99^{th} European Study Group with Industry

Optimization of ATMs filling-in with cash

Problem presented by

Credit Agricole Bank, Serbia

Executive Summary

This report presents an approach for modeling daily cash demand for all ATMs in the Credit Agricole Bank network in Serbia. The approach is based on time series and regression methods for forecasting an optimal amount of money that should be placed daily in the ATMs in order to meet customers' demands and mimimize costs of the bank. Three different types of costs were considered: cash freezing costs, transportation costs and insurance costs. The performance of the resulting forecasts were compared with results of the application that bank uses for prediciton of the time and the amount of filling-in for each ATM based on historical data.

Report author

Piotr Broda, Tijana Levajković, Milena Kresoja, Marina Marčeta, Hermann Mena, Mladen Nikolić, Tijana Stojančević

Contributors

Piotr Broda (University of Warshaw),
Tijana Levajković (University of Belgrade),
Milena Kresoja (University of Novi Sad),
Marina Marčeta (University of Novi Sad),
Hermann Mena (University of Innsbruck),
Mladen Nikolić (University of Belgrade),
Tijana Stojančević (University of Novi Sad)

ESGI99 was jointly organised by
Department of Mathematics and Informatics, Faculty of Sciences,
University of Novi Sad
European Consortium for Mathematics in Industry

Contents

1 Introduction

Automated Teller Machines (ATMs) are 24-hour self-service machines that enable bank customers conducting their financial transactions without visiting the bank branch. In spite of online banking facilities expansion, need for ATMs transactions remains high over years and makes ATMs an irreplaceable devices in everyday life. In order to meet growing cash needs of bank clients, banks have to increase continually the number of their ATMs in different location to make cash available. The number of ATMs in the world in 2013 was over 2.2 million. In the last years phenomenon of filling-in ATMs with cash has attracted many researchers, but ATMs transactions are far from being completely analyzed. ([2, 3, 4]) Banks' basic challenge in this liquidity cash management framework is to predict amount of money which must be supplied to each ATM to satisfy the customer unknown demand. While forecasting cash demand, an important issue must be kept in mind: serving ATMs network is everything but cheap mission. There are three major types of costs involved in serving ATMs network: cash freezing costs, transportation costs and insurance costs.

This kind of problem statement was proposed to ESGI 99 by Credit Agricole Bank in Serbia. The Bank operates on the whole territory of Serbia and has total of 81 branches in the country. In order to make cash withdrawals accessible per principle 24x7, the Bank supplies ATMs with cash, per need. Figure 1 shows a sketch of ATMs network in Novi Sad.

Bank operates with different types of ATMs, based on their activity and location. According to their activity, ATMs can be active or currently inactive. For our problem, location of ATMs is crucial. Thus, it is important to make distinguish between internal (inside of the branches) and external (outside of the branches). External ATMs can be isolated or clustered into sectors, where each sector consists of 2, 3, 4 or 5 closely located ATMs. In order to keep costs of transports as low as possible, clustered ATMs are being sup-

Figure 1: Credit Agricole Bank ATMs network in the Novi Sad

plied with cash in the same time (route). Consequently, important issue to be considered beside amount of money that needs to be placed in each ATMs is frequency of filling-in ATMs. Of course, Bank wants to avoid situations when ATMs are out of the cash.

While supplying ATMs with cash, Bank faces with minimizing of total costs. Total costs are consisted of 3 basic parts:

1. cost for unwithdrawn cash in the ATM itself (cash freezing);
2. cost for transport from the branch to the ATM;
3. cost for insurance of the cash in the ATM.

There's negative correlation between frequency of filling-in ATMs and costs of cash freezing. If ATMs are often supplied with cash the costs of cash, freezing are lower and if they are rarely filled-in, costs for cash freezing are higher.

