Volume 3, Issue 11, November 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Hybrid Data Mining Metaheuristic Approach for Anomaly Detection

Meenakshi, Sandeep Jaglan
N.C. College of Engg
Panipat,Haryana, India

Abstract: Along with great increase in credit card transactions, credit card fraud has become increasingly unbridled in recent years. Fraud is one of the major causes of great financial losses, not only for merchants, individuals clients are also affected. In this paper two techniques support vector machine and particle swarm optimization is used as a hybrid approach for fraud detection. Using data from a credit card issuer, a support vector machine based fraud detection system was trained on a large sample of labeled credit card transactions and tested on a proffer dataset that consisted of all account activity over a subsequent time period. Particle swarm optimization is used to optimize the high dimensional space. In the present study, we combine these two methods and introduce a new method for anomaly detection.

1. Introduction

Data mining is the process of automatically discovering useful information in large data repositories. Data mining (dm) techniques are deployed to scour large database in order to find novel and useful patterns.dm techniques are clustering, association rule, classification and support vector machine etc. One of the important areas of data mining is fraud detection. Fraud differs and varies from act to act likes Telecommunications Fraud ,Insurance Fraud and Health Care Fraud, Money Laundering ,Computer Intrusion, Identity Theft ,Credit Card Fraud. The use of credit cards is prevalent in modern day society. Detecting credit card fraud is a difficult task when using normal procedures, so the development of the credit card fraud detection model has become of significance, whether in the academic or business community ecently. The majority of the loss due to credit card fraud is suffered by the USA alone. This is not surprising since 71% of all credit cards are issued in the USA only. In 2005, the total fraud loss in the USA was reported to be $2.7 billion and it has gone up to $3.2 billion in 2007 [7]. Another survey of over 160 companies revealed that online fraud (committed over the Web or phone shopping) is 12 times higher than offline fraud (committed by using a stolen physical card) [8]. credit card fraud has broader twigs, as such fraud helps fund organized crime, international narcotics trafficking, and even terrorist financing [4,5]. Over the years, along with the fruition of fraud detection methods, doers' fraud has also been evolving their fraud practices to avoid detection [6]. Therefore, credit card fraud detection methods need stable novelty. In this study, we evaluate two techniques support vector machines and particle swarm optimization, as a part of an attempt to better detects credit card fraud. The study is based on spending behavior of user. The spending profile of a cardholder suggests his normal spending behavior. Cardholders can be broadly categorized into three groups based on their spending habits, namely, high-spending (hs) group, medium-spending (ms) group, and low-spending (ls) group. Cardholders, who belong to the hs group, normally use their credit cards for buying high priced items. Fraud detection methods have been divided into two broad categories: supervised and unsupervised [6]. In supervised fraud detection methods, models are projected based on the samples of deceitful and genuine transactions, to classify new transactions as deceitful or genuine. In unsupervised fraud detection, outliers or unusual transaction are identified as potential cases of deceitful transaction. Credit card fraud detection is an application of anomaly detection (looking for buying patterns different from typical behavior).

Anomaly detection refers to detecting patterns in a proffer data set that do not conform to an established, normal behavior. The patterns thus detected are called anomalies, outliers, surprise, aberrant, deviation, peculiarity, etc. Anomaly detection techniques are three broad categories:

Supervised anomaly detection techniques become skilled at a classifier, using labeled occurrence belonging to customary and anomaly classes, and then disperse a customary or anomalous label to a test occurrence [1, 2].Unsupervised anomaly detection techniques detect anomalies in an unlabeled test data set, under the postulation that the preponderance of occurrence in the data set are customary [1,2].Semi-supervised anomaly detection techniques erect a model representing customary behavior from a proffer training data set [1–3].To detect anomaly we use here a data mining technique.

Support vector machine is used to work in high dimensional feature space, in this no upper limit for the number of attributes.SVM can model text and real image classification, handwriting recognition and speaker identification. In this paper support vector machine is used to classify the genuine and deceitful transaction. Particle swarm optimization is used to optimize the high dimensional space. The innovative Particle Swarm Optimization (PSO) is a different algorithm discovered at some stage in requisite social model simulation, which is effectual in nonlinear optimization problems and simple to implement, only a few input parameters require to be familiar because the update process is based on simple equation.PSO can be used competently used on large data sets. The remainder of the paper is organized as follows. Section 2 reveals the existing literature for different types of credit card fraud detection and fraud detection techniques. The next Section Proposed fraud detection system describes the SVM and PSO based heuristic local search

procedure. Section 4 will discuss about the Implementation and results, which we have obtained. At the last section of the paper we derive some conclusion.

