

## CASH DEMAND FORECASTING FOR ATM USING NEURAL NETWORKS AND SUPPORT VECTOR REGRESSION ALGORITHMS

Rimvydas Simutis<sup>1,2</sup>, Darius Dilijonas<sup>2</sup>, Lidija Bastina<sup>3</sup>

<sup>1</sup>*Kaunas University of Technology, Studentų g. 48-327, 51368 Kaunas*

<sup>2</sup>*Vilnius University, Kaunas faculty, Muitinės g. 8, 44280 Kaunas*

<sup>3</sup>*JSC Penkių kontinentų bankinės technologijos, Kalvarijų g. 143, Vilnius*

*E-mail: <sup>1</sup>Rimvydas.Simutis@ktu.lt; <sup>2</sup>Darius.Dilijonas@vukhf.lt; <sup>3</sup>LidijaB@5ci.lt*

**Abstract:** In this paper two different methods are used to forecast the daily cash demand for automatic teller machines (ATM). The first method is based on flexible artificial neural network (ANN). The generalization properties of this ANN were improved using special adaptive regularization term. The second forecasting method employs the support vector regression (SPR) algorithm. Performed simulation studies and experimental tests showed tolerable forecasting capacities using the both proposed methods. Despite the today's overenthusiastic beliefs about the capabilities of SPR, our investigation showed however that for this application slightly better result can be achieved using forecasting method based on flexible ANN. At this stage the forecasting schema based on flexible ANN is in the implementing phase for intelligent cash management in ATM network.

**Keywords:** cash demand forecasting, automatic teller machines, neural networks, support vector regression.

### 1. Introduction

Automatic teller machines (ATM) are computerized telecommunication device which provide a financial institution's customers a method of financial transactions in a public space without the need for a human clerk. According the estimates developed by ATMIA (ATM Industry Association) the number of ATMs worldwide in 2007 was over 1.6 million (Snellman and Viren, 2006; Bounie and Francois, 2006).

As interest rate rises and greater operating efficiencies become paramount, more banks are turning their attention to driving greater efficiency in how they manage their cash in ATM networks (Westland, 2002; Drehmann *et al.*, 2002). Some banks typically maintain as much as 40 % more cash at their ATMs than what's needed, even though many experts consider cash excess of 15 % to 20 % to be sufficient. Cash-related costs represent about 35–60 % of the overall costs of running an ATM. Through currency management optimization, banks can avoid falling into the trap of maintaining too much cash and begin to profit by mobilizing idle cash. Therefore, it is very important to develop advanced algorithms to accurately predict currency demand for ATM. Based on cash demand forecasting an intelligent cash management system can then provide the bank the opportunities to lower its operational expenses and improve the return on its cash assets. In Lithuania, the ATM networks are expanding strongly in last time and the development of intelligent systems for monitoring and optimization of ATM networks becomes very relevant.

By the end of 2006 the JSC “Penkių kontinentų bankinės technologijos” received financial support from the EU structural funds for development of intelligent cash management and optimization system for an ATM network. In this paper we present some results in solving this task.

This paper is structured as follows. After short introduction in this section, the problem formulation and existing approaches are discussed in section 2. In section 3, cash demand forecasting model, based on flexible artificial neural network is presented and in section 4, support vector regression algorithm is introduced. Section 5 describes the experimental set-up, also some experimental results are presented and analyzed. Finally, the main results of this work are discussed in section 6, followed by conclusions and future work.

## 2. Existing approaches for ATM cash demand forecasting and management

The basic element in development of efficient ATM cash management system is a cash demand forecasting model for every ATM. Generally this forecasting model is created based on historical cash demand data. The historical cash demand for every ATM varies with time and is often overlaid with non stationary behaviour of users and with additional factors, such as paydays, holidays, and seasonal demand in a specific area. Cash drawings are subject to trends and generally follow weekly, monthly and annual cycles. For example, people tend to draw relatively large sums of cash at the beginning of each month. Before Christmas, drawing rates soar, whereas in August, during the summer holidays, rates tend to drop considerably. ATMs that are located in shopping centres, for example, are most heaped on Fridays and Saturdays. Typical example of cash demand for real ATM during three years time interval is presented in figure 1.