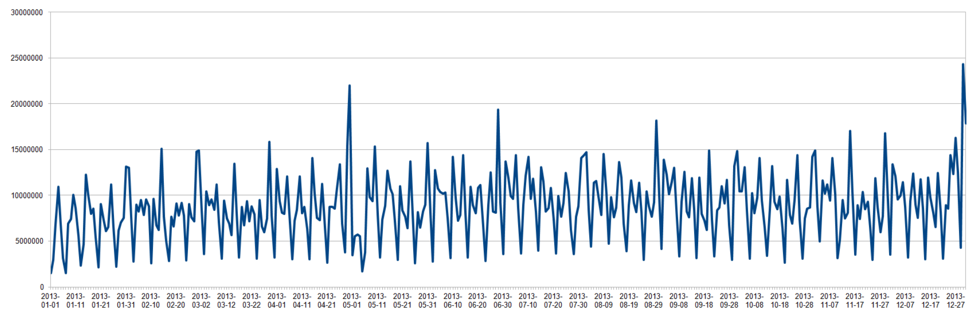

Figure 2: ATMs daily withdrawals

Furthermore, costs for insurance are proportional depending on amount of money in ATM. When it comes to the cost of transport, the Bank deals with the situation on the following way. In order to decrease costs for filling-in, Bank always organizes cash transport for multiple subjects in the neighborhood, therefore the transport is done in the loop (for 3-4 subjects) and not as star (for each subject separately). The Bank uses an application which proposes the time and the amount of fling-in for each ATM based on historical data.

The Bank is seeking for approach for minimizing total costs and would like to have an optimal balance of costs. The frequency of ATM's filling-in should be in function of the needs with minimal cash freezing and of course, to avoid ATM to be out of cash especially in critic periods (ex. New Years Eve, Christmas and Easter holidays..).

Data used in the report were provided by the Bank and consists of daily ATMs transactions on total 40 ATMs, out of which 20 are the top used and 20 the least used ATMs. Transactions were made in 2013. Unfortunately, only data sets for 3 external and 12 internal ATMs were useful for the analysis.

Figure 2 presents cash withdrawals time series for 2013 year. It would be useful to identify calendar effects in ATM cash transactions series. Time

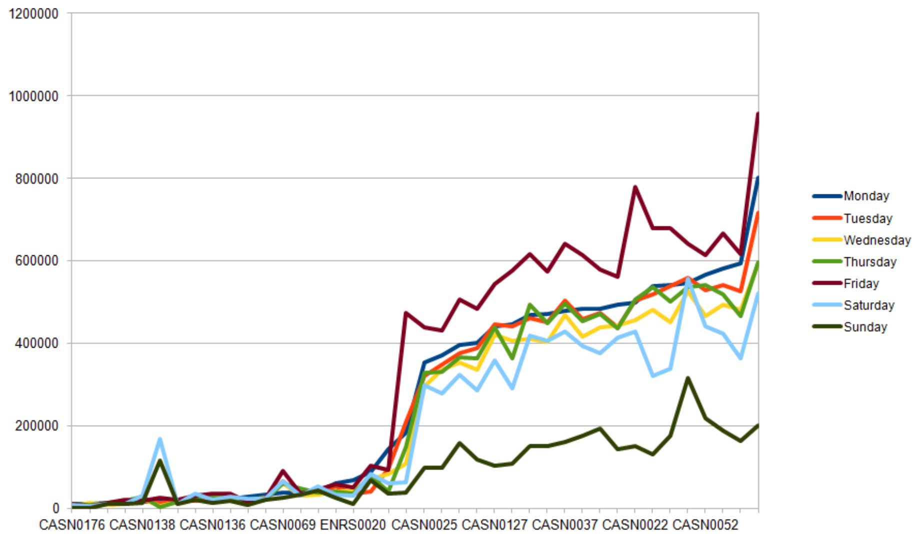

Figure 3: Average value of cash withdrawals for diffrent days of the week in 2013

series trend and volatility given in graph prevent us to identify seasonality. Using available data, it is observable that cash withdrawals are not same for all days of the week. Mondays and Fridays are the days in which the largest amounts are withdrawn and Saturdays and Sundays the days with the lowest withdrawn amounts(3). Important and interesting part in daily ATM withdrawals analysis is also identifying monthly and holidays seasonality. In particular, we could expect summer and winter seasons to be the periods of highest withdrawn amount.([4])

The report is organized as follows. Section 2 is referred to modeling cash demand. It presents time series and regression models for daily cash demand predictions at external ATMs and theoretical proposal for predictions at internal ATMs. Section 3 is devoted to determining optimal quantity of cash to be stored in ATM based on the short term prediction from section 2. Last section 4 discusses the decision making process.