2. Literature Survey

Anomaly detection systems work by annoying to recognize anomalies in an environment [13]. At the untimely stage, the research center lies in using rule-based expert systems and statistical approaches. But when encountering bigger datasets, the results of rule-based expert systems and statistical approaches become shoddier. Thus, many data mining techniques have been commenced to solve the problem. Among these techniques, the Artificial Neural Network (ANN) is extensively used and has been successful in solving many complex practical problems [14].An important part of anomaly detection methods is alert on computer intrusion detection. The charge of an intrusion detection system is to protect a computer system by sensing and analyzing stabed breaches of the integrity of the system [16]. The process of involuntarily constructing models from data is not trivial, particularly for intrusion detection problems. This is because intrusion detection faces problems, such as gigantic network traffic volumes, highly unprovoked data distribution, the difficulty of realizing resolution boundaries between normal and abnormal behavior, and a requisite for continuous adaptation to a constantly changing environment. Artificial intelligence and machine learning have shown confines in achieving high detection accuracy and fast processing times when confronted with these rations [2]. More recently, some research has begun to employ the clustering competencies of neural networks in fraud detection [18]. Self-organizing maps or Kohonen maps, kinfolk of neural networks, are being suggested for forming customer profiles and evaluating fraud patterns. In this research, all transactions in the payment system (PS) are classified into genuine and deceitful sets.

Aleskerov et al. [15] present a database mining system (CARDWATCH) intended for fraud detection in credit card. This system is base on a neural learning module which grants an edge to a variety of commercial databases .It makes the network process the current spending pattern to distinguish possible anomalies.

Suvasini Panigrahi [17] proposed a narrative approach for credit card fraud detection in which current as well as past behavior of customer is studied. The fraud detection system (FDS) consists of four components, to be precise, rule-based filter, Dempster–Shafer adder, transaction history database and Bayesian learner. The distrust level of each incoming transaction based on the extent of its variation from good pattern is determined by rule-based component. Dempster–Shafer's theory is used to combine multiple such indications and an initial belief is computed so that transaction can be classified as normal, abnormal or apprehensive. Once a transaction is found to be apprehensive, belief is further strengthened or diluted according to its similarity with deceitful or genuine transaction history using Bayesian learning.

Another issue, as noted by Provost [19], is that the value of fraud detection is a utility of time. The quicker a fraud gets detected, the greater the preventable loss. However, most fraud detection techniques need history of card holders' behavior for approximating models. Past research suggests that fraudsters try to exploit spending within short periods before frauds get detected and cards are withdrawn [20].

[21] Proposed a assessment of random forests and support vector machines, together with logistic regression, for credit card fraud detection. Performance is also measured across some measures.Chen, Q. [13] when encountering bigger datasets, the results of rule-based expert systems and statistical approaches become shoddier. Ambareen [16] faces Fraud detection problems, such as gigantic network traffic volumes, highly unprovoked data distribution, the difficulty of realizing resolution boundaries between normal and abnormal behavior. Zaslavsky [18] produce cost of clustering is very high, due to any samples of deceitful and genuine transactions, to classify new transactions as deceitful or genuine its takes more processing steps and time.

This paper proposes a model (SPSO that is combination of SVM and PSO) in which a supervised learning technique (Support Vector machine) is used for detection of pattern or abnormal behavior of user from huge amount of data. The innovative Particle Swarm Optimization (PSO) is a different algorithm discovered at some stage in requisite social model simulation, which is effectual in nonlinear optimization problems.

3. Proposed Fraud Detection System

In the proposed FDS, a number of steps are used to analyze the deviation of each incoming transaction from the normal profile of the cardholder. This paper works on shown fig.3.1. In this model dataset is attained from internet. After adding some deceitful transaction we call this dataset. To compare credit card fraud prediction using different techniques, we needed sets of transaction of both known deceitful and undetected or observed genuine transactions. Firstly convert the categorical data in to numerical. The dataset consist of name field which are converted to numerical integer values. The conversion is done by first evaluating the total number of words in a name. The string is converted into number using relation $[(\text{radix})^{\text{place value}} * \text{face value}) \bmod m$ by [12].

3.1 Support Vector Machine

The Support Vector Machine (SVM) was first anticipated by Vapnik and has since engrossed a high degree of curiosity in the machine learning research community [9]. Several recent studies have reported that the SVM (support vector machines) generally are competent of delivering higher concert in stipulations supervised learning methods used for classification. A special assets of SVM is, SVM at the same time diminish the empirical classification error and exploit the geometric margin. So SVM called Maximum Margin Classifiers. SVM is based on the Structural risk Minimization (SRM). SVM atlas input vector to a elevated dimensional space where a maximal separating hyper plane is constructed. Two equivalent hyper planes are constructed on each side of the hyper plane that detaches the data. The separating hyper plane is the hyper planes that exploit the distance between the two equivalent hyper planes.

