method were efficient for churn prediction. High computational complexity of GA is an issue to be resolved, and other sampling techniques of the dataset should be considered in future studies.

Prediction of churn using customer product buying patterns is a new approach. Subramaniam and Thangavelu (2011) developed the beehive-based approach using customer buying patterns for different products to find the customer churn. Data was collected by questionnaire survey from selling departments for a period of four months of marketing firms. The model outperforms traditional classification and the GA approach. Customer buying patterns are important to predict churn with the proposed method. However, acquiring accurate buying patterns is difficult and is a topic for future research. The sequential data of customer transaction history can also be mined using EC techniques for predicting churn.

Basiri and Taghiyareh (2012) developed the Colonial competitive Rule-based classifier (CORER) based on ICA to enhance the accuracy of prediction of churn. Data from Teradata Center at Duke University was analyzed using the model, which yielded better accuracy than Local Linear Model Trees (LOLIMOT), Artificial Neural Networks and Decision trees. Rules for CORER classifier can be constructed randomly. Also, an ensemble of CORER with other techniques is worth investigating.

Comparison of various DM techniques and DMEL for customer churn prediction in the telecom industry is necessary. Huang et al. (2012) utilized DMEL of Au et al. (2003) for customer churn prediction in telecommunications along with other classification algorithms. Ireland's telecom data was analyzed, and the results indicate that DMEL has poor performance for large datasets. To improve the performance, feature selection can be employed, and effective sampling techniques for dealing with an imbalanced dataset can be considered in future work.

Social networks are an important source of information for building churn prediction models. Sotoodeh (2012) measured churn in social networks with the Swarm Intelligence Algorithm inspired by ants' foraging behavior and analyzed Autonomous System (AS) Dataset collected from 2004 to 2007. The communications used in the simulation result from mining a dataset including real communications and results obtained confirm the model. Topology and scalability changes can also be accommodated in future research, and other advanced techniques for churn prediction can also be considered.

Faris et al. (2014) applied self-organizing feature map (SOM), for outlier elimination, and then later used GP to build a classification tree employing dataset from a major cellular telecommunication company in Jordan. The proposed method demonstrated a better classification of the above dataset. In the future, other similar methods can be employed for churn prediction.

Gaming relation (Bi-level) exists between a company and its customers. Gao et al. (2014) applied a bi-level decision model with PSO for customer churn analysis. Data from the Company B, a large telecommunication company in Australia was analyzed for the above-proposed model. Bi-level decision model solution approach provides reasonably adequate solutions for designing service plan features for the purpose of reducing customer churn rate. Besides churn prediction, this model can be applied to other problems. Further, dealing with bi-level problems that do not have a mathematical form is a future area of research.

Customer satisfaction has an implication on customer churn. So Bandaru et al. (2015) developed a quantitative method for assessing customer satisfaction that, in turn, was used for churn prediction using NSGA-II. Vehicle sales and vehicle service data are used. The results show the usefulness of after-sales and warranty data in predicting churn accurately. As the proposed method is generic and novel, it can be applied to other consumer products.

4.7. Default prediction

The default detection models make use of the present and historical data of a customer to forecast the customer's capacity to pay back on time. An accurate default detection system is essential for the company's profitability. Here, one should realize that both retail and corporate customers can default. Very few works appeared in this area.

In this subsection, we review the works on default detection using an EC technique, i.e., GA. Surprisingly, not much work is reported in this important area of CRM.

Classification rules can be set by using GA, just like discriminative analysis for default forecasting. Bozsik (2010) developed a GA-based default forecasting model on data from 200 companies and compared it with the models currently in use. The results obtained were reliable and achieved a classification accuracy of 84.5%. A further bigger sample of the dataset can be used in the model. Hybrid EC techniques can be used, whereas formulating the problem in a multi-objective framework is also an important future research direction.

It is important to pick up the signs of financial distress on time to evaluate firms that are going to default. Gordini (2013) proposed a GA-based approach for default prediction of small enterprises. Data from defaulting 6200 firms of Italy by the end of 2009 was analyzed. Results obtained are the best in terms of predicting defaulting firms. Models other than GA that capture the non-linear relationship and are capable of extracting rules can also be used.

4.8. Direct marketing

Data analysis in direct marketing identifies the most promising individuals to mail to and maximize the returns. Direct Marketing does not include advertisements set on the web, TV, or the radio. Types of direct marketing materials include catalogs, mailers, and fliers. Here in this subsection, we review the works reported in direct marketing using EC techniques, i.e., GA, MOGA.

Bhattacharyya (1999) applied a GA for Direct Marketing performance modeling based on the depth of mailing. Real Life data was utilized. Results demonstrate that the method performed well. Given large volumes of data in direct marketing, parallel implementation is also possible. Also, a better fitness function has to be formulated.

As an extension, EC techniques other than GA can be employed in the future. Further, the problem may be formulated as a multi-objective optimization problem. Bhattacharyya (2000) proposed a MOGA in combination with DMAX approach for direct marketing up to specified file depth on real-world data from the cellular phone. The results obtained are beneficial for linear and non-linear models. Further research considers over-fitting of MOGA and tools for visualization of various tradeoffs amongst multiple objectives.

The customer base can be segmented for marketing. This includes the problems of forming and distinguishing the actions taken towards different segments. Jonker et al. (2004) applied a GA for segmentation and direct marketing. Dutch charitable organization data was analyzed. The results outperformed CHAID segmentation. Further segmentation could be done by combining the score of RFM dimensions into a variable and using it to segment rather than the break points. Also, other EC techniques can be tried out.

Customer profitability should also be considered along with response probabilities. Thus, Cui et al. (2010) applied constrained optimization using a GA for direct marketing to maximize profitability on the dataset of recent promotion as well as purchased history over 12 years taken from a U.S.-based catalog company. The proposed model outperformed the unconstrained model and decile-maximizing (DMAX) model in terms of higher profitability.

Direct Marketing identifies targets that are from the top percentage of customers who are most likely to respond and purchase a greater amount. Cui et al. (2015) used partial order constrained optimization with penalty weight and GA as a tool for stochastic optimization to select the total sales at the top deciles of a customer list. The model analyzed large direct marketing dataset from U.S. Catalog Company of Direct Marketing Educational Foundation. The results outperformed those of other competing methods. The limitation is that this method is a partial order method. Further, other EC techniques can be used as an alternative; more experimentation can be carried out with other datasets, and the multi-objective optimization framework can be attempted.

4.9. Customer campaigning

Most organizations run dozens to several hundreds of campaigns or sub-campaigns, where sometimes all of them must be managed simultaneously. Campaigns are of various types. There are campaigns to pull in new customers, campaigns to sustain leads until they get to be sales-ready, campaigns to support qualified prospects through the purchasing cycle, campaigns to target special customers, and campaigns to up-sell and cross-sell existing customers.

In this subsection, we review the only paper available in the literature involving the honey bee behavior (HBB) algorithm (a variant of Bee Swarm) applied to customer campaigning. Dehuri et al. (2008) developed the HBB multi-agent approach for multiple campaign assignment problems on data randomly generated and compared the results with the preference matrix of the email marketing company Optus Inc. The results show a clear edge over random and independent methods created using artificially constructed customer campaign preference matrix. This model can be made more robust by combining other EC techniques with HBB. There is an immense scope to apply other EC techniques for feature selection and association rule mining for cross-sell. This is a excellent research opportunity.

4.10. Sentiment analysis

Sentiment analysis, also known as opinion mining, is the procedure of identifying the relevant polarity of opinions from a set of documents called customer reviews on products and services. Of late, it has gained prominence because of e-word of mouth, using which people typically buy a product/service based on the opinions/reviews expressed online by other customers of that product/service.

The relation between time scales and time series properties are important for a speculator to react based on investor sentiment. Yamada and Ueda (2005) developed a genetic learning model of investor sentiment. The model revealed the time series properties of S&P 100, Nikkei225, Gold Silver Index, USD/JPY, GBP/JPY(1971–2003), and USD/JPY-1998 and1999 data. Results describe the quality and amount of information to determine time series properties. However, the proposed model to the accepted time scale must be adequate for the speculator to react.

Possibilities for the use of GA for describing investor sentiment are a research area by itself. Yamada and Terano (2006) studied the application of GA describing investor sentiment and time series analysis on one-term and eight-term stock data, and results reveal dynamics reported in the early studies. The investors react to the piece of information and adjust their view quickly based on the proposed model.

Marketing actions can be based on the opinions obtained from social media. Kaiser et al. (2010) simulated the spreading of opinions within online social networks using the ant-based algorithm. Data from the German online community Gamestar.de was analyzed, and the model produced encouraging results. More experiments are planned by applying further social network communities to the model and advice can be taken from opinion leaders for improving the results.

Stock scoring models have been developed based on investor sentiment. Huang et al. (2011) devised a stock selection model based on investor sentiment using GA. Data from Taiwan Economic Journal Co. Ltd. Database for the period of twenty-four months from 2009 to 2010 (from Taiwan Stock Exchange) was used, and the proposed model surpassed existing benchmark results. In the future, other advanced techniques apart from GA can be employed by including recent-past data on stocks.

E-marketing campaign is also carried out by opinion mining using EC techniques. Banati and Bajaj (2012) extracted features from the opinions expressed online using Firefly Algorithm. They also performed market segmentation using the Firefly Algorithm and, in the end, promoted the product. Data from reviews of 10 users of a digital camera was analyzed, and the proposed approach is capable of attracting large web users by a small advertising budget.

GA can also be integrated into Sentiment analysis for certain optimization tasks. Guo et al. (2013) designed a multidimensional sentence modeling algorithm (MSMA), and GA was used for optimizing the weighting scheme while determining feature importance. The data collected from Amazon.com and Cnet.com of Apex, Canon, Creative, Nikon and Nokia was used. Experiments on the data were promising and showed a significant improvement over past studies in terms of precision, recall, and f-measure. Other aspects of sentence position, discourse relation and the semantic factors are involved in the feature extraction process on which future research is necessary.

Paradigm words are important for the polarity of text to be classified as positive, negative or neutral. Carvalho et al. (2014) proposed a statistical method where paradigm words are selected using GA on Stanford (STD)-based Tweets and Health Care Reform Data. The devised method not only outperformed others. but was also more flexible and demonstrated variation of the paradigm words according to the domain. Further, we can investigate the effect of different genetic operators for final classification.

There is a phenomenal growth of e-commerce reviews for products or services. Gupta et al. (2015) performed feature selection using PSO for aspect extraction and sentiment classification. Data from the domain of restaurants and laptops was analyzed, and proposed methods yielded promising accuracies. Further, more features can be included into account classifiers, and multi-objective PSO can also be considered for solving the above problem.

Hochreiter (2015) applied EC techniques for computing optimal rule-based trading strategies on sentiment data from the Dow Jones Industrial Average (DJIA) from 2010 to the end of 2013. The portfolio obtained outperformed classical strategies. When transaction costs were included, the proposed evolutionary strategy performed poorly. Apart from GA, we can also utilize GP and other indices data.