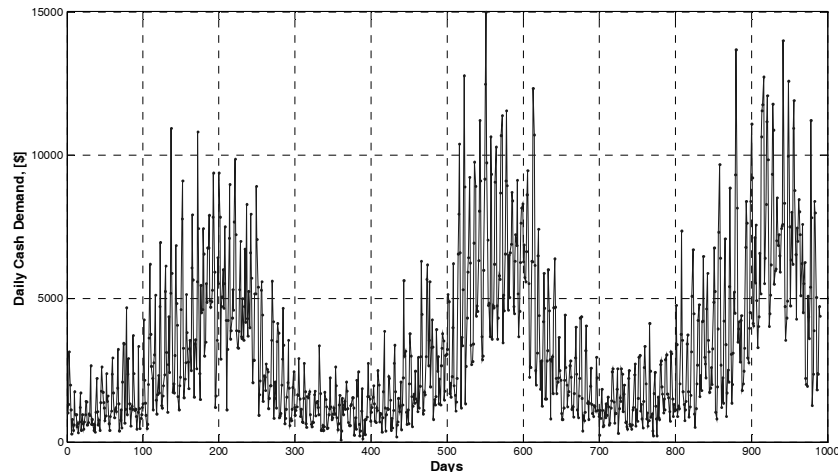


Fig. 1. Cash demand for real ATM during three years time interval

The development of the forecasting model is complicated procedure, because it must consider the changing behaviour of users and various input variables.

Based on the cash demand forecasting model the optimization procedure determines the optimum cash amount for each ATM by calculating the transport and money upload costs against interest rates. Cash management system has to guarantee the availability of cash in the ATMs network, should estimate optimal amount of stocked money plus efficiently manage and control day-to-day cash handling, transportation with reducing of currency transportation and servicing costs. The system should be flexible enough to allow the bank to re-forecast future demand, perform *WHAT – IF* analyses, and optimize the network as the cash distribution environment evolves. Most known cash management systems for ATM network are iCom (Carreker Corporation), MorphisCM (Morphis, Inc), OptiCa\$h (Transoft International), Pro Cash Analyser (Wincor Nixdorf), (Simutis *et al.*, 2007). The forecasting algorithms involved in these products have the following common drawbacks:

- Cash demand forecast for every ATM generally is based on linear regression models with seasonality coefficients. The development of such models is relatively complicated and differs for various ATM. Therefore preparation of forecasting models for whole ATM network is difficult task for owners of machines.

- The parameters of forecasting models are determined in the system implementation stage and are hold constant during the operation phase. On the other hand, business environment changes continually in a real world and, consequently, the model parameters must be also adapted to the changing environment. To eliminate these weaknesses, we propose to use forecasting methods based on flexible artificial neural networks and support vector regression algorithms. The functioning principles of these methods are discussed below.

## 3. A flexible neural network for cash demand forecasting

Artificial Neural Networks are universal and highly flexible function approximators first used in pattern recognition, classification and time series forecasting (Haykin, 1999; Bishop, 2006; Nrgaard and

Norgaard, 2006). In recent years, ANN becomes increasingly popular in financial markets. The key to all forecasting applications is to capture and process the historical data so that they provide insight into the future. In this paper, we are concentrating on application of artificial neural networks for cash demand forecasting problem. The general idea behind the ANN is to allow the network to map the nonlinear relationships between various factors affecting the cash withdrawal and the actual cash demand. Once this relationship between inputs and outputs is identified, it gives the output variable – cash demand forecast using values of various input variables.