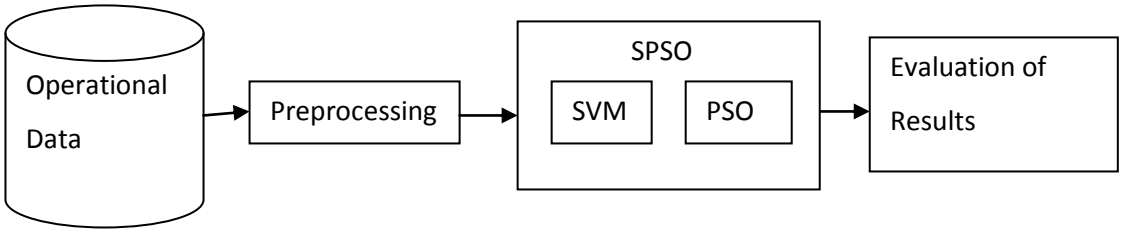

Fig.3.1: Mining and optimization of data

An postulation is made that the larger the margin or distance between these equivalent hyper planes the better the generalization error of the classifier. In SVMs, the classification function is a hyper-plane extrication the different classes of data.

$(W, x)+b=0$

The notation (w, x) characterizes the dot product of the coefficient vector w and the vector variable x. The solution to a classification problem is then precise by the coefficient vector w. The optimization problem obtainable in the foregoing section is difficult to solve because it depends on $\|\mathbf{w}\|$, which involves a square root. Providentially it is possible to alter the equation by

substituting $\|\mathbf{w}\|$ with $2/\|w\|$ without changing the solution.

$Yi\ (WX_i +b)>=1.$

For any i=1,2…….n. In doing so all the points which can be separated as

$Yi\ (WX_i +b)>=1$

do not matter since we must set the resultant to zero. This dilemma can now be solved by standard quadratic programming techniques and programs. Applications of SVM include bioinformatics, machine vision, text categorization, and time series analysis [23]. SVM are set of associatedThe "stationary" Karush–Kuhn–Tucker condition implies that the solution can be articulated as a linear combination of the training vectors $W=\sum\lambda_i.Y_i.X_i$.Only a few λ_i will be greater than zero. The solution of the above quadratic programming dilemma is a computationally rigorous task, which can be a limiting factor in using SVM with very large data. The consequent are exactly X_i the support vectors, which lie on the margin and satisfy

$Yi\ (WX_i -b) =0.$

From this one can originate that the support vectors also satisfy

$Yi\ (WX_i -b) =1/ Y_i$, and

$b= Yi-(WX_i)$. This allows one to define the offset b. In practice, it is more vigorous to average over all support vectors.SVM can applied to a number of provinces, together with handwritten digit identification, object recognition, and speaker identifications, as well as benchmark time series prophecy tests.

3.2 Particle swarm optimization

A metaheuristic is a set of concepts that can be used to define heuristic methods that can be applied to an ample set of different problems. In other words, a metaheuristic can be seen as a general algorithmic framework which can be applied to different optimization problems with moderately few modifications to make them adapted to a specific problem" [11]. Therefore instead of performing a boorish search over the complete fitness landscape, the metaheuristics guide the underlying heuristics to regions of better fitness according to the procedure defined by the metaheuristics. The common metaheuristics techniques are Genetic algorithm, Simulated Annealing and Tabu search, PSO, Scatter Search, Cuckoo Search, Ant colony Optimization, Local Search etc. In this paper PSO is used for search space optimization.

The Particle Swarm Optimization (PSO) algorithm, as one of the latest algorithms inspired from the nature, was introduced in the mid 1990s and since then, it has been utilized as an optimization tool in diverse applications, ranging from biological and medical applications to computer graphics and music composition. Kennedy and Eberhart [10], considering the behavior of swarms in the nature, such as birds, fish, etc. developed the PSO algorithm. The PSO has particles ambitious from natural swarms with communications based on evolutionary computations. PSO merges self-experiences with social experiences. In this Algorithm, a candidate solution is obtainable as a particle. It uses a collection of flying particles (changing solutions) in a search area (current and possible solutions) plus the movement towards a promising area in order to search out to a global optimum. Each particle keeps track of its coordinates in the solution space which are allied with the best solution (fitness) that has achieved so far by that particle. This value is called personal best (pbest***)***.