Table 11 presents the details of classification problems which use accuracy, sensitivity, specificity as the measures of evaluation as well as whether the rules are generated or not. Sensitivity is called true positive rate or recall. Specificity called true negative rate.

Table 12 describes target marketing utilizing lift. Lift here is the ratio of response in the target group over response in the random group. The higher this ratio, the better the rate of prospects converting into customers.

Table 11
Results of classification problems.

Reference	Accuracy (EC method)	Sensitivity (EC method)	Specificity (EC method)	Compared (standard DM technique)	Dataset type	Benchmark dataset	Rule generated (mentioned in the article)
(Kožený, 2015)	75.85%	85.28%	53.85%	–	Bank	No	No
(Cai et al., 2009)	–	91.18%	56.28%	–	Credit scoring	Yes	No
(Ong et al., 2005)	88.27%, 77.34% (GP) (Australian, German)	–	–	MLP* (87.93, 75.51), CART* (85.81, 7059), C4.5 (87.06, 73.17), Rough Sets (83.72, 74.57), LR* (86.19, 75.40)	Credit scoring	Yes	Yes
(Huang et al., 2006)	79.49% (2SGP)	–	–	MLP* (75.51), CART* (70.59), C4.5 (73.17), Rough sets (74.57), LR* (75.40)	Credit scoring	Yes	Yes
(Abdou, 2009)	83.28%, 85.82% (GP_p, GP_t)	91.89%, 91.07% (GP_p, GP_t)	65.25%, 74.94% (GP_p, GP_t)	WOE_a* (54.99, 98.30, 34.08) WOE_1* (54.44, 98.30, 33.25), PA* (81.93, 67.40, 88.95), PA_1* (81.62, 67.15, 88.60)	Bank	No	No
(Chen et al., 2007)	89.83% (MGP), 88.73% (Normalized Inputs GP (NGP))	–	–	86.76% (Back Propagation Neural Network (BPN))	Bank	No	Yes
(Wang and Huang, 2009)	100% (MOEA with LDA*)	–	–	92.5%(Chi-Square), 99.3%(EAFS*), 92.0%(None), 99.4%(Relief) With LDA*	Credit scoring (Australian)	Yes	No
(Dong et al., 2009)	72.9% (SAREA)	91.57% (SAREA)	29.3% (SAREA)	LDA* (72.1%, 73.3%, 68.75%) QDA* (65.9%, 64.9%, 67.5%) KNN* (71.6%, 98.0%, 7.66%) C4.5 (70.4%, 78.6%, 49.9%)	Credit scoring	Yes	Yes
(Aliehyaei and Khan, 2014)	77.5%	–	–	–	Credit scoring	Yes	Yes
(Duman and Elikucuk, 2013)	–	82.78% (Average) (MBO)	–	82.73% (Average) (GASS*)	Bank	No	No
(Soltani et al., 2012)	–	91%, 100%	33%, 13%	–	Retail	No	No
(Brabazon et al., 2010)	90.14%	96.55% Max. of (Sen.+Spec)/2	9.35% Max. of (Sen.+Spec)/2	–	Bank	No	No
(Yang et al., 2011)	73.56% (Rank+), 73.10% (Rank)	–	–	70.90% (FOIL*), 69.10% (CPAR*), 72.00% (CMAR*)	Credit scoring	Yes	Yes
(Au et al., 2003)	95.67%, 99.9% (DMEL)	–	–	82.6%, 94.6% (C4.5), 58.9%, 58.9% (Michigan Style (SCS*)), 59.6%, 94.2% (Pittsburgh Style (GABL*))	Bank, company	Yes	Yes
(Gordini, 2013)	76.9% (GA)	–	–	71.3% (MDA*), 72.1%(LR*)	Company	No	Yes
(Bozsik, 2010)	84.5%	–	–	–	Company	No	Yes

*MLP=Multi-layer perceptron.
*CART=Classification and regression tree.
*LR=Logistic regression.
*FOIL=First order inductive learner.
*CPAR=Classification based on predictive association rule.
*CMAR=Classification based on multiple-class association rules.
*KNN=K-nearest neighbor.
*GASS=Genetic algorithm hybridized with scatter search.
*WOE=Weight of evidence.
*PA=Probit analysis.
*EAFS=Evolutionary algorithm for feature selection.
*LDA=Linear discriminant analysis.
*QDA=Quadratic discriminant analysis.
*MDA=Multiple discriminant analysis.
*SCS=Simple classifier system.
*GABL=GA batch concept learner.

Table 12
Results of direct marketing with lift measures.

References	Decile (start) (lift type)	Decile (end) (lift type)
(Cui et al., 2015) (10%)	630.8 (1) (profit lift), 374 (1) (response lift)	356.2 (2) (profit lift), 274 (2) (response lift)
(Cui et al., 2015) (40%)	600.0 (1) (profit lift), 347.7 (1) (response lift)	385.1 (2) (profit lift), 259.6 (1) (response lift)
(Bhattacharyya, 1999)	284.3 (1)	125.7 (7)
(Cui et al., 2010)	629 (1) (profit lift), 309 (1) (response lift)	100 (10) (profit lift), 100 (10) (response lift)
(Bhattacharyya, 2000)	304.9 (1) (churn lift), 261.7 (1) ($-lift)	138.8 (7) (churn lift), 126.9 (7) ($-lift)

Table 13 presents the performance of association rule mining using various measures of evaluation like support, confidence, comprehensibility, interestingness, and the number of rules considered for a mining task.

5. Discussion of the review and future directions

While EC techniques have made steady inroads into the area of analytical CRM, they have not done as much as they have in other fields of application. Our survey indicates that analytical CRM still offers fertile ground, where EC techniques either in isolation or in combination have a tremendous role to play. In this context, we listed some important open research problems and future directions, as follows:

- The most striking observation from the review is that EC techniques turn out to be viable alternatives for classical DM techniques like Neural Networks, Support Vector Machines, Decision Trees, K-means, etc. as sometimes they offer better results in less time.
- As the trend of applying EC techniques to solve CRM tasks has been increasing from 1998 to 2015, it will be a booming area of research in future.
- EC techniques solve these tasks by generating 'if-then' rules from the underlying data. These rules can be considered as early warning expert systems in tasks such as fraud detection, churn detection, default detection, credit scoring, etc. In this sense, EC techniques, are termed as 'transparent systems' as opposed to black boxes like logistic regression, MLP, SVM, etc. Decision trees, rule sets, rough set based methods, fuzzy logic based methods, fall in the category of transparent systems. However, in many of the papers reviewed here, quite surprisingly, these rules are not at all presented making it difficult to judge the value of the contributions.
- It is conspicuous from the survey that for some reason, GA has been employed extensively for solving CRM tasks. Apart from GA, other EC techniques that are proved to be better in dealing with multi-modal, non-linear, and non-differential objective functions and non-convex search spaces can also be employed. For instance, DE, which outperformed GA in many problems in other domains, can be employed to investigate its potential for solving CRM tasks. Likewise, the potential of other EC techniques can also be explored.
- Further, hybrid EC techniques have immense potential in solving these tasks in less time with higher accuracies. However, the challenge is the correct choice of constituent EC techniques and the suitable design of the hybrid.
- Many CRM tasks can be formulated as multi-objective optimization problems by suitably identifying the multiple objectives. Then, Evolutionary Multi-objective optimization techniques can be employed leading to innovation in decision making.
- Real-world data sets rather than synthetic data should be analyzed by the EC techniques so that the practicing community gets convinced about the utility of these techniques. Further, benchmark datasets should be used.
- Customer Life Time Value (LTV), Default detection, and Customer Campaigning have a tremendous scope for further research involving the application of EC techniques, as not much work is reported in this area. CLV estimation can be posed as a regression problem and solved using EC.
- In future, EC techniques can be applied to solve new CRM tasks such as sentiment classification and class association rule mining for sentiment analysis, fraud detection, etc.
- Fraud Detection is of interest to many financial institutions because the world is becoming more and more technology-dependent for performing financial transactions and services. Further, the problem becomes more complicated because of inherent imbalance that the datasets are replete with. In other words, the companies have a disproportionately small number of fraudulent transactions than the number of genuine transactions. Therefore, it is very difficult to detect fraud without resorting to data balancing techniques. While one-class classification is the ideal way to solve this problem, the application of EC techniques in building one-class classifiers is worth investigating. Another direction is the application of data balancing before invoking EC techniques for binary classification.
- Further, MBA being an important task using which one can grow one's business, EC techniques-driven class association rule mining will yield insights as to what attributes together lead to successful cross-sell and thereby build better recommendation engines. These are alternative to traditional association rules.
- It is found that high utility itemset mining, generates high amount of return on investment as frequent itemsets are not always interesting and sometimes rare itemsets yield high profit (Krishnamoorthy, 2015; Lin et al., 2011; Tseng et al., 2015, 2016). In this connection, we observe that evolutionary algorithms find an interesting application in this area as well, with utility as the objective function.
- Since, quantum-inspired EC techniques are gaining popularity, their potential in solving CRM tasks is worth investigating.
- Credit recovery analytics involves the use of predictive analytics to identify those defaulted customers, who have the propensity to repay the loans but failed to do so for some valid and legitimate reasons. This is one problem that is typically formulated as a classification problem and solved using data mining and text mining techniques (Ravi et al. 2015), but the application of EC techniques is conspicuously absent. This problem can also be explored in future.
- Operational Risk pertains to all other risks different from credit and market risk. It is the risk emanating from people, process, and technology. Predicting the financial loss due to operational risk collapses into a regression task, wherein the potential loss from an untoward event becomes the target variable, whereas the data on losses from past events, event triggers, etc. become predictor variables. These events include fraudulent online transactions, cyber attacks, etc. Banks started collecting data associated with operational risk. This regression problem can also be solved using EC techniques.
- We did not consider pure and hybrid soft computing techniques for the review. However, EC techniques have a tremendous role to play in that direction too. Since many analytical CRM tasks are binary classification problems, the issues of wrapper-based feature subset selection, classification/regression/association/clustering can be solved using two-tiered EC architectures.
- Finally, one caveat is in order. It is well known that user-defined parameters have to be tuned carefully for the EC techniques. It consumes time and effort to fine tune and identify the optimal

Table 13
Results of association rule mining (M.B.A.).