One of the most important components in the success of neural network solution is the structure of the ANN and the data necessary to train the network. In this study, we used simulated data and real data for training and evaluation of the artificial neural networks. For every ATM machine a separate three-layer feed-forward neural network was designed. The neural network was trained using Levenberg-Marquardt optimization method and RMS (root mean square) error between predicted and real value. Regularization term was also included in the training criterion (Haykin, 1999; Bishop, 2006). The input variables for ANN were coded values of weekday, day of the month, month of the year, holiday effect value and average daily cash demand for ATM in last week. The output variable of ANN was cash demand for the ATM for the next day. For simplification purpose the ANN structure for all ATM in the network was chosen the same (the same inputs and the same number of hidden units in ANN). The number of neurons in the hidden layer was chosen relative big (15 hyperbolic tangents neurons in hidden layer). Such neural network can approximate very complicated relationships between input output variables, but the generalization properties of neural network can be very pure. Therefore we proposed a special flexible neural network design procedure for cash demand forecasting for every local ATM. The realization of the proposed procedure is executed in the following steps:

- 1) Assemble input-output data from every local ATM (historical data from 2–3 years is necessary for reliable training of ANN).
- 2) Divide assembled data in training (70 %) and testing sets (30 %).
- 3) Train ANN using Levenberg – Marquard optimization method using various values of regularization term  $D$ , which is included in the training criteria

$$E = \frac{1}{2N} \cdot (y_d - y)^2 + \frac{1}{2N} \mathbf{w}^T \cdot \mathbf{D} \cdot \mathbf{w}, \quad \mathbf{D} = \alpha \mathbf{I}, \quad (1)$$

$y_d, y$  – desired and real output,  $\mathbf{I}$  – unit matrix,  $\alpha$  – regularization term,  $N$  – number of data.

- 4) Estimate the normalized sum square error (NSSE) of ANN for test data set.
- 5) Choose the regularization term  $\alpha$  which gives the minimum of NSSE in test data and use it as optimal regularization term.
- 6) Repeat ANN training on whole data sets using optimal regularization term and use this ANN as basis for cash demand prediction.
- 7) As the additional portion of new data about functioning of local ATM is available (typically in one week), repeat the steps 2–6 with less number of training iteration.
- 8) Use the fresh adapted ANN for cash demand prediction in chosen time interval.

The proposed algorithm adapts the ANN parameters (weights) following the new observation; therefore the designed ANN is always tuned to the current situation observed in the business environment. Having the models to forecast the daily (or weekly) cash demand for every ATM, it is possible to plan and to optimize the cash loads for a whole ATM network.

#### 4. Cash demand forecasting using Support Vector Regression method

Support vector machine (SVM) is a new and promising technique for data classification and regression (Vapnik, 1998).

In this section we briefly introduce support vector regression (SVR), which can be used for cash demand prediction in ATM machines. The basic idea of SVM is to map the linear non-separating training data from the input space into a higher dimensional feature space via special function  $\Phi$  and then construct a separating hyperplane with maximum margin in the feature space. Consequently, although we cannot determine linear function in the input space to decide what class the given data is, we can easily determine a hyperplane that can discriminate between two classes of data. Support vector regression is special modification of support vector machines techniques dedicated for solving of regression problems. Given training data  $(x_1, y_1), (x_2, y_2) \dots (x_l, y_l)$ , where  $x_i$  are input vectors and  $y_i$  are the associated output value of  $x_i$ , the support vector regression solves an optimization problem:

$$\begin{aligned}
 & \min_{w, b, \xi, \xi^*} \quad \frac{1}{2} w^T w + C \sum_{i=1}^l (\xi_i + \xi_i^*) & (2) \\
 & \text{subject to} \quad y_i - (w^T \Phi(x_i) + b) \leq \varepsilon + \xi_i, \\
 & \quad \quad \quad (w^T \Phi(x_i) + b) - y_i \leq \varepsilon + \xi_i^*, \\
 & \quad \quad \quad \xi_i, \xi_i^* \geq 0, i = 1, \dots, l,
 \end{aligned}$$

where  $x_i$  is mapped to a higher dimensional space by the function  $\Phi$ ,  $\xi_i$  and  $\xi_i^*$  are the upper and lower training errors subject to the  $\varepsilon$  – insensitive tube  $|y - (w^T \Phi(x) + b)| \leq \varepsilon$ ,  $w$  – model parameter vector. The parameters which control the regression quality are the cost of error  $C$ , the width of the insensitive tube  $\varepsilon$ , and the mapping function  $\Phi$ . These parameters must be set by user. The constraints in equation (2) imply that it is necessary to put most data  $x_i$  in the tube  $|y - (w^T \Phi(x) + b)| \leq \varepsilon$ . If  $x_i$  is not in the tube, there is an error  $\xi_i$  or  $\xi_i^*$  which we would like to minimize in the objective function. Graphical illustration of  $\varepsilon$  – insensitive loss function is shown in Fig. 2.