Another best value that is tracked by the PSO is the best value acquired so extreme by any particle in the neighborhood of that particle. This value is called global best (gbest).The basic concept of PSO lies in hastening each particle toward its pbest and the gbest locations, with a random weighted acceleration at each time. The modification of the particle's position can be mathematically modeled according the following equation:

$V_i= wV_{i-1} +c_1\ rand_1 x\ (pbest_i-S_{i-1}) + c_2\ rand_2 x\ (gbest-S_{i-1})$ ….. (1)

For update the velocity of a particle equation (1) is used.

w: weighting function

c_1, c_2: weighting factor,

$rand_1$,$rand_2$: uniformly distributed random number between 0 and 1

$pbest_i$: pbest of agent i
gbest: gbest of the group.
S_{i-1} : current position of agent i
V_{i-1} : velocity of agent i
The following weighting function is usually utilized in (1)
w = wMax-[(wMax-wMin) x iter]/maxIter ----- (2)
where wMax= initial weight,
wMin = final weight,
maxIter = maximum iteration number,
Iter = current iteration number.
$S_{i=} S_{i-1 +} V_i$ -------- (3)
$V_{i=}$ modified velocity
S_i =current searching position.
P_i =Particle best position.

Pso based local search procedure

1) Initialize the swarm by assigning a random position to each particle in the search space. For each particle i=1, 2… S. (S be the no of particle in the swarm.
a) Initialize the particle's best known position to its initial position.
b) $P_i \leftarrow S_i$
2) Appraise the fitness function for each particle.
a) If ($f(P_i)<f(gbest)$) then update the swarm's best known position.
3) Update the particle's velocity and position with the equation (1) and (3).
4) If [($f(S_i)<f(P_i)$)] then update the particle best known position
$P_i \leftarrow S_i$

5) If [($f(p_i)<f(gbest)$)] then update the particle best known position
gbest$\leftarrow P_i$
Now g holds the best solution.
6) Repeat step 3 to 5 until a stopping criterion is met (eg maximum no of iterations performed)
A large inertia weight (w) facilitates a global search while a small inertia weight facilitates a local search.PSO is effective in nonlinear optimization problems. It is easy to implement. Only a few input parameters need to be adjusted in PSO. Because the update process in PSO is based on simple equations, PSO can be efficiently used on large data sets. PSO has been successfully applied to many areas: function optimization, artificial neural network training, fuzzy system control.

PSO is effective in nonlinear optimization problems. It is straightforward to execute. Only little input parameters necessitate to be adjusted in PSO. As the update process in PSO is stand on simple equations, PSO can be competently used on large data sets. PSO has been fruitfully applied to a lot of areas: function optimization, artificial neural network training, fuzzy system control and other areas where GA can be applied [22].PSO is used with the outcome of svm so it's called SPSO.

4. Implementation and Results

This paper uses 7.8.0.347(R2009a) version of mat lab for implementation of data. Matlab is a high-level language and interactive environment for numerical computation visualization, and programming. Using Matlab, data can analyze, develop algorithms, and create models and applications. In this section various performance measures are defined with respect to the confusion matrix below in table 1[21], where Positive corresponds to Fraud cases and Negative corresponds to non-fraud cases. Accuracy: It represents the fraction of total number of transactions (both genuine and deceitful) that have been detected correctly.
True positive: It represents the fraction of deceitful transaction correctly identified as deceitful and genuine transactions correctly identified as genuine. False positive: It represents the fraction of genuine transaction identified as deceitful and deceitful transactions identified as genuine.

Table 1

	Predicted positive	**Predicted negative**
Actual positive	True positive (tp)	False negative (fp)
Actual negative	False positive (fp)	True negative (tn)

True Positive rate=TP/TP+FN
Accuracy=TP+TN/TP+TN+FP+FN
False Positive rate=FP/FP+TN

The performance of Support vector machine and particle swarm optimization on traditional measures is shown in Table2. We examine the performance of the different algorithms on four training dataset (data1, data2, data3, data4) having 100,200,500 and 700 transactions.
The traditional measures of classification, performance, however may not passably address performance requirement of specific applications.

Table 2

	svm	**spso**
Accuracy	81%	95%
TP rate	82%	95%
FP rate	20%	5%

In fraud detection ,cases envisaged as potential fraud are taken up for exploration or some further action, which involves a cost .Accurate identification of fraud cases helps avoid costs arising from deceitful activity, which are generally larger than the cost of inspecting a potential deceitful transaction. Fig 4.1 depicts the accuracy and tp rate are high than the fp rate, which shows better result performance. This table shows better results of spso as compare to svm. In this tp rate is high in case of spso as compare to svm.