References	Support (EC method)	Confidence (EC method)	Comprehensibility (EC method)	Interestingness (EC method)	No. of rules (EC method)	Compared (standard DM technique)
(Cheng, 2005)	3%	85%	–	–	–	–
(Ghosh and Nath, 2004)	–	–	Yes	Yes	40	–
(Chien and Chen, 2010) (I) (Associative classification)	1%, 2%, 5%, 10%	60%, 70%, 80%, 90%	–	–	Top 3, 4, 5 rules taken with 90% Accuracy	–
(Chien and Chen, 2010) (II) (GA used)	–	–	–	–	Top 3 rules from population of 20 (GA used) with 85% Accuracy	Data Distribution Method (DDM)
(Luna et al., 2013)	70% (G3PARM)	90% (G3PARM)	–	–	Population Size of 50	–
(Yang et al., 2011)	1% (GNP parameters)	50% (GNP parameters)	–	–	Max. no, of rules is 80,000	FOIL*, CPAR*, CMAR*
(Kuo and Shih, 2007)	–	–	–	–	Pheromone Threshold of 15 to generate 10 rules	Apriori
(Sarath and Ravi, 2013)	1.5%	40%	–	–	10	Apriori, FP-growth*
(Bhugra et al., 2013)	1.5% to 0.6%	–	–	–	–	–
(Khademolghorani, 2011)	15.87%	38.24%	–	–	10	–
(Ganghishetti and Vadlamani, 2014) (MO-BPSO)	2.04%	54.25%	68.32%	20.95%	–	–
(Ganghishetti and Vadlamani, 2014) (MO-BPSO-TA)	2.15%	53.79%	67.94%	18.41%	–	–
(Ganghishetti and Vadlamani, 2014) (MO-BFFO-TA)	1.11%	28.62%	64.71%	6.93%	–	–

*FOIL = First order inductive learner.
*CPAR = Classification based on predictive association rule.
*CMAR = Classification based on multiple-class association rules.
*FP-Growth = Frequent pattern-growth.

Table A1
Evolutionary computation (EC) techniques with advantages and disadvantages.

S.No.	EC technique	Basic idea	Advantages	Disadvantages
1	**Genetic Algorithm** (Golberg, 1989)	A Genetic Algorithm (GA) is a search heuristic that mimics the Darwinian principle called natural selection within a population of candidate solutions to solve an optimization problem. There are four procedural steps in the implementation of GA. 1. Initialization 2. Selection 3. Genetic Operators 4. Termination	1. It can handle optimization problems with the variable length chromosome encoding and the average fitness of the population never decreases as the number of runs increase. 2. GA can solve multi-modal, multi-dimensional, non-differential, discontinuous, and non-linear problems.	1. GA consumes a lot of time in order to converge. 2. There is no proof that a GA will converge to the global optimum.
2	**Genetic Programming** (Koza, 1992)	Genetic Programming (GP) is an EC-based strategy to discover computer programs that can solve an optimization problem.	1. It needs little domain-dependent information. 2. It generates solutions that look like programs. 3. As with GA's, GP's computer programs are independent of the domain, where the problem is being solved.	1. Parameters are tuned by trial and error.
3	**Interactive Genetic Algorithm** (Takagi, 2001)	An interactive GA (IGA) is characterized as a GA that uses human intervention. It belongs to the family of Interactive EC (IEC) techniques. It is used when it is difficult to find a computational fitness function: for instance, developing pictures, music, different imaginative outlines, and structures to match a client's inclinations. Interactive techniques can utilize diverse representations, both straight (as in conventional GA) and tree-like ones (as in GP).	1. Interactive data can be used whenever fitness values cannot be calculated accurately with the available data. 2. IEC executions can simultaneously be used by numerous users.	1. IGA depends on humans to get quality assessments. 2. Moreover, human assessments are moderate and costly when compared to fitness value calculations.
4	**Genetic Network Programming** (Mabu et al., 2010)	Genetic Network Programming (GNP) is a development over GP. The first inspiration for creating GNP is for more broad representation capacity of graphs than that of trees. The objective is to create GNP, which can manage dynamic situations productively and successfully. The attributes of GNP are as follows: 1) GNP programs are made out of various nodes guided by the connections to one another. 2) The graph structure empowers GNP to re-use nodes. 3) The nodes of GNP are executed by node associations with no terminal node. Therefore, the past node moves are utilized in making the next move, and this is used as a memory function.	1. The point of interest of GNP is an automatic selection of the important number of inputs and actions depending upon the situations. 2. GNP can be applied to dynamic and complex problems.	1. One of the limitations of GNP is that its basic gene structure may turn out to be excessively complex after iterations. 2. Sometimes, the memory required in GNP to find the optimum will be of maximum capacity.
5	**Particle Swarm Optimization** (Kennedy and Eberhart, 1995)	Particle Swarm Optimization (PSO) is a population-based evolutionary optimization method. Here we have a population of particles moving around in the search space by iteratively updating the particle's position and speed using simple heuristics. Every particle's position is impacted by its local best-known position but at the same time is guided toward the global best-known positions in the search space. The above process is required to move the swarm toward the best positions.	1. The benefits of PSO are that PSO is simple to code, and there are very few parameters to tune. 2. PSO can be parallelised. 3. PSO is originally designed for continuous optimization problems. However, combinatorial version has also been reported.	1. The algorithm gets trapped at local optima in some cases 2. It is not very good at diversification
6	**Bee hive** (Karaboga and Akay, 2009)	The bee hive has four stages- dance, auditorium, dispatch and outside. Moving among the stages gives honey bees a current source. The honey bees share their sources in dance stage, and in the auditorium, honey bees see the dance experts. The dispatch room has sources by which honey bees can begin their quest for sustenance. Outside the hive is an extremely broad state, and the accurate conduct of the honey bee in this state is not characterized.	1. This algorithm can be applied to problems of search where the exploitation capacity enhancement and leveled improvement of search are needed. 2. A bee hive is fault tolerant, scalable, and depends totally on nearby, or territorial data.	1. For the most part, this is applied to routing problems, even though is can be applied to other problems 2. Information about the neighborhood is needed
7	**Honey bee** (Karaboga and Akay, 2009)	The honey bee algorithm is made of three fundamental exercises: exploration, recruitment and harvest. The honey bee goes from an inactive state to the exploration state in which the "scouts" fly out extensively to explore potential sources, and afterward return and move to select foragers.	1. The point of interest of utilizing the honey bee search is its strength against anomalies. 2. The issue of getting trapped in local optima is defeated and thus will be able to discover an optimal solution.	1. It utilizes an excess of tuneable parameters. 2. Parameter setting is hard.

		Once the investigation begins, the recruitment and harvest stages are introduced, and the entire cycle repeats, changed by the present necessity of the hive.		
8	**Ant Colony Optimization** (Colorni et al., 1991)	The ant colony optimization is discovering ideal ways of ants looking for food. To start with, when an ant discovers a food source, it leaves "markers" (pheromones) to indicate the way leading to the food. At this point when different ants 'smell' the markers, they follow the trial with a certain probability. In case they are successful, they then populate the way with their markers as they bring the food back. This process continues until the food source becomes empty. In the process, they also take the shortest path.	1. There is inherent parallelism. 2. Positive feedback represents the fast discovery of better solutions.	1. Probability with which ant moves changes by generation. 2. Exploration is experimental rather than based on theory. 3. The time for convergence is hard to determine, even though convergence is ensured.
9	**Ant Cluster Algorithm** (Handl and Meyer, 2002)	The Ant Cluster Algorithm begins with an initial stage, in which (i) all information items are haphazardly scattered around the space, (ii) Every agent arbitrarily gets one information item, and moves to new location in the space. (iii) Every operator is put at an arbitrary position on the network which then chooses whether to drop the information item at a particular position or not. This process is performed on all information items.	1. Points of interest are numerous in perspective, for example, self-association, adaptability, no need for earlier data, and decentralization.	1. Picking up likelihood and dropping likelihood of a group are the two user- characterized parameters, and it is difficult to determine user characterized parameters ahead of time to frame the group due to different properties of the dataset.
10	**Honey Bee Mating Optimization (HMBO)** (Karaboga and Akay, 2009)	A honey bee colony consists of a queen(s), drones, workers, and broods. The queen leaves the hive for mating with drones. Mating flight starts by giving energy to the ruler and flight stops as energy gradually decreases.	1. Its parameters are inspired by the biological ideas. 2. Performs well in problems having both discrete and continuous variables.	1. Numerous parameters are required to be estimated. 2. Gets trapped in local optima. Convergence to global optimum takes a long time
11	**Migrating Birds Optimization** (Duman et al., 2012)	A migrating birds optimization (MBO) meta-heuristic is very good at energy minimization. This begins with the first arrangement (with respect to the leader bird in the V flight formation) and advancing on the lines towards the end. Every solution attempts to be enhanced by its neighbor solutions.	1. Solutions can be obtained in parallel. 2. Between solutions, there is a benefit mechanism.	1. It is a neighborhood search process. 2. It is moderate in speed
12	**Simulated Annealing (SA)** (Kirkpatrick et al., 1983)	Annealing is known as a thermal process for getting low energy states of a metal in heat bath. The SA comprises the two stages 1. Expand the temperature of the heat bath to the greatest value at which the metal melts; 2. Deliberately lower the temperature of the heat bath until the particles cool down to the ground condition of the metal.	1. It can avoid local minima. 2. It measurably ensures an optimal solution. 3. It is simple to code.	1. You have to recognize what the energy state is for a specific problem. 2. It is hard to find optima, and it is moderate in speed if the objective function is hard to compute. 3. It performs a point based search.
13	**Imperialist Competitive Algorithm** (Atashpaz-Gargari and Lucas, 2007)	Countries are initially random solutions. The objective function decides the power of every country. Taking into account their power, a portion of the best countries get to be imperialists and begin taking control of other countries thereby structuring the empires. Two primary administrators are Assimilation and Revolution. Assimilation makes colonies get closer to the Imperialistic country in the space of socio-political attributes. Revolution achieves random sudden changes in the position of some countries in the search space. Imperialistic Competition is another phase of the algorithm. Every one of the empires attempts to win this competition and take ownership of colonies of different empires.	1. They will not be caught on the local minimum. 2. They can handle an objective function that is non-convex, discontinuous in a domain and can also come out of a few local minima that are strong in the search space.	1. They are infeasible with the higher dimensional search spaces and computation cost thereby increases.
14	**Biogeography Based Optimization** (Simon and Member, 2008)	Species move between "islands" using debris, the wind, flying, and swimming. Habitat Suitability Index (HSI): Some islands are more suitable for a living than others. Suitability Index Variables (SIVs): Habitability is controlled by precipitation, geography, differences in vegetation, temperature, and so on As habitat suitability moves forward: 1. The species count rises. 2. Emigration increases 3. Immigration decreases	1. It is an efficient meta-heuristic and doesn't take much computational time. 2. Great in finding solutions. that don't come to the end or die with each generation.	1. Limitations are that it is not suitable for higher dimensional spaces and a large number of infeasible solutions are generated. 2. Exploitation is done poorly, and best individuals from every generation are not selected.