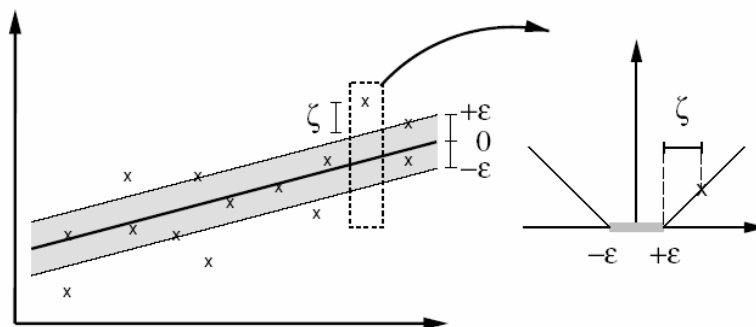


Fig. 2. Graphical illustration of  $\varepsilon$  – insensitive loss function

SVR technique avoids underfitting and overfitting the training data by minimizing the training error  $C \sum (\xi_i + \xi_i^*)$  as well as the regularization term  $\frac{1}{2} w^T w$ . In case of using classical least-square regression technique  $\varepsilon$  is always zero and original data are not mapped into higher dimensional spaces.

Equation (2) can be minimized by solving a quadratic programming problem, which is uniquely solvable. This is very important characteristic of support vector regression techniques, because the training of a SVM involves only the solving a quadratic optimisation task, which has one unique solution and does not involve the random initialisation of weights as training ANN does.

There are a lot of public available software libraries for realization of SVR methods. We use well known support vector machine public software library LIBSVM (Chang and Lin, 2001) in our experimental investigations. When training an SVR model, user must choose some important parameters. They would influence the performance of an SVR model. In order to get a satisfactory model, these parameters need to be selected properly. As we already mentioned the most important parameters are: mapping function, cost of error  $C$  and the width of the  $\varepsilon$  – insensitive tube. Surprisingly, no explicit construction of the nonlinear mapping function  $\Phi$  is needed. Mapping function in SVR algorithms is replaced by kernel function employing so called kernel trick (Vapnik, 1998; Suykens *et al.*, 2002). Kernel functions enable the dot product of transformation  $\Phi(x_i)^T \Phi(x_j)$  to be performed in high-dimensional feature space using low-dimensional space data input without knowing the transformation  $\Phi$ . All kernel functions must satisfy Mercer’s condition that corresponds to the inner product of some feature space. In the first investigation phase we tested various popular kernel functions: linear, polynomial and radial basis function for cash demand forecasting algorithms. Later, in accordance with prediction quality, we used only radial basis function (RBF) in our algorithms. The RBF function is one of the most commonly used kernel function in SVR technique. It has the property that  $\Phi(x_i)^T \Phi(x_j) = e^{-\gamma \|x_i - x_j\|^2}$ , where parameter  $\gamma$  has to be tuned together with parameters  $C$  and  $\varepsilon$ .

## 5. Experimental set-up and experiments

To test the possibilities of artificial neural network to forecast the ATM's cash demand simulated data and experimental data from 15 real ATMs were used. Artificial neural network models and SVR models were trained using data records from 2 years. Cross-validation technique was used to determine suitable values of regularization term  $D$  for ANN models and  $C$ ,  $\varepsilon$  and  $\gamma$  values for SVR models. After that, the models were used to forecast the daily cash demand for every ATM. We estimated mean average proportional error (MAPE) of daily cash demand forecasting for following 50 days. Table 1 presents the results of this investigation. MAPE of daily cash demand prediction for simulated environment was very good: 0.76 % using flexible ANN model and 4.1 % using SVR model. Forecasting results for real ATMs were of poorer quality. MAPE using flexible ANN models varied between 15–28 % and for SVR models between 17–40 %.