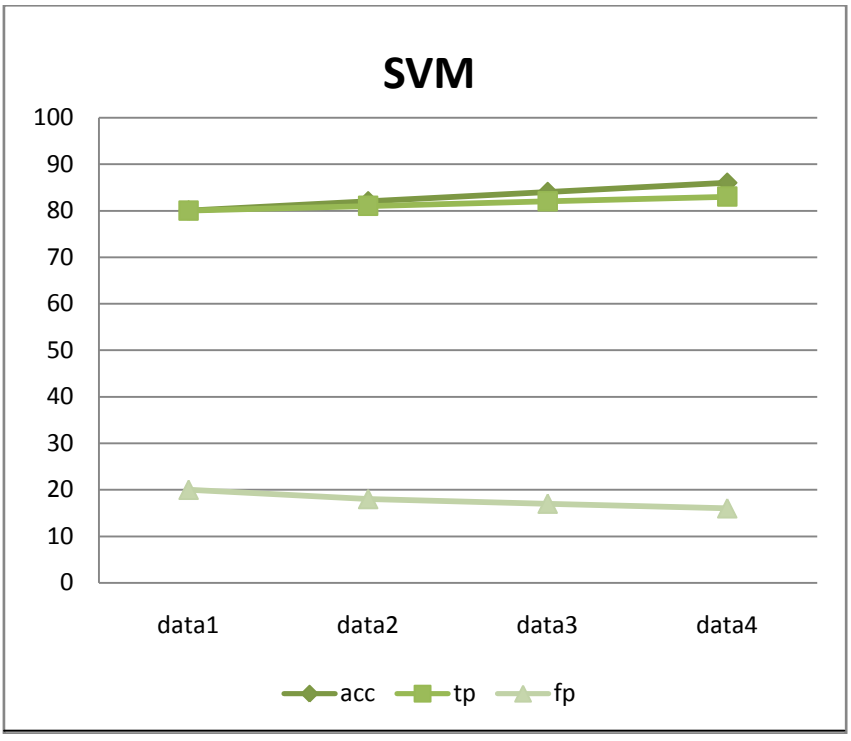

Fig. 4.1 Performance across Support vector machine

Fig 4.2 depicts the accuracy and tp rate are high than the fp rate, which shows better result performance. As the transaction increases in (data1) ds1 to (data4), the acc and tp also increases and fp decreases.

Fig 4.2 performance across particle swarm optimization

This graph shows the result of pso. In which different data are used for different employees. Data4 gives high performance as compare to data1.

Table 3: Comparison between SPSO and SVM in dataset1

	SPSO	**SVM**
Accuracy	92%	79%
TP Rate	84%	83%
FP Rate	13%	24%

Table 3 shows the difference in performance measures of dataset1 in case of SVM and SPSO. This is a comparison between these two techniques, which shows one is better than other. In this data set1 contain the thousands number of transactions. We get the performances measures like accuracy, TP rate and FP rate. On the basis of these measures a difference is shown between these two approaches. This table shows high accuracy and TP rate in both cases of SPSO and SVM.

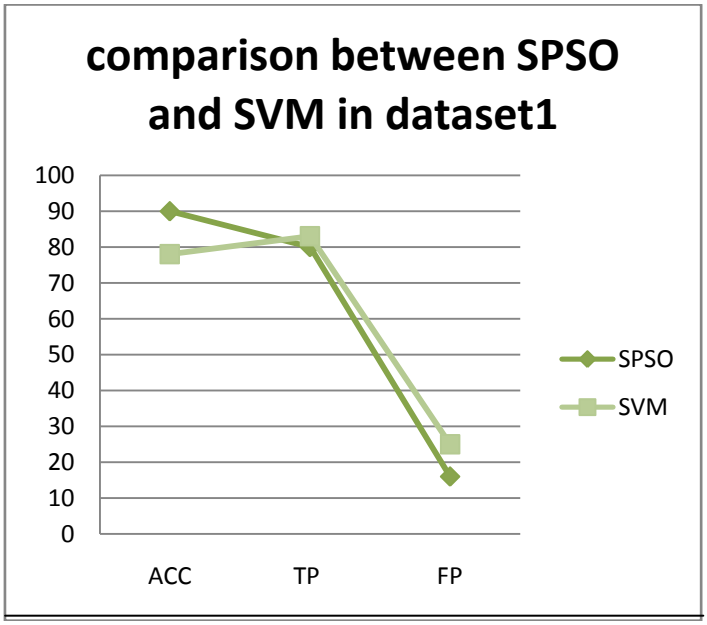

Fig. 4.3 Comparison between SPSO and SVM in dataset1

Fig 4.3 depicts the performance measures of dataset1 using the techniques SPSO and SVM. In this SPSO shows high accuracy and TP rate and low FP rate. Data set1 gives the high accuracy in SPSO i.e. 90% as compare to SVM which

gives only 78%.similarly TP rate is high in SPSO as compare to SVM. FP rate is low in SPSO, while in SVM is high. FP rate should be minimum. So we conclude that SPSO is better than the SVM.