Table A1 (*continued*)

S.No.	EC technique	Basic idea	Advantages	Disadvantages
15	**DMEL** (Au et al., 2003)	Data mining by evolutionary learning (DMEL), is used for classification with precision where every forecast is assessed. In performing its tasks, DMEL looks at the conceivable space of rules utilizing an evolutionary approach. 1) First it generates the first order rule set utilizing the probabilistic strategy, and then higher order rules are obtained, 2) finding the rules that are interesting, 3) the chromosome's fitness is characterized by the probability that the attribute values of a record can be effectively decided to utilize the encoding of rules and 4) the probability of predictions made is estimated by likelihood.	1. DMEL can adequately find interesting rules. 2. Better than non-EC algorithms as it generates rules.	1. It has poor prediction capability with large datasets.
16	**Artificial Immune System (AIS)** (Castro and Timmis, 2002; Dasgupta, 1993)	AIS depends on a computational framework inspired by the human immune structure. Algorithms are developed based on human immune structure qualities.	1. AISs are information-driven self-adaptive routines. 2. They are good at function approximation. 3. AISs are nonlinear models, which makes them adaptable in more complex search spaces.	1. Parameters become hard to tune when attempting to utilize them.
17	**Multi-Objective Genetic Algorithm** (Konak et al., 2006)	A Single Objective GA can be altered to locate an arrangement of multiple non-dominated solutions. The capacity of GA to find various areas of solutions for non-convex, discontinuous, and multi-modal search spaces. The crossover of GA may find good solutions regarding different objectives to make a new non-dominated solution in unexplored parts of the Pareto front. Moreover, most multi-objective GAs don't require them to organize, scale, or measure objectives. Accordingly, GA has been the most well-known heuristic way to deal with multi-objective optimization.	1. MOGA can be effectively implemented for combinatorial optimization problems. 2. Suitable if a spread of Pareto-optimal arrangements are needed in the objective space.	1. MOGA may be biased towards the shape of the Pareto-Optimal front and to the density of solutions. 2. It is slow to converge or not good to locate a decent spread in the Pareto-Optimal front.
18	**NSGA-II** (Deb et al., 2002)	Multi-Objective Evolutionary Algorithm (EA) that utilize non-dominated sorting and sharing have three main difficulties. 1) $O(MN^3)$ computational complexity where M is the number of objectives and N is the population size 2) non-elitist methodologies, and 3) the requirement for indicating a sharing parameter. Non-dominated sorting-based multi-objective EA (MOEA), called non-dominated sorting genetic algorithm II (NSGA-II), overcomes the above three challenges. In particular, there is a fast non-dominated sorting methodology with $O(MN^2)$ computational time. Likewise, a selection is exhibited that makes a mating pool by joining the parent and offspring populations and selecting the best.	1. Diversity is preserved. 2. NSGA-II has $O(MN^2)$ complexity 3. Elitism does not permit a discovered Pareto-Optimal solution to be discarded.	1. Crowding confines the convergence. 2. Non-dominated sorting is of the size 2N.
19	**SPEA-2** (Zitzler et al., 2001)	An enhanced variant, specifically SPEA-2, is proposed, which uses its ancestors: a fine-grained fitness assignment, a density estimation strategy, and an enhanced archive truncation. 1. An improved fitness assignment is considered for every individual based on the number of individuals it dominates and by which it is dominated. 2. The nearest neighbor density estimation strategy is fused with SPEA, which permits a more exact direction of the search process. 3. New archive truncation techniques ensure the conservation of boundary solutions.	1. It is easy to realize solutions in the obtained Pareto-Optimal front. 2. Clustering methodology is employed. SPEA-2 performs better than its previous versions on all problems.	1. SPEA presents an additional parameter M, the external population size. A harmony between the normal population size N and this external population size M is important in the effective working of the SPEA and generating Pareto-Optimal front.
20	**Tabu search** (Glover, 1989a, 1989b)	Tabu search meta-heuristic is a point-based local heuristic search method that searches the neighborhood optimum by utilizing a Tabu list. Essentially it is utilized to solve combinatorial optimization problems and is a dynamic neighborhood search technique. It uses an adaptable	1. Permits non-improving solutions to escape from local optima. 2. Can be used to both discrete and continuous spaces.	1. There are numerous parameters to be evaluated. 2. Runs could be vast. 3. Global Optimum may not be found

		memory to confine the subsequent solutions to a subset of the neighborhood of a current solution.		
21	**Grammar Guided Genetic Programming(G3P)** (McKay et al., 2010)	Grammar-guided genetic programming (GGGP) is an expansion of GP frameworks. The distinction lies in the way that they utilize context-free grammars (CFGs) that produce all the conceivable answers. For a given problem of sentences, build up the formal meaning of syntactic problem constraints and utilize the derivation trees for every sentence to encode these solutions.	1. It permits adaptable methods to find the solutions in the search space with restrictions.	1. The difficulty in utilizing GGGP is the necessity for defining the grammar.
22	**Threshold accepting** (Dueck and Scheurer, 1990)	It is a deterministic variant of simulated annealing. This is a point-based local search technique that begins with a random feasible solution and then investigates its neighborhood in the solution space by moving from its present position, accepting another solution if it is better than the existing one. TA wriggles out of local minima by accepting new solutions, if they are not much worse than the current solution.	1. It easily overcomes getting stuck at local optima. 2. It is easy to implement 3. It is computationally efficient as there is no need for probabilistic acceptance, unlike simulated annealing.	1. It has a very slow convergence as it is a local search technique.
23	**Differential Evolution** (Storn and Price, 1997)	Differential Evolution is a stochastic, parallel, direct search evolution strategy method. At every iteration, new solutions are created with the combination of solutions chosen from the present set of solutions called mutation. The solutions obtained at this stage are combined with a specific target solution in a process called crossover. At last, based on fitness, the solutions are selected.	1. Differential Evolution is capable of handling non-linear, non-differentiable, and multi-modal objective functions. 2. It is fairly fast and robust. 3. Though designed to solve problems with real values, it can be used in the case of discrete valued variables also.	1. The encoding scheme has to be defined if non-real-valued decision variables are present in a problem.
24	**Firefly Algorithm** (Yang, 2010)	Firefly Algorithm consists of the following steps: 1. All fireflies are attracted to other fireflies. 2. Attractiveness depends on brightness and distance. If there is no other firefly near to one, then it will move randomly. 3. The brightness of the firefly depends on the decision space.	1. There is an automatic subdivision of the whole population into subgroups. 2. There is an in built capability of dealing with multi-modal objective functions. 3. There is high ergodicity and diversity in the solutions.	1. It gets trapped into several local optima. 2. Parameters are fixed and are not dynamically changed. 3. The Firefly Algorithm does not memorize any history of the better solution.

Table A2
Objective functions, properties of objective function and parameters set for the EC techniques.

Reference	EC technique	Objective function used	(Single- or multi-objective)/(maximization or minimization)/(discrete or continuous)/(constrained or unconstrained)/(unimodal or multimodal)	Parameter values set
(Ong et al., 2005)	GP	Mean absolute error	Single/minimization/discrete/unconstrained/NM	Population size (PS)=40, Crossover rate (CR)=0.9, Mutation rate (MR)=0.01, Maximum generations (MG)=100
(Huang et al., 2006)	Two Stage - GP	Mean absolute error	Single/minimization/NM/unconstrained/NM	PS=100, CR=0.9, MR=0.01, MG=1000
(Kožený, 2015)	GA	Fitness function=Weighted bit mask and fitness score=Accuracy	Single/maximization/discrete/NM/NM	PS=200, CR=0.8, MR=0.15
(Fogarty, 2012)	GA	Scoring system maintenance function	NM/NM/NM/NM/NM	NM
(Abdou, 2009)	GP	$\alpha(\textit{Sum of square error})+\beta(\textit{Classification Error})$	Single/minimization/NM/unconstrained/NM	CR=50%, MR=95%
(Chiu and Kuo, 2010)	PSHBMO	Mean square error	Single/minimization/continuous/unconstrained/NM	NM
(Brusco et al., 2002)	SA	Identifiability and responsiveness	Multi (Bi-objective)/maximization/continuous/constrained/NM	Cooling factor=0.95, Temperature length=10* (Number of customers)
(Mahdavi et al., 2011)	GA	$\sum_k \textit{Weight}_k \times \sum_k \textit{Customer Covered}_k$ *where k is the catalog number*	Single/maximization/discrete/constrained/multimodal	NM
(Gordini, 2013)	GA	Accuracy	Single/maximization/NM/NM/multimodal	NM
(Subramaniam and Thangavelu, 2011)	Beehive	Quality of data	Single/maximization/discrete/NM/NM	NM
(Hoogs et al., 2007)	GA	Accuracy	Single/maximization/NM/NM/multimodal	NM
(Kuo et al., 2011)	PSO	Confidence$\times\log(\textit{Support} \times \textit{length} + 1)$	Single/maximization/discrete/constrained/NM	PS=5; 10; 20; 30; 40; 50
(Kuo and Shih, 2007)	ACO	Support and frequency	Single/maximization/discrete/constrained/multimodal	Pheromone threshold=15
(Sarath and Ravi, 2013)	BPSO	Support × confidence	Single/maximization/discrete/unconstrained/multimodal	NM
(Shenoy et al., 2005)	GA	$\frac{(\textit{No.of matching bits} - \textit{No.of nonmatching bits})}{\textit{Total No.of bits}}$	Single/maximization/discrete/NM/multimodal	NM
(B. Huang et al., 2010)	NSGA-II	1. Overall accuracy 2. Accuracy of true churn and 3. Accuracy of true non-churn	Multi/maximization/NM/unconstrained/multimodal	PS=300, MG=500
(Wang and Huang, 2009)	NSGA-II	1. Chi-square 2. Relative correlation 3. Relief	Multi/NM/NM/unconstrained/multimodal	PS=50, CR=0.9, MR=0.01, MG=100 and 1000
(Chan, 2008)	GA	LTV=Current value+potential value	Single/maximization/NM/constrained/NM	PS=100, CR=0.9, MR=0.5, MG=50
(Chiu et al., 2009)	IACA	Euclidean distance	Single/minimization/NM/NM/NM	NM
(Jonker et al., 2004)	GA	CLTV=Total discounted net revenue	Single/maximization/NM/NM/NM	PS=100, MR=0.1, MG=100
(Au et al., 2003)	DMEL	Probability of attribute values of the record that can be correctly determined by the rule.	Single/maximization/discrete/NM/multimodal	CR=0.25; 0.5; 0.75
(Huang et al., 2012)	DMEL	High true positive with low false positive	Single/maximization/discrete/NM/multimodal	PS=30, CR=60%, MR=0.1%, MG=300
(Gao et al., 2014)	PSO	Customer churn (follower) prevented by features provided by service provider (leader) so that 1. Minimum Cost 2. Better service	Multi (Bi-objective)/NM/NM/constrained/NM	NM
(Cui et al., 2015)	GA	M(e) is the money spent by customer 'e'	Single/maximization/NM/constrained/NM	PS=256, CR=1, MR=0.1, MG=2000, Tournament Size (TS)=10
(George et al., 2013)	GA	CLV of single customer	Single/maximization/NM/constrained/NM	NM
(Soltani Halvaiee and Akbari, 2014)	AIS	Fraud Score based on 1. Distance function and 2. Cost	Single/minimization/NM/NM/NM	AIS
(Ghosh and Nath, 2004)	MOGA	1. Comprehensibility and 2. Interestingness	Multi/NM/discrete/constrained/multimodal	PS=40, CR=0.8, MR=0.02, MG=100; 200; 300