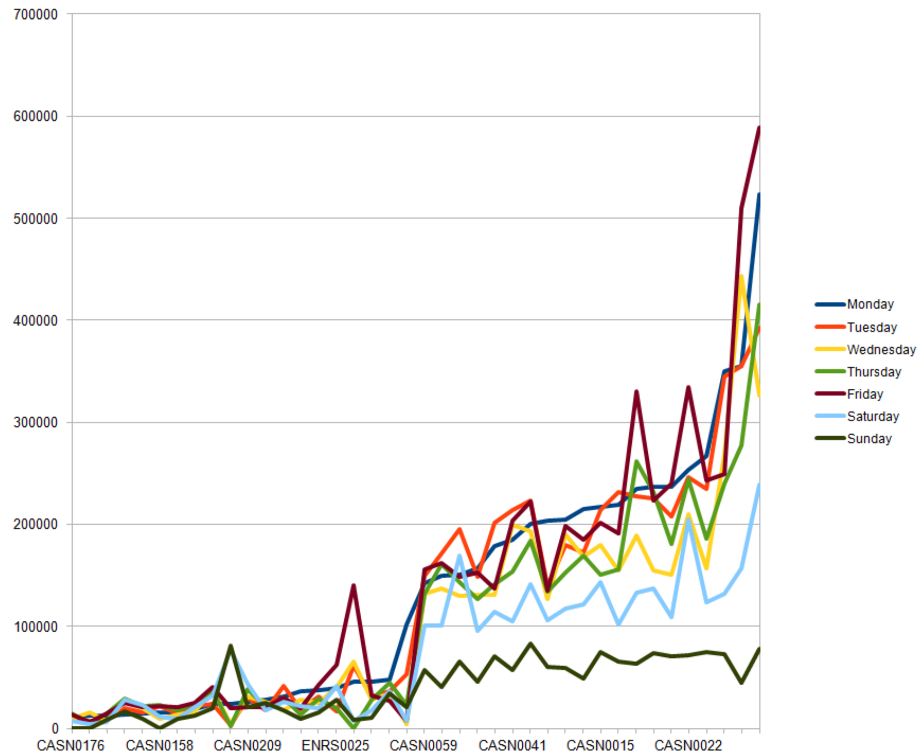

Figure 4: Standard deviation of cash withdrawals for diffrent days of the week in 2013

2 Modeling Daily Cash Demand

Probably the most natural approach to model daily cash demand is to fit time series models which would allow forecasting of future demand. We tried several such approaches. Another reasonable approach that we explored is to model cash demand by a linear model based on various statistics of demand history as predictors.

2.1 Time Series Analysis

Daily cash demand represents a natural time series data. However, it can be organized in time series in different ways. We considered the following ways to organize the data into time series:

- Demand values for all days in the considered period are used to form a single time series. This is the simplest way to arrange the data and therefore a natural first choice.

- Separate time series were formed, consisting of demand values on specific week days. It is reasonable to assume that the demand behavior on Friday, which precedes the weekend, is distinct from Wednesday which is in the middle of the week. By modeling these behaviors separately, one might hope to avoid complex behavior in the data obtained by merging several simpler ones.

- Separate time series were formed for days that comprise the first, second, third and fourth week in a month (e.g. all days of first week of January, first week of February, first week of March, etc. are used to form one time series). This approach aimed at capturing dynamics within different parts of the months which may differ due to factors like paycheck arrival and similar.

- Three time series were formed by different resolution of data aggregation — daily, weekly, and monthly time series. The forecasts of these three would be averaged to provide a prediction, the aim being to provide more robustness to simple daily estimates.

The methods used to model time series include popular seasonal ARIMA and exponential smoothing models. All of these are implemented in `forecast` package of `R` statistical computing environment which provides model fitting and forecasting capabilities including fitting all admissible models and returning the forecast provided by the best one. ([1]) The problem with all approaches explored was that the lengths of confidence intervals for all forecasted values were order of magnitude larger than the values themselves, indicating that even though the values exhibited similar patterns as observed in the data, the obtained forecasts were to unreliable to be of practical importance. Therefore, this direction was not pursued further.

2.2 Linear Regression

Predicting cash demand can also be approached as a regression problem. We considered linear model of form:

$$f(d) = \sum_i \alpha_i h_i(d, l_i)$$

where α_i are the parameters of the model to be estimated from the data, and h_i are different statistics of cash demand history. The model is evaluated for specific day d. Values l_i are metaparameters which govern the computation of statistics h_i, like the length of the history considered. Specifically, we considered the following statistics:

- The average recent demand history (h_1) represents the average of cash demand in l_1 days prior to given day d.
- The average weekly history (h_2) represents the average of cash demand on the day corresponding to d in l_2 weeks prior to the one d belongs to. For instance, if d is a Monday, h_2 could be an average of cash demand on previous 3 Mondays.
- The average monthly history (h_3) represents the average of cash demand on the day corresponding to d in l_3 months prior to the one d belongs to. For instance, if d is 15th of the current month, then h_3 could be an average of cash demand on 15th days of previous three

months. In case that previous months do not contain the corresponding day (e.g., if d is 31st of July, the problem arises since there is no 31 of June), the information on cash demand on the last day of that month is used instead.