Table 4: Comparison between SPSO and SVM in dataset2

	SPSO	**SVM**
Accuracy	90%	78%
TP Rate	80%	83%
FP Rate	16%	25%

Table 4 shows the difference in performance measures of dataset2 in case of SPSO and SVM. In which accuracy and TP rate is high in case of SPSO and FP rate is low. This table shows better results of SPSO than SVM. In this data set twelve hundred transactions are used which gives the high accuracy and TP rate as compare to data set1.SVM gives high FP rate i.e. 24% where SPSO gives 13% which shows best results. As accuracy and TP rate is increases as well as FP rate is decreases in both cases of SPSO and SVM. In case of SPSO accuracy is high as compare to SVM. SPSO gives 92% accuracy while SVM gives 79% which is lower than the SVM.
Similarly SPSO gives high TP rate as compared to SVM. SPSO gives 84% TP rate while SVM gives only 83%, which is slightly less than SPSO. In both cases of accuracy and TP rate SPSO gives high value as compare to SVM. So it is clear that SPSO gives better results as compared to SVM. In this dataset2 SPSO gives better results as compare to SVM.
If we compare both datasets i.e. dataset1 and dataset2 than we can easily says that SPSO is better as compare to SVM in both datasets. While dataset2 gives better result as compare to dataset1, because dataset2 gives high accuracy and TP rate as compare to dataset1 and low FP rate as compare to dataset1. Accuracy and TP rate should be high and low FP rate, which gives better results. SPSO fulfills this condition so it is better.

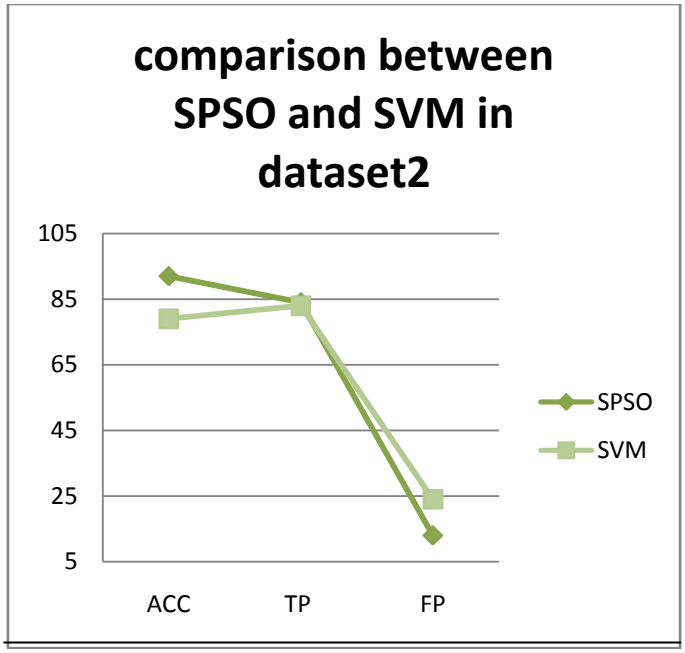

Fig. 4.4 Comparison between SPSO and SVM in dataset2

The graphs in Fig. 4.4 highlight the comparison between the techniques across dataset2, on various performance measures. These dataset contain the different number of transactions from previous data set. This data set contains the twelve hundred transactions. SPSO shows the high accuracy i.e. 94% as compared to SVM accuracy. TP rate of SVM is 83% while SPSO has 84%. There is small variation between the TP rate of SPSO and SVM. In case of TP rate both techniques gives similar results, but the basis of difference it is clear that SPSO is better than SVM.

Table 5 Comparison between SPSO and SVM in dataset3

	SPSO	**SVM**
Accuracy	94%	80%
TP Rate	89%	84%
FP Rate	12%	23%

Table5 depicts the differences in performance measures of dataset3 in case of SPSO and SVM. In which accuracy and TP rate is high in case of SPSO and FP rate is low. In this SVM shows the 80% accuracy which is lower than the SPSO.

Accuracy of SPSO is 94% which is much better than the SVM. There is a large variation in case of accuracy. SVM shows low TP rate as compare to other one. So we can easily say this table shows better results of SPSO than SVM.