(Luna et al., 2013)	G3P, G3P-NSGA-II, G3P-SPEA-2	1. Cartesian distance 2. Support 3. Confidence	Single and multi/NM/discrete/constrained/multimodal	PS=50, CR=90%, MR=15%, MG=0.1, Support=70%, Confidence=90%, Pareto Front Size=20
(Yang et al., 2011)	GNP	Accuracy of rules	Single/maximization/discrete/NM/NM	PS=5, CR=0.2, MR=0.1, MG=5, Support=0.01, Confidence=0.5, Maximum antecedent size=3, Maximum no. of rules=80000
(Bhattacharyya, 1999)	GA	Weighted sum of 1. Decile maximization and 2. Model is fitting well	Single/maximization/NM/constrained/NM	PS=50, CR=0.7, MR=0.2
(Bandaru et al., 2015)	NSGA-II, GA	1. Mean standard deviation and 2. Skewness if customer satisfaction index (CSI)	Single/minimization/continuous/constrained/NM	PS=2000, Binary CR=0.9, Single CR=0.9, Polynomial MR=0.05, Bitwise MR=0.15, MG=50 TS=2
(Guo et al., 2013)	GA	$W_1\times$(Sentence position)+$W_2\times$(Sentence resemblance in title)+$W_3\times$(Sentence inclusion of noun)+$W_4\times$(Sentence inclusion of numeric data)+$W_5\times$(Bushy path of node)+$W_6\times$(Summation of similarities of each node)	Single/maximization/continuous/NM/NM	NM
(Hansen et al., 2010)	GA	$\sum\sum P_{jk}\times X_{jk}$ Where P is profit of item and X is number of facings of item	Single/maximization/continuous/constrained/NM	PS=Shelves × Products, MG=5 × Shelves × -Products 0.2, CR=0.99, MR=0.03
(Corazza et al., 2014)	PSO	1. Spearman rank correlation coefficient and 2. Net flow	Single/minimization/continuous/unconstrained/NM	No. of particles=100; 200
(Liu et al., 2008)	GP	Information value	Single/maximization/discrete/unconstrained/NM	NM
(Cai et al., 2009)	GA	Error rates of judging rightly and wrongly good and bad customers	Single/maximization/discrete/unconstrained/NM	PS=1000, CR=0.5, MR=0.01, MG=1000
(Dong et al., 2009)	SA	Percentage correctly classified × similarity function × e (e is a coefficient)	Single/maximization/discrete/unconstrained/NM	Initial temperature=10000, Final temperature=0.01, Cooling rate=0.95, Iteration at each temperature=100
(Aliehyaei and Khan, 2014)	GP, ACO, GP+ACO	Mean absolute error	Single/minimization/nm/unconstrained/NM	NM
(Waad et al., 2014)	GA	$\sum_i Weight_i \times Distance(\sigma,L_i)$	Single/maximization/discrete/NM/NM	NM
(Zhang, 2008)	ACA	Errors in computing distances of clusters	Single/minimization/continuous/unconstrained/NM	NM
(Eiben et al., 1998)	GP	Coefficient of concordance (distinctive power of model)	Single/maximization//NM/NM/NM	CR=0.5, MR=0.9
(Bozsik, 2010)	GA	$Discrminant\ Function = \sum_i W_i \times x_i$ W_i is the weight and x_i is the attribute value	Single/maximization/continuous/unconstrained/NM	MR=20%
(Ozçelik et al., 2010)	GA	Correctly classified transactions	Single/maximization/NM/NM/NM	NM
(Assis et al., 2014)	GP	Accuracy × sensitivity × specificity	Single/maximization/NM/NM/NM	PS=500, CR=75%, MR=2.5%, MG=100
(Yamada and Terano, 2006)	GA	F_a, F_b, F_c all receive+1 if prediction is right	Single/NM/discrete/NM/NM	CR=0.6; 0.8, MR=0.01;0.05
(Yamada and Ueda, 2005)	GA	F_a, F_b, F_c all receive+1 if prediction is right	Single/maximization/discrete/NM/NM	CR=0.6; 0.8, MR=0.01;0.05
(Huang et al., 2011)	GA	Average monthly return of portfolio	Single/maximization/discrete/NM/NM	NM
(Carvalho et al., 2014)	GA	$\frac{(No.\ of\ correctly\ classified\ tweets)}{Total\ tweets\ already\ classified}$	Single/maximization/discrete/NM/NM	PS=100, CR=0.9, MR=0.01, MG=20
(Hochreiter, 2015)	GA	Optimal rules based on classification (investment rules)	Single/minimization/discrete/constrained/NM	PS=100, Elitist=10
(Chen et al., 2007)	GP	Sensitivity+specificity	Single/maximization/discrete/unconstrained/NM	PS=1000, CR=0.9, MR=0.1, MG=50
(Sotoodeh, 2012)	ACO	User time window for churn based on previous activity	Single/maximization/NM/constrained/NM	Pheromone level=1 to 10, Pheromone evaporation=0.3, Average number of neighbors=8, Average time Period=8, Evaporation time period for each user=1; 2; 4
(Qiwan and Min, 2008)	IGA	$\sum_i W_i \times F_i(x)$ Calculation of non-inferior solutions	Multi/minimization/NM/constrained/NM	NM
(Basiri and Taghiyareh, 2012)	ICA	1-($\alpha\times$(accuracy)+$\beta\times$(coverage))	Single/maximization/discrete/constrained/NM	No. of iterations=3000, No. of initial rules=50, Initial imperialists=8, α=0.95 and β=0.5
(Soltani et al., 2012)	AIS	Fraud ratio	Single/maximization/NM/NM/NM	NM
(Bhugra et al., 2013)	BBO	No. of rules satisfying minimum support and high confidence	Single/maximization/discrete/NM/NM	Immigration rate=∞,

Table A2 (*continued*)

Reference	EC technique	Objective function used	(Single- or multi-objective)/(maximization or minimization)/(discrete or continuous)/(constrained or unconstrained)/(unimodal or multimodal)	Parameter values set
(Khademolghorani, 2011)	ICA	$\alpha_1 \times \frac{SUP(A \cup C)}{SUP(A)} \times \frac{SUP(A \cup C)}{SUP(C)} \times \left[1-\frac{SUP(A \cup C)}{\lvert D\rvert}\right]+\alpha_1\frac{NumberField(i)}{MaxField}$	Single/maximization/continuous/NM/NM	$S_{max}=\infty$, $s_{min}=1$, Emigration rate = rules with minimum support No. of Countries = 70, Maximum Generations = 100, No. of imperialists = 10
(Birtolo et al., 2013)	GA	$\sigma \times (degreeofcompliancewithconstarints)+(1-\sigma)\times(retailerneeds)$	Single/maximization/discrete/constrained/NM	CR = 0.8, Tournaments = 3, MG = 1000, elitists = 3, σ=0. 8
(Christian and Martin, 2010)	GA	Interestingness factor × Completeness factor	Single/maximization/discrete/NM/NM	NM
(Cui et al., 2010)	GA	M(e) = Amount of money spent by customer 'e'	Single/maximization/discrete/constrained/NM	NM
(Bhattacharyya, 2000)	GA, GP	Churn per customer and revenue per customer	Single/NM/continuous/NM/NM	NM
(Gupta et al., 2015)	PSO	F-score	Single/maximization/discrete/NM/NM	No. of particles = 10; 25, Maximum Iterations = 200, Inertia = 0.89; 0.95; 0.3928
(Kaiser et al., 2010)	ACO	Rules with high accuracy	Single/maximization/NM/NM/NM	NM
(Dehuri et al., 2008)	HBB	$\sum_i W_i \times \Sigma(campaign\ preferenc \times campaign\ assigned\ or\ not\ \times recommendations)$	Single/maximization/NM/constrained/NM	No. of iterations = 100, No. of scout bees = 25, No. of sites selected = 100, No. of best sites = 5, No. of bees recruited for best sites = 10, No. of bees recruited for other sites = 5
(Cunha and Castro, 2013)	AIS	$W_1 \times support + W_2 \times confidence$	Single/maximization/discrete/constrained/NM	NM
(Mzoughia and Limam, 2014)	MOGA	1. No. of transactions and 2. CLTV	Multi/maximization/NM/NM/NM	NM
(Cheng, 2005)	GA	Most total lost profit 'R_S' and least total lost profit 'R_L'	Single/maximization/discrete/NM/NM	NM
(Duman and Elikucuk, 2013)	MBO	Savings limit and true positive rate	Single/maximization/NM/NM/NM	No. of birds = 15, No. of neighbors = 25, No. of flaps = 100, No. of overlaps = 5,
(Faris et al., 2014)	GP	Mean square error	Single/minimization/NM/NM/NM	Maximum tree depth = 10, Maximum tree length = 50, Elites = 1, PS = 1000, MG = 100, MR = 15%
(Wong et al., 2012)	AIS	Detection rate, false positive rate, fraud ratio	Single/maximization/NM/NM//NM	NM
(Brabazon et al., 2010)	AIS	Heterogeneous value distance metric	Single/maximization/continuous/NM/NM	NM
(Gadi et al., 2008)	AIS, GA	Cost = \$100 × false negative + \$10 × false positive + \$1 × true positive	Single/maximization/NM/NM/NM	Parameters of AIS optimized using GA
(R. Huang et al., 2010)	AIS	True positive and true negative	Single/maximization/NM/NM/NM	NM
(Ganghishetti and Vadlamani, 2014)	MO-BPSO, MO-BFFO-TA, MO-BPSO-TA	$\prod_{i=1}^{8} Measurevalue(i)$ Support, confidence, interestingness, comprehensibility, lift, leverage, conviction, coverage	Multi/maximization/discrete/unconstrained/NM	NM
(Brun et al., 2009)	DE	Cost = y + (0.8 × coverage) 'y' depends on accuracy and certainty	Single/maximization/NM/NM/NM	NM
(Banati and Bajaj, 2012)	Firefly Algorithm	User interest	Single/maximization/NM/NM/NM	NM
(Chien and Chen, 2010)	GA	1. Confidence 2. Support × confidence 3. $\sqrt{Support}\ \times Confidence$ 4. Confidence × $\log(Support)$	Single/maximization/discrete/NM/multimodal	PS = 20, MG = 50, CR = 0.7, MR = 0.001, Minimum support = 0.002, Minimum confidence = 70%, Top K rules; K = 3

*NM = Not mentioned.

combination of parameters for the EC techniques in a given application. Similar to all other non-parametric, intelligent techniques, this is one limitation of these techniques.