- The average yearly history (h_4) represents the average of cash demand on the day corresponding to d in l_4 years prior to the one d belongs to. For instance, if d is a Christmas in the current year, h_4 could be an average of cash demand on three previous Christmases.

The estimation of parameters α_i was performed by ordinary least squares method. In case of strongly correlated predictors, ridge regression would be advised, but our data exhibited no such behavior. We performed experiments by fitting the proposed linear model to the data for various values of metaparameters l_i. Specifically, we used $l_1 \in \{1, \dots, 30\}$, $l_2 \in \{1, \dots, 52\}$, $l_3 \in \{1, \dots, 12\}$, and $l_4 = 1$. The last choice was forced by the amount of the data at our disposal. For all models roughly described by constraints $l_1 > 20$, $l_2 > 20$, and $l_3 = 12$, the coefficient of determination R^2 (a standard measure of goodness of fit for linear models, bounded in the interval [0,1]) was approximately 0.4 with small differences. No model achieved lower than 0.32. On the other hand, to our knowledge, current Bank practices exploit only the information on average yearly history. Using only such information, one obtains $R^2 = 0.226$. Even though there is a lot of room for improvement to the proposed model, this clearly shows that there is useful information in the data which is still not exploited. Since the obtained results are promising and indicate that there is the information in the data unused by current Bank practices, it is reasonable pursue further improvement of the regression based cash demand prediction. We envisioned several directions of research:

- Formulating new statistics (h_i) to be used as predictors. Four statistics used in our experiments were the first reasonable guesses which were based on the idea of capturing recent activity and three different kinds of seasonality. Based on the expert knowledge and experience of Bank's employees, it is probably possible to formulate more reasonable statistics which would be correlated with ongoing cash demand.

- Incorporation of outside knowledge into the model (e.g., a sports event is announced in the vicinity of the ATM) by adding new α parameters for such predictors. One would need to provide such historical data in order to estimate the corresponding parameters.

- Replacing the proposed model with a family of models, each corresponding to different parts of the year. For instance, a separate model could be estimated for the days of each month (e.g., the parameters for January model would be trained only on data consisting of cash demand on days of several previous Januaries).

- Formulating the model relying on nonlinear transformation of predictors or a nonlinear model. Use of such models would be justified if previous efforts fail in achieving significant boost of prediction quality over the model already evaluated.

2.3 Compound Poisson Process

Although Bank's representatives suggested focusing only on withdrawals at external ATMs, in this subsection we propose simple model for forecasting cash demand at internal ATMs in theoretical manner. Motivation for this was found in [2]. Basic idea was to simulate customer's uncertain demand using some stochastic processes. Capturing randomness of our problem requires modeling the following processes:

- number of customers arrived at ATM at the time interval $(0, t)$;

- total amount withdrawn in the time interval $(0, t)$.

First process (numbers of ATM users during each day) represent a situation when outcome variable is a count. Such variable takes on a limited number of values and are never negative. Additionally, their mean and variance are often related (which isn't true for normally distributed variables). Common approach for describing counting processes is to use a Poisson process. Poisson process is one of the most important models used in queuing theory. It is a simple stochastic process used for modeling the times at which arrivals enter a system.

Formally, the Poisson process is described by the so called counter process N_t. The counter tells the number of arrivals that have occurred in the interval $(0, t)$ i.e:

$$N_t = \text{number of ATM's customers in the interval } (0, t).$$

We assume that customers arrive at ATM at the times of a Poisson process with a rate λ per day and consider day as a time unit. Parameter λ is the average of withdrawals made of ATM customers in a day. Upon this assumption, important properties of the process can be derived. The most important one is that the number of arrivals N_t in a finite interval of length t obeys the Poisson distribution of parameter λt which means that we are able to calculate the probability of n ATM customers in the time t by formula:

$$P\left\{N_t = n\right\} = \frac{(\lambda t)^n}{n!} \exp(-\lambda t).$$

Moreover, the number of arrivals in intervals (t_1, t_2) and (t_3, t_4) in non-overlapping intervals $(t_1 \leq t_2 \leq t_3 \leq t_4)$ are independent. Mean and variance of N_t are both equal to λt.