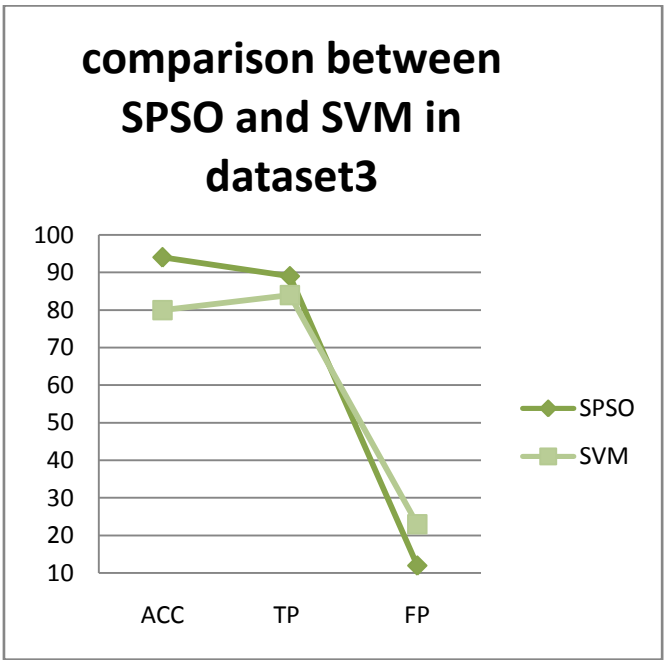

Fig. 4.5 Comparison between SPSO and SVM in dataset3

Fig 4.5 shows the Accuracy and TP rates are high and low FP rate, which shows better result performance in case of SPSO than SVM. This figure shows the performance measure differences in case of dataset3.The dataset3 contain more transaction as compare to previous two datasets. FP rate of SPSO is lower than the SVM. SPSO produce 12% FP rate, while SVM gives 24%.which is more than the SPSO. If FP rates is low than it give high accuracy which gives better results.

Table 6: Comparison between SPSO and SVM in dataset4

	SPSO	**SVM**
Accuracy	95%	81%
TP Rate	90%	84%
FP Rate	10%	23%

Table 6 shows the differences in performance measures of dataset4 in case of SPSO and SVM. In which accuracy and TP rate is high in case of SPSO and FP rate is low.

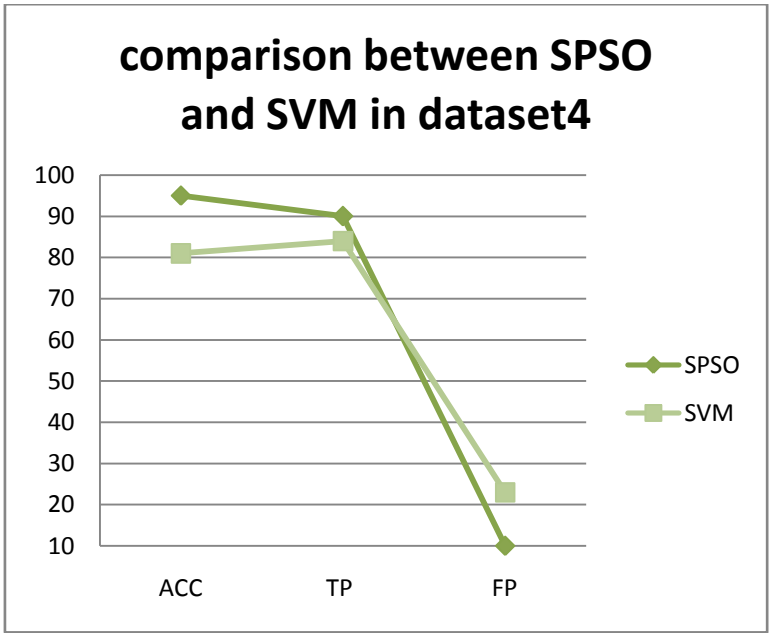

Fig. 4.6 Comparison between SPSO and SVM in dataset4

The graphs in Fig. 4.6 highlight the comparison between the techniques SPSO and SVM across dataset4, on various performance measures. These datasets contain the more number of transactions as compare to remaining datasets. FP rate of SVM is 23% while SPSO has 10% which is much lower than the SVM. Low FP rate gives better results. Accuracy of SPSO is 95% which is much better than the SVM and previous data sets results. So we can say that SPSO gives better results in each case of data sets.

Techniques (SVM and SPSO) showed adequate ability to model fraud in the considered data. Performance with different levels of under sampling was found to vary by technique and also on different performance measures. SPSO being computationally efficient and with only two adjustable parameter (Velocity and particle position) which can be set at commonly considered updated values, are also attractive from a practical usage standpoint.