6. Conclusions

In this survey, we reviewed all the articles from 1998- to 2015 that are related to the application of EC techniques to CRM tasks. This paper has provided a detailed review based on ten (10) CRM tasks like MBA, fraud detection, churn detection, sentiment analysis, direct marketing, customer segmentation, Customer Life Time Value (LTV), default detection, service marketing and customer campaigning. The paper also details how EC techniques can be used to improve the relationship with the customer, thereby increasing the return on investment to the company.

We conclude that EC techniques have a spectacular role to play in solving CRM tasks and that this survey will be useful to practitioners from service industries, academic researchers working in the CRM domain. Survey will also be useful for data miners working in the field of optimization, in general and EC techniques, in particular. The list of future directions is going to be very useful for budding research scholars in this area.

Appendix A

Evolutionary computation (EC) techniques in review

EC techniques which are included in the survey are summarized in the Tables A1 and A2 . There are some point based optimization techniques like Simulated Annealing, etc., also alongside population-based EC techniques. EC Techniques used in the review are summarized with the basic idea, advantages, and disadvantages in the form of a Table A1. EC techniques utilizing various objective functions along with the properties of objective functions and parameters that are set in each research article are summarized in Table A2. Though the tables span many pages they present a detailed overview of the EC techniques surveyed.

References

Abdou, Ha, 2009. Genetic programming for credit scoring: the case of Egyptian public sector banks. Expert Syst. Appl. 36, 11402–11417. http://dx.doi.org/10.1016/j.eswa.2009.01.076.

Aliehyaei, R., Khan, S., 2014. Ant Colony Optimization, Genetic Programming and a hybrid approach for credit scoring: a comparative study. In: Proceedings of the 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014). IEEE, pp. 1–5. 〈http://dx.doi.org/10.1109/SKIMA.2014.7083391http://dx.doi.org/10.1109/SKIMA.2014.7083391〉.

Artun, Ö., Levin, D., 2015. Predictive Marketing: Easy Ways Every Marketer Can Use Customer Analytics and Big Data. John Wiley & Sons, Inc, Hoboken, NJ, USA.

Assis, C.A.S., Pereira, A.C.M., Pereira, M.A., Carrano, E.G., 2014. A genetic programming approach for fraud detection in electronic transactions. In: Proceedings of the 2014 IEEE Symposium on Computational Intelligence in Cyber Security (CICS). IEEE, pp. 1–8. 〈http://dx.doi.org/10.1109/CICYBS.2014.7013373http://dx.doi.org/10.1109/CICYBS.2014.7013373〉.

Atashpaz-Gargari, E., Lucas, C., 2007. Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition. In: Proceedings of the 2007 IEEE Congress on Evolutionary Computation. pp. 4661–4667. 〈http://dx.doi.org/10.1109/CEC.2007.4425083http://dx.doi.org/10.1109/CEC.2007.4425083〉.

Au, W., Chan, K.C.C., Yao, X., 2003. A novel evolutionary data mining algorithm with applications to churn prediction. IEEE Trans. Evol. Comput. 7, 532–545. http://dx.doi.org/10.1109/TEVC.2003.819264.

Back, T., Fogel, D., Michalewicz, Z., 1997. Handbook of Evolutionary Computation. IOP Publishing Ltd., Bristol, UK.

Banati, H., Bajaj, M., 2012. Promoting products online using firefly algorithm. In: Proceedings of the 2012 12th International Conference on Intelligent Systems Design and Applications (ISDA). IEEE, pp. 580–585. 〈http://dx.doi.org/10.1109/ISDA.2012.6416602http://dx.doi.org/10.1109/ISDA.2012.6416602〉.

Bandaru, S., Gaur, A., Deb, K., Khare, V., Chougule, R., Bandyopadhyay, P., 2015. Development, analysis and applications of a quantitative methodology for assessing customer satisfaction using evolutionary optimization. Appl. Soft Comput. 30, 265–278. http://dx.doi.org/10.1016/j.asoc.2015.01.014.

Basiri, J., Taghiyareh, F., 2012. An application of the CORER classifier on customer churn prediction. In: Proceedings of 2012 Sixth International Symposium on Telecommunications (IST). pp. 867–872. 〈http://dx.doi.org/10.1109/istel.2012.6483107http://dx.doi.org/10.1109/istel.2012.6483107〉.

Bhattacharyya, S., 1999. Direct marketing performance modeling using genetic algorithms. INFORMS J. Comput. 11, 248–257. http://dx.doi.org/10.1287/ijoc.11.3.248.

Bhattacharyya, S., 2000. Evolutionary algorithms in data mining. In: Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '00. ACM Press, New York, New York, USA, pp. 465–473. 〈http://dx.doi.org/10.1145/347090.347186http://dx.doi.org/10.1145/347090.347186〉.

Bhugra, D., Goel, S., Singhania, V., 2013. Association rule analysis using biogeography based optimization. In: Proceedings of the 2013 International Conference on Computer Communication and Informatics. pp. 1–5. 〈http://dx.doi.org/10.1109/ICCCI.2013.6466106http://dx.doi.org/10.1109/ICCCI.2013.6466106〉.

Birtolo, C., De Chiara, D., Losito, S., Ritrovato, P., Veniero, M., 2013. Searching optimal product bundles by means of GA-based Engine and Market Basket Analysis. In: Proceedings of IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013 Joint. pp. 448–453.

Bozsik, J., 2010. Genetic algorithm in default forecast. In: Proceedings of the 8th International Conference on on Applied Informatics. pp. 41–51.

Brabazon, A., Cahill, J., Keenan, P., Walsh, D., 2010. Identifying online credit card fraud using Artificial Immune Systems. In: Proceedings of IEEE Congress on Evolutionary Computation. IEEE, pp. 1–7. 〈http://dx.doi.org/10.1109/CEC.2010.5586154http://dx.doi.org/10.1109/CEC.2010.5586154〉.

Brun, A.D.M., Pereira Pinto, J.O., Carvalho Pinto, A.M.A., Sauer, L., Colman, E., 2009. Fraud detection in electric energy using differential evolution. In: Proceedings of the 2009 15th International Conference on Intelligent System Applications to Power Systems. pp. 1–5. 〈http://dx.doi.org/10.1109/ISAP.2009.5352917http://dx.doi.org/10.1109/ISAP.2009.5352917〉.

Brusco, M., Cradit, J., Stahl, S., 2002. A simulated annealing heuristic for a bicriterion partitioning problem in market segmentation. J. Mark. Res.

Cai, W., Lin, X., Huang, X., Zhong, H., 2009. A genetic algorithm model for personal credit scoring. In: Proceedings of the 2009 International Conference on Computational Intelligence and Software Engineering. pp. 1–4. 〈http://dx.doi.org/10.1109/CISE.2009.5366208http://dx.doi.org/10.1109/CISE.2009.5366208〉.

Carvalho, J., Prado, A., Plastino, A., 2014. A statistical and evolutionary approach to sentiment analysis. In: Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT). pp. 110–117. 〈http://dx.doi.org/10.1109/WI-IAT.2014.87http://dx.doi.org/10.1109/WI-IAT.2014.87〉.

Castro, L. De, Timmis, J., 2002. Artificial immune systems: a new computational intelligence approach. Springer Science & Business Media, Springer-Verlag, London.

Chan, C.C.H., 2008. Intelligent value-based customer segmentation method for campaign management: a case study of automobile retailer. Expert Syst. Appl. 34, 2754–2762. http://dx.doi.org/10.1016/j.eswa.2007.05.043.

Chen, Q., Zhang, D., Wei, L., Chen, H., 2007. A modified genetic programming for behavior scoring problem. In: Proceedings of IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2007. pp. 535–539. 〈http://dx.doi.org/10.1109/CIDM.2007.368921http://dx.doi.org/10.1109/CIDM.2007.368921〉.

Cheng, Y., 2005. Genetic algorithm for item selection with cross-selling. In: Proceedings of the 2005 International Conference on Machine Learning and Cybernetics. pp. 18–21.

Chien, Y.W.C., Chen, Y.L., 2010. Mining associative classification rules with stock trading data – a GA-based method. Knowl. Based Syst. 23, 605–614. http://dx.doi.org/10.1016/j.knosys.2010.04.007.

Chiu, C.Y., Kuo, I.T., 2010. Applying particle swarm optimization and honey bee mating optimization in developing an intelligent market segmentation system. J. Syst. Sci. Syst. Eng. 19, 182–191. http://dx.doi.org/10.1007/s11518-010-5135-9.

Chiu, C.-Y., Kuo, I.-T., Lin, C.-H., 2009. Applying artificial immune system and ant algorithm in air-conditioner market segmentation. Expert Syst. Appl. 36, 4437–4442. http://dx.doi.org/10.1016/j.eswa.2008.05.005.

Chorianopoulos, A., 2015. Effective CRM Using Predictive Analytics. John Wiley & Sons, Ltd, Chichester, UK.

Christian, A.J., Martin, G.P., 2010. Optimization of association rules with genetic algorithms. In: 2010 XXIX International Conference of the Chilean Computer Science Society. pp. 193–197. 〈http://dx.doi.org/10.1109/SCCC.2010.32http://dx.doi.org/10.1109/SCCC.2010.32〉.

Colorni, A., Dorigo, M., Maniezzo, V., 1991. Distributed optimization by ant colonies. In: Proceedings of the European Conference on Artificial Life. pp. 134–142.

Corazza, M., Funari, S., Gusso, R., 2014. Particle swarm optimization for preference disaggregation in multicriteria credit scoring problems. In: Corazza, M., Pizzi, C. (Eds.), Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer International Publishing. pp. 119–130.

Cui, G., Wong, M.L., Wan, X., 2015. Targeting high value customers while under resource constraint: partial order constrained optimization with genetic algorithm. J. Interact. Mark. 29, 27–37. http://dx.doi.org/10.1016/j.intmar.2014.09.001.