**Table 1.** Forecasting quality (MAPE) of daily cash demand for simulated and real ATMs

ATM number	MAPE, % ANN Model	MAPE, % SVR Model
Simulated data	0.76	4.1
Nr. 1	28.3	40.9
Nr. 2	16.7	27.6
Nr. 3	19.1	25.4
Nr. 4	15.6	17.1
Nr. 5	28.8	28.2
Nr. 6	24.1	25.9
Nr. 7	23.8	28.0
Nr. 8	20.8	26.9
Nr. 9	19.4	24.3
Nr. 10	19.1	20.6
Nr. 11	16.1	21.1
Nr.12	20.3	18.27
Nr.13	19.8	20.1
Nr.14	17.2	29
Nr. 15	20.9	23.8
Mean, $\sigma$ (Nr.1–15)	M=20.6, $\sigma$ =4.0	M=25.1, $\sigma$ =5.7

The obtained forecasting results are in some confrontation with the today's opinion about the possibilities of SVR techniques. Nowadays the SVM/SVR is assumed to be as "next generation" technique and some sort of "panacea" for classification and forecasting tasks. Despite of our extensive attempts to find the best parameters for SVR model our experimental results could not confirm this opinion. In the future we are planning to make more extensive studies to improve the quality of SVR models.

## 6. Conclusions

The proposed daily cash demand forecasting methods for automatic teller machines (ATM) showed tolerable forecasting quality. Despite the today's overenthusiastic beliefs about capabilities of SPR, our investigation showed, that for this application slightly better result can be achieved using forecasting method based of the flexible ANN. In the future we are planning to make more extensive investigations to improve the quality of the SVR model. At this stage the forecasting schema, based on flexible ANN, is in implementing phase to enrich the professional cash management software ASOMIS, developed by JSC "Penkių kontinentų bankinės technologijos" group.

## Acknowledgments

This research work has been supported by means of EU structural funds (Project No. BPD2004-ERPF-3.17-06/0045).

**References**

- Snellman, H.; Viren, M. 2006. ATM networks and cash usage, *Research Discussion Papers. Bank of Finland*, 21: 1–33.
- Bounie, D.; Francois, A. 2006. *Cash, Check or Bank Card: The effect of transaction characteristics on the use of payment instruments*. Telecom Paris, Working papers in Economics and Social Sciences, ESS-06-05.
- Westland, Chr. J. 2002. Preference ordering cash, near cash and electronic cash, *Journal of Organizational Computing and Electronic Commerce* 12(3): 223–242.
- Drehmann, M.; Goodhart, C.; Krueger, M. 2002. The challenges facing currency usage: will the traditional transaction medium be able to resist competition from the new technologies? *Economic Policy* 17(34): 195–227.
- Simutis, R.; Dilijonas, D.; Bastina, L.; Friman, J.; Drobinov, P. 2007. Optimization of cash management for ATM network, *Informacinės technologijos ir valdymas* [Information technology and control]. Kauno technologijos universitetas. Kaunas, 1A: 117–121. ISSN 1392-124X.
- Haykin, S. 1999. *Neural networks: a comprehensive foundation*. Prentice-Hall, Inc.
- Bishop, C. M. 2006. *Pattern recognition and machine learning*. Springer.
- Nrgaard, M.; Norgaard, P. M. 2006. *Neural Networks for Modelling and Control of Dynamic Systems: A Practitioner's Handbook (Advanced Textbooks in Control and Signal Processing)*. Springer.
- Vapnik, V. 1998. *Statistical Learning Theory*. Wiley, New York, NY.
- Suykens, J. A.; Gestel, T.; Brabanter, J.; Moor, B.; Vanderwalle, J. 2002. *Least squares support vector machines*. World Scientific, New Jersey.
- Chang, C. C.; Lin C. J. 2001. *LIBSVM: a library for support vector machines*.