5. Conclusion

Today anomaly detection methods are of major interest to the world and are used in very different and various domains like computer intrusion detection, credit card and telephone fraud detection, spam detection, and so on. Here, we have introduced a new supervised method for anomaly detection, based on a combination of a Support vector machine and Particle Swarm Optimization that fuse information from various sources. It is a simple, time and space consuming method that can be used in different domains. We compare the results of SVM for different datasets. Datasets having larger transactions shows better results. We optimize the results using metaheuristic technique i.e Particle Swarm Optimization. PSO is applied on the outcome of SVM that will optimize our results.

References

[1] Anural, S. and Christian, W.O. ‘‘Performance comparison of particle swarm optimization with traditional clustering algorithms used in self-organizing map’’, International Journal of Computational Intelligence, 5(1), pp. 32–41 (2009).

[2] Shelly Xiao an, Wu and Banshee, W. ‘‘The use of computational intelligence in intrusion detection systems: a review’’, Review Article Applied Soft Computing, 10(1), pp. 1–35 (2010).

[3] Xiaoping, Zhu ‘‘Semi-supervised learning literature survey’’, Computer Sciences TR 1530, University of Wisconsin–Madison, Last modified on July 19 (2008)

[4] C. Everett, Credit Card Fraud Funds Terrorism, Computer Fraud and Security, May,1, 2009.

[5] S. McAlearney, TJX Data Breach: Ignore Cost Lessons and Weep, CIO, August 07, 2008.

[6] R.J. Bolton, D.J. Hand, Unsupervised profiling methods for fraud detection, Conference on Credit Scoring and Credit Control, Edinburgh, 2001.

[7] Statistics for General and Online Card Fraud, 20 June, 2007. <http:// epaynews.com/statistics/fraud.html>.

[8] Online fraud is 12 times higher than offline fraud, 20 June, 2007. <http://sellitontheweb.com/ozone/news0434.shtml>.

[9] V. Vapid. The Nature of Statistical Learning Theory. NY: Springer-Vela. 1995.

[10] Kennedy, J. and Eberhart, R.C., (1995), “Particle Swarm Optimization”, Proc. IEEE Int. Conf. on N.N., pp. 1942-1948.

[11] Li, H.-Q., and Li, L., (2007), “A Novel Hybrid Particle Swarm Optimization Algorithm Combined with Harmony Search for High Dimensional Optimization Problems”, Proc. IEEE/IPC, pp. 94-97.

[12] Rajiv Arora, Pahma Payal ,Shubha “Alliance Rules for Data Warehouse Cleansing”, International Conference on Signal Processing Systems, IEEE Explore no. D01 10.1109/ICSPS.2009.133, pages 743-747.

[13] Chen, Q. and Aickelin, U. Dempster–Shafer for anomaly detection, Proceedings of the International Conference on Data Mining DMIN 2006, Las Vegas, USA, pp. 232–238 (2006).

[14] Wang, Gang, Hao, Jinxing, Ma, Jian and Huang, Lihua ‘‘A new approach to intrusion detection using artificial neural networks and fuzzy clustering’’,Original Research Article Expert Systems with Applications, 37(9), pp. 6225–6232 (2010).

[15] E. Aleskerov, B. Freisleben, and B. Rao, “CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection,” Proc. IEEE/IAFE: Computational Intelligence for Financial Eng., pp. 220-226, 1997.

[16] Ambareen, S., Susan, M. and Bridges, R.B.V. ‘‘Fuzzy cognitive maps for decision support in an intelligent intrusion detection system’’, National Science Foundation Grant# CCR-9988524 and the Army Research Laboratory Grant # DAAD17-01-C-0011.

[17] Suvasini Panigrahi a, Amlan Kundu a, Shamik Sural a, A.K. Majumdar “Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning”, Information Fusion 10,pp.354–363, 2009.

[18] Zaslavsky, V., & Strizhak, A. (2006). Credit card fraud detection using self-organizing maps. Information and Security, 18, 48–63.

[19] F. Provost, Comment on Bolton and Hand, Statistical Science 17 (2002) 249–251.

[20] R.J. Bolton, D.J. Hand, Statistical fraud detection: a review, Statistical Science 17 (3) (2002) 235–249.

[21] Siddhartha Bhattacharyya, Sanjeev Jha, Kurian Tharakunnel, J. Christopher Westland, "Data mining for credit card fraud: A comparative study", Decision Support Systems 50 pp. 602–613,2011.

[22] Xiang, X., Ernst, R.D., Russell, E., Zina, B.M. and Robert, J.O. ''Gene clustering using self-organizing maps and particle swarm optimization'',Ipdps, International Parallel and Distributed Processing Symposium IPDPS'03, p. 154b (2003).

[23] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines andOther Kernel-based Learning Methods, Cambridge University Press, 2000.