Cui, G., Wong, M.L., Wan, X., 2010. Constrained optimization with genetic algorithm: improving profitability of targeted marketing. In: Proceedings of the

2010 International Conference on Management of E-Commerce and E-Government. pp. 26–30. 〈http://dx.doi.org/10.1109/ICMeCG.2010.14http://dx.doi.org/10.1109/ICMeCG.2010.14〉.
Cunha, D.S. da, Castro, L.N. de, 2013. Bioinspired algorithms applied to association rule mining in electronic commerce databases. In: Proceedings of the 2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence. pp. 189–194. 〈http://dx.doi.org/10.1109/BRICS-CCI-CBIC.2013.40http://dx.doi.org/10.1109/BRICS-CCI-CBIC.2013.40〉.
Dasgupta, D., 1993. An overview of artificial immune systems and their applications. In: Daspupta, D. (Eds.), Artificial immune systems and their applications. Springer Berlin Heidelberg. pp. 3–21.
Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. 6, 182–197. http://dx.doi.org/10.1109/4235.996017.
Dehuri, S., Cho, S.-B., Jagadev, A.K., 2008. Honey bee behavior: a multi-agent approach for multiple campaigns assignment problem. In: Proceedings of the 2008 International Conference on Information Technology. pp. 24–29. 〈http://dx.doi.org/10.1109/ICIT.2008.14http://dx.doi.org/10.1109/ICIT.2008.14〉.
Dong, G., Lai, K.K., Zhou, L., 2009. Simulated annealing based rule extraction algorithm for credit scoring problem. In: Proceedings of the 2009 International Joint Conference on Computational Sciences and Optimization. pp. 22–26. 〈http://dx.doi.org/10.1109/CSO.2009.450http://dx.doi.org/10.1109/CSO.2009.450〉.
Dueck, G., Scheurer, T., 1990. Threshold accepting: a general purpose optimization algorithm. J. Comput. Phys. 90, 161–175. http://dx.doi.org/10.1016/0021–9991(90)90201-B.
Duman, E., Uysal, M., Alkaya, A.F., 2012. Migrating birds optimization: a new metaheuristic approach and its performance on quadratic assignment problem. Inf. Sci. 217, 65–77. http://dx.doi.org/10.1016/j.ins.2012.06.032.
Duman, E., Elikucuk, I., 2013. Applying migrating birds optimization to credit card fraud detection. In: Li, J., Cao, L., Wang, C., Tan, K.C., Liu, B., Pei, J., Tseng, V.S. (Eds.), Trends and Applications in Knowledge Discovery and Data Mining. Springer Berlin Heidelberg. pp. 416–427.
Dyche, J., 2002. The CRM Handbook: A Business Guide to Customer Relationship Management. Addison-Wesley Professional, USA.
Eiben, A.E., Koudijs, A.E., Slisser, F., 1998. Genetic modelling of customer retention. In: Proceedings of the First European Workshop on Genetic Programming. pp. 178–186. 〈http://dx.doi.org/10.1007/BFb0055937http://dx.doi.org/10.1007/BFb0055937〉.
Faris, H., Al-Shboul, B., Ghatasheh, N., 2014. A genetic programming based framework for churn prediction in telecommunication industry. In: Hwang, D., Jung, J.J., Nguyen, N.T. (Eds.), Computational Collective Intelligence Technologies and Applications. Springer International Publishing. pp. 353–362.
Fogarty, D., 2012. Using genetic algorithms for credit scoring system maintenance functions. Int. J. Artif. Intell. Appl. 3, 1–8.
Gadi, M.F.A., Wang, X., Lago, A.P. Do, 2008. Credit card fraud detection with artificial immune system. In: Proceedings of 7th International Conference (ICARIS 2008). pp. 119–131. 〈http://dx.doi.org/10.1007/978-3-540-85072-4_11http://dx.doi.org/10.1007/978-3-540-85072-4_11〉.
Ganghishetti, P., Vadlamani, R., 2014. Association rule mining via evolutionary multi-objective optimization. In: Murty, M.N., He, X., Chillarige, R.R., Weng, P. (Eds.), Multi-Disciplinary Trends in Artificial Intelligence. Springer International Publishing. pp. 35–46.
Gao, Y., Zhang, G., Lu, J., Ma, J., 2014. A bi-level decision model for customer churn analysis. Comput. Intell. 30, 583–599. http://dx.doi.org/10.1111/coin.12008.
George, M., Kumar, V., Grewal, D., 2013. Maximizing profits for a multi-category catalog retailer. J. Retail. 89, 374–396. http://dx.doi.org/10.1016/j.jretai.2013.05.001.
Ghosh, A., Nath, B., 2004. Multi-objective rule mining using genetic algorithms. Inf. Sci. 163, 123–133. http://dx.doi.org/10.1016/j.ins.2003.03.021.
Glover, F., 1989a. Tabu search Part I. INFORMS J. Comput. 1, 190–206.
Glover, F., 1989b. Tabu search - Part II. ORSA J. Comput 2 (1), 4–32. http://dx.doi.org/10.1287/ijoc.2.1.4.
Golberg, D., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addion Wesley Longman Publishing Co., Inc., Boston, MA, USA.
Gordini, N., 2013. Genetic Algorithms for Small Enterprises Default Prediction: Empirical Evidence from Italy, Handbook of Research on Novel Soft Computing Intelligent Algorithms: Theory and Practical Applications.
Greenberg, P., 2009. CRM at the Speed of Light: Social CRM 2.0 Strategies, Tools, and Techniques for Engaging your Customers. McGraw Hill Professional, United States of America.
Guo, J.-L., Peng, J.-E., Wang, H.-C., 2013. An opinion feature extraction approach based on a multidimensional sentence analysis model. Cybern. Syst. 44, 379–401. http://dx.doi.org/10.1080/01969722.2013.789649.
Gupta, D., Reddy, K., Ekbal, A., 2015. PSO-ASent: feature selection using particle swarm optimization for aspect based sentiment analysis. In: Biemann, C., Handschuh, S., Freitas, A., Meziane, F., Métais, E. (Eds.), Natural Language Processing and Information Systems. Springer International Publishing. pp. 220–233.
Han, J., Kamber, M., Pei, J., 2011. Data Mining: Concepts and Techniques: Concepts and Techniques. Elsevier, Waltham, MA, USA.
Handl, J., Meyer, B., 2002. Improved ant-based clustering and sorting in a document retrieval interface. Parallel Probl. Solving Nat. VII, 913–923. http://dx.doi.org/10.1007/3-540-45712-7_88.
Hansen, J.M., Raut, S., Swami, S., 2010. Retail shelf allocation: a comparative analysis of heuristic and meta-heuristic approaches. J. Retail. 86, 94–105. http://dx.doi.org/10.1016/j.jretai.2010.01.004.
Hochreiter, R., 2015. Computing trading strategies based on financial sentiment data using evolutionary optimization. In: Mendel 2015. Advances in Intelligent Systems and Computing. pp. 181–191.
Hoogs, B., Kiehl, T., Lacomb, C., Senturk, D., 2007. A genetic algorithm approach to detecting temporal patterns indicative of financial statement fraud. Intell. Syst. Account. Financ. Manag 15, 41–56. http://dx.doi.org/10.1002/isaf.284.
Huang, B., Buckley, B., Kechadi, T.-M., 2010. Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications. Expert Syst. Appl. 37, 3638–3646. http://dx.doi.org/10.1016/j.eswa.2009.10.027.
Huang, B., Kechadi, M.T., Buckley, B., 2012. Customer churn prediction in telecommunications. Expert Syst. Appl. 39, 1414–1425. http://dx.doi.org/10.1016/j.eswa.2011.08.024.
Huang, J.-J., Tzeng, G.-H., Ong, C.-S., 2006. Two-stage genetic programming (2SGP) for the credit scoring model. Appl. Math. Comput. 174, 1039–1053. http://dx.doi.org/10.1016/j.amc.2005.05.027.
Huang, C.-F., Chang, C.-H., Chang, B.R., Hsieh, T.-N., 2011. A genetic-based stock selection model using investor sentiment indicators. In: Proceedings of the 2011 IEEE International Conference on Granular Computing. pp. 262–267. 〈http://dx.doi.org/10.1109/GRC.2011.6122605http://dx.doi.org/10.1109/GRC.2011.6122605〉.
Huang, R., Tawfik, H., Nagar, A.K., 2010. A novel hybrid artificial immune inspired approach for online break-in fraud detection. In: Sloota, P.M.A., Dick van Albadaa, G., Dongarra, J. (Eds.), Procedia Computer Science. Elsevier, pp. 2733–2742. 〈http://dx.doi.org/10.1016/j.procs.2010.04.307〉.
Iriana, R., Buttle, F., 2007. Strategic, operational, and analytical customer relationship management: attributes and measures. J. Relatsh. Mark. 5, 23–42.
Jin, P., Zhu, Y., Li, S., Hu, K., 2006. Application architecture of data mining in telecom customer relationship management based on swarm intelligence. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 4099 LNAI, pp. 960–964.
Jonker, J.J., Piersma, N., Van Den Poel, D., 2004. Joint optimization of customer segmentation and marketing policy to maximize long-term profitability. Expert Syst. Appl. 27, 159–168. http://dx.doi.org/10.1016/j.eswa.2004.01.010.
Jungwon, K., Ong, A., Overill, R.E., 2003. Design of an artificial immune system as a novel anomaly detector for combating financial fraud in the retail sector. In: Proceedings of the 2003 Congress on Evolutionary Computation, CEC 2003 – Proceedings. pp. 405–412. 〈http://dx.doi.org/10.1109/CEC.2003.1299604http://dx.doi.org/10.1109/CEC.2003.1299604〉.
Kaiser, C., Kröckel, J., Bodendorf, F., 2010. Ant-based simulation of opinion spreading in online social networks. In: Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. pp. 537–540. 〈http://dx.doi.org/10.1109/WI-IAT.2010.12http://dx.doi.org/10.1109/WI-IAT.2010.12〉.
Karaboga, D., Akay, B., 2009. A survey: algorithms simulating bee swarm intelligence. Artif. Intell. Rev. 31, 61–85. http://dx.doi.org/10.1007/s10462-009-9127-4.
Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proceedings of the 1995 IEEE International Conference on Neural Networks. Vol 4, pp. 1942–1948. 〈http://dx.doi.org/10.1109/ICNN.1995.488968http://dx.doi.org/10.1109/ICNN.1995.488968〉.
Khademolghorani, F., 2011. An effective algorithm for mining association rules based on imperialist competitive algorithm. In: Proceedings of the Sixth International Conference on Digital Information Management. pp. 6–11. 〈http://dx.doi.org/10.1109/ICDIM.2011.6093350http://dx.doi.org/10.1109/ICDIM.2011.6093350〉.
Kibler, D., Aha, D.W., Albert, M.K., 1989. Instance-based prediction of real-valued attributes. Comput. Intell. 5, 51–57. http://dx.doi.org/10.1111/j.1467–8640.1989.tb00315.x.
Kirkpatrick, S., Jtr, C.G., Vecchi, M., 1983. Optimization by simulated annealing. Science (80). http://dx.doi.org/10.1126/science.220.4598.671.
Konak, A., Coit, D.W., Smith, A.E., 2006. Multi-objective optimization using genetic algorithms: a tutorial. Reliab. Eng. Syst. Saf. 91, 992–1007. http://dx.doi.org/10.1016/j.ress.2005.11.018.
Kotler, P., Keller, K.L., 2012. Marketing Management. Printice Hall, New Jersey, USA http://dx.doi.org/10.1080/08911760903022556.
Koza, J., 1992. Genetic programming: on the programming of computers by means of natural selection. MIT press, MA, USA.
Kožený, V., 2015. Genetic algorithms for credit scoring: alternative fitness function performance comparison. Expert Syst. Appl. 42, 2998–3004. http://dx.doi.org/10.1016/j.eswa.2014.11.028.
Kracklauer, A.H., Mills, D.Q., Seifert, D., 2004. Collaborative Customer Relationship Management: Taking CRM to the Next Level. Springer Science & Business Media, Berlin, Heidelberg.
Krishnamoorthy, S., 2015. Pruning strategies for mining high utility itemsets. Expert Syst. Appl. 42, 2371–2381.
Kumar, V., Reinartz, W., 2012. Customer Relationship Management: Concept, Strategy, and Tools. Springer Science & Business Media, Berlin, Heidelberg.
Kuo, R.J., Shih, C.W., 2007. Association rule mining through the ant colony system for national health insurance research database in Taiwan. Comput. Math. Appl. 54, 1303–1318. http://dx.doi.org/10.1016/j.camwa.2006.03.043.
Kuo, R.J., Chao, C.M., Chiu, Y.T., 2011. Application of particle swarm optimization to association rule mining. Appl. Soft Comput. 11, 326–336. http://dx.doi.org/10.1016/j.asoc.2009.11.023.
Lin, C.W., Hong, T.P., Lu, W.H., 2011. An effective tree structure for mining high utility itemsets. Expert Syst. Appl. 38, 7419–7424.
Liu, S., Wang, Q., Lv, S., 2008. Application of Genetic Programming in credit scoring. In: Proceedings of the 2008 Chinese Control and Decision Conference. IEEE, pp.

1106–1110. ⟨http://dx.doi.org/10.1109/CCDC.2008.4597485http://dx.doi.org/10.1109/CCDC.2008.4597485⟩.

Luna, J.M., Romero, J.R., Ventura, S., 2013. Grammar-based multi-objective algorithms for mining association rules. Data Knowl. Eng. 86, 19–37. http://dx.doi.org/10.1016/j.datak.2013.01.002.

Mabu, S., Tjahjadi, A., Sendari, S., Hirasawa, K., 2010. Evaluation on the robustness of genetic network programming with reinforcement learning. In: Proceedings of the 2010 IEEE International Conference on Systems Man and Cybernetics (SMC). pp. 1659–1664. ⟨http://dx.doi.org/10.1109/ICSMC.2010.5642323http://dx.doi.org/10.1109/ICSMC.2010.5642323⟩.

Mahdavi, I., Movahednejad, M., Adbesh, F., 2011. Designing customer-oriented catalogs in e-CRM using an effective self-adaptive genetic algorithm. Expert Syst. Appl. 38, 631–639. http://dx.doi.org/10.1016/j.eswa.2010.07.013.

McKay, R.I., Hoai, N.X., Whigham, P.A., Shan, Y., O'Neill, M., 2010. Grammar-based genetic programming: a survey. Evol. Mach. Genet. Progr. 11, 365–396. http://dx.doi.org/10.1007/s10710-010-9109-y.

Mzoughia, M. Ben, Limam, M., 2014. A multicriterion segmentation approach based on CLV components. In: Proceedings of the 2014 IEEE 15th International Symposium on Computational Intelligence and Informatics (CINTI).

Ngai, E.W., 2005. Customer relationship management research (1992–2002) an academic literature review and classification. Mark. Intell. Plan. 23, 582–605.

Ngai, E.W.T., Xiu, L., Chau, D.C.K., 2009. Application of data mining techniques in customer relationship management: a literature review and classification. Expert Syst. Appl. 36, 2592–2602. http://dx.doi.org/10.1016/j.eswa.2008.02.021.

Ong, C., Huang, J., Tzeng, G., 2005. Building credit scoring models using genetic programming. Expert Syst. Appl. 29, 41–47. http://dx.doi.org/10.1016/j.eswa.2005.01.003.

Ozçelik, M.H., Duman, E., Isik, M., Cevik, T., 2010. Improving a credit card fraud detection system using genetic algorithm. In: Proceedings of the 2010 International Conference on Networking and Information Technology. pp. 436–440. ⟨http://dx.doi.org/10.1109/ICNIT.2010.5508478http://dx.doi.org/10.1109/ICNIT.2010.5508478⟩.

Padmanabhan, B., Tuzhilin, A., 2003. On the use of optimization for data mining: theoretical interactions and eCRM opportunities. Manag. Sci. 49, 1327–1343. http://dx.doi.org/10.1287/mnsc.49.10.1327.17310.

Palshikar, G., 2002. The hidden truth-frauds and their control: a critical application for business intelligence. Intell. Enterp. 5, 46–51.

Provost, F., Fawcett, T., 2013. Data Science for Business: What you Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media, Inc, Sebastopol, CA, USA.

Qiwan, W., Min, Z., 2008. Study on the dual optimization of service marketing and customer churn management. In: Proceedings of the 2008 International Seminar on Business and Information Management. pp. 324–328. ⟨http://dx.doi.org/10.1109/ISBIM.2008.49http://dx.doi.org/10.1109/ISBIM.2008.49⟩.

Ravi, V., Koparkar, S., Raju, N.P., Sridher, S., 2015. Improving retail banking loans recovery via data mining techniques: acase study of Indian market. Int. J. Electron. Cust. Relatsh. Manag. 9, 189–201.

Rygielski, C., Wang, J.C., Yen, D.C., 2002. Data mining techniques for customer relationship management. Technol. Soc. 24, 483–502.

Sarath, K.N.V.D., Ravi, V., 2013. Association rule mining using binary particle swarm optimization. Eng. Appl. Artif. Intell. 26, 1–9. http://dx.doi.org/10.1016/j.engappai.2013.06.003.

Shenoy, P.D., Srinivasa, K.G., Venugopal, K.R., Patnaik, L.M., 2005. Dynamic association rule mining using genetic algorithms. Intell. Data Anal. 9, 439–453.

Simon, D., Member, S., 2008. Biogeography-based optimization. Evol. Comput. IEEE Trans. 12, 702–713. http://dx.doi.org/10.1109/TEVC.2008.919004.

Soltani Halvaiee, N., Akbari, M.K., 2014. A novel model for credit card fraud detection using artificial immune systems. Appl. Soft Comput. 24, 40–49. http://dx.doi.org/10.1016/j.asoc.2014.06.042.

Soltani, N., Akbari, M.K., Sargolzaei Javan, M., 2012. A new user-based model for credit card fraud detection based on artificial immune system. In: Proceedings of the 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012). pp. 029–033. ⟨http://dx.doi.org/10.1109/AISP.2012.6313712http://dx.doi.org/10.1109/AISP.2012.6313712⟩.

Sotoodeh, H., 2012. Churn calculation based on swarm intelligence algorithms in social networks. In: Proceedings of the 2012 IEEE 2nd International eConference on Computer and Knowledge Engineering (ICCKE). pp. 173–179.

Storn, R., Price, K., 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. J. Glob. Optim., 341–359. http://dx.doi.org/10.1023/A:1008202821328.

Subramaniam, Sakthi, Thangavelu, Arunkumar, 2011. CCHUM – an adaptive business intelligence strategy to improve Customer Retention using bee-hive classification algorithm. Int. J. Res. Rev. Inf. Technol. 1, 35–40.

Takagi, H., 2001. Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation. Proc. IEEE 89, 1275–1296. http://dx.doi.org/10.1109/5.949485.

Teo, T.S., Devadoss, P., Pan, S.L., 2006. Towards a holistic perspective of customer relationship management (CRM) implementation: a case study of the Housing and Development Board, Singapore. Decis. Support Syst. 42, 1613–1627.

Thabtah, F., 2007. A review of associative classification mining. Knowl. Eng. Rev. 22, 37–65.

Tseng, V.S., Wu, C.W., Fournier-Viger, P., Philip, S.Y., 2015. Efficient algorithms for mining the concise and lossless representation of high utility itemsets. IEEE Trans. Knowl. Data Eng. 27, 726–739.

Tseng, V.S., Wu, C.W., Fournier-Viger, P., Philip, S.Y., 2016. Efficient algorithms for mining top-k high utility itemsets. IEEE Trans. Knowl. Data Eng. 28, 54–67.

Van den Poel, D., Larivière, B., 2004. Customer attrition analysis for financial services using proportional hazard models. Eur. J. Oper. Res. 157, 196–217. http://dx.doi.org/10.1016/S0377-2217(03)00069-9.

Waad, B., Ghazi, B.M., Mohamed, L., 2014. A New Feature Selection Technique Applied to Credit Scoring Data Using a Rank Aggregation Approach Based on: Optimization, Genetic Algorithm and Similarity, Knowledge Discovery Process and Methods to Enhance Organizational Performance.

Wahlberg, O., Strandberg, C., Sandberg, K.W., 2009. Trends, topics and under-researched areas in CRM research – a literature review. Int. J. Public Inf. Syst. 5.

Wang, C.M., Huang, Y.F., 2009. Evolutionary-based feature selection approaches with new criteria for data mining: a case study of credit approval data. Expert Syst. Appl. 36, 5900–5908. http://dx.doi.org/10.1016/j.eswa.2008.07.026.

Wolpert, D.H., Macready, W.G., 1997. No free lunch theorems for optimization. IEEE Trans. Evol. Comput. 1, 67–82.

Wong, N., Ray, P., Stephens, G., Lewis, L., 2012. Artificial immune systems for the detection of credit card fraud: an architecture, prototype and preliminary results. Inf. Syst. J. 22, 53–76. http://dx.doi.org/10.1111/j.1365-2575.2011.00369.x.

Xu, M., Walton, J., 2005. Gaining customer knowledge through analytical CRM. Ind. Manag. Data Syst. 105, 955–971.

Yamada, T., Terano, T., 2006. Price formations in genetic learning model of investor sentiment. In: Proceedings of the 9th Joint International Conference on Information Sciences.

Yamada, T., Ueda, K., 2005. Explanation of binarized time series using genetic learning model of investor sentiment. In: Proceedings of the 2005 IEEE Congress on Evolutionary Computation. IEEE, pp. 2437–2444. ⟨http://dx.doi.org/10.1109/CEC.2005.1554999http://dx.doi.org/10.1109/CEC.2005.1554999⟩.

Yang, G.F., Mabu, S., Shimada, K., Hirasawa, K., 2011. An evolutionary approach to rank class association rules with feedback mechanism. Expert Syst. Appl. 38, 15040–15048. http://dx.doi.org/10.1016/j.eswa.2011.05.042.

Yang, X., 2010. Nature-inspired metaheuristic algorithms. Luniver Press, Bristol, UK.

Zhang, X.-X., 2008. Applying incremental ant colony clustering to value-based customer segmentation. In: Proceedings of the 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing. IEEE, pp. 1–4. ⟨http://dx.doi.org/10.1109/WiCom.2008.2588http://dx.doi.org/10.1109/WiCom.2008.2588⟩.

Zitzler, E., Laumanns, M., Thiele, L., 2001. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. ⟨http://dx.doi.org/10.1.1.28.7571⟩.