Assuming that the successive withdrawn amounts are independent and identically distributed random variables, we can describe the other process of our interest (total amount withdrawn in the time interval $(0, t)$) by using generilized Poisson process.

Let A_t be the total amount withdrawn in the time interval $(0, t)$ and N_t the number of customers to come to the ATM in the interval $(0, t)$ as described above. If we denote the amount of the n^{th} withdrawal with A_n, then A_t can be represented as the following random sum of random variables:

$$A_t = \sum_{i=1}^{N_t} A_n.$$

Process $\{A_t\}$ is called a compound Poisson process. Using this assumtion we are able to derive following characteristics of the process:

$$E[A_t] = \lambda t E[A_1]$$

and

$$var[A_t] = \lambda t E[A_1^2].$$

3 Optimization of ATM Filling

Assuming a given estimation of demand for a single ATM (or a group of ATMs that is treated as a single ATM for purposes of filling-ins) one can find an exact optimal solution of costs associated with filling-ins in a given period. A simple algorithm which provides such a minimal cost executes an exhaustive search of solution space, which is feasible if the optimized period is small enough (e.g. 20 days). In its most basic form the algorithm considers total costs of every combination of days at which the fills occur, which makes its complexity exponential. As such, it is not realistically possible to optimize for longer periods. However, it can be assumed that eventual errors in forecasting will require occasional recomputation of optimized schedule, so the long term solution may not be absolutely necessary.

Besides the optimization algorithm which works on precomputed data, a real-time approach will be also required for dealing with unexpected trends in withdrawals. This approach has to take into account the current amount of money in the ATM as well as expected withdrawals in the next few days. Depending on whether there is too much or too little money and e.g. a shortage can occur, the optimization may then be redone starting with the current day.

Since a real-time approach is secondary to the optimization the latter is more important for the solution. The optimization algorithm requires quantified costs associated with the filling-in schedule:

- cost of insurance - dependent on the amount of money filled-in at the given ATM,
- cost of cash freezing - dependent on the amount of money at any given time in the ATM,
- cost of filling-in - any filling-in has a fixed cost associated with the distance to the closest branch from which the cash for the ATM can be supplied.

The result based on those would include the amounts of money as well as days at which filling-in should take place. To take into account a margin of money that should be at all times in the ATM an additional set of assumptions can be included if desired to fine tune the results. These assumptions can include:

- setting a minimum and maximum limit of the number of days which can occur between sequential filling-ins,
- extra percentage of money that an ATM should be filled with (e.g. 20% so that assuming the forecast hold 20% of the filled-in money will be left when the next fill occurs).

An example algorithm works by first assuming that a filling-in occurs every day and then checking other possibilities. If no other than basic assumptions are given then Tables 1-3 show a sample of three steps conducted by the algorithm when optimizing total costs of filling-in schedule.

Date	Forecast	Fill-in (morning)
2013-01-01	6000	6000
2013-01-02	9000	9000
2013-01-03	3000	3000
2013-01-04	1000	1000
2013-01-05	10000	10000

Table 1: Step one of the algorithm

Date	Forecast	Fill-in (morning)
2013-01-01	6000	6000
2013-01-02	9000	9000
2013-01-03	3000	3000
2013-01-04	1000	11000
2013-01-05	10000	

Table 2: Step two of the algorithm

Date	Forecast	Fill-in (morning)
2013-01-01	6000	6000
2013-01-02	9000	9000
2013-01-03	3000	4000
2013-01-04	1000	
2013-01-05	10000	10000

Table 3: Step three of the algorithm

The steps of the algorithm are as follows:

1. assume that the filling-ins occur in a given order;

2. calculate the amount of money to be filled in as a sum of forecasted withdrawals in the consecutive days between every filling-in and next filling-in (here for example a 20% extra amount can be added in);

3. calculate the total cost for assumed schedule;

4. if total cost is less than previous minimal cost then set the discovered schedule as optimal;

5. modify schedule;

6. repeat from step 1.

The final results will describe at what days should the filling-in occur and what amount of money should be filled in. Example showed in Table 4. gives three such days with accordingly 15000, 4000 and 10000 money units to be filled in at those days. Since the algorithm checks every solution the result is the global minimum cost for the schedule optimized for these days.

Date	Forecast	Fill-in (morning)
2013-01-01	6000	15000
2013-01-02	9000	
2013-01-03	3000	4000
2013-01-04	1000	
2013-01-05	10000	10000

Table 4: Example final results of the algorithm

A real-time approach will than check the amount of money left at the ATM at the end of the day (e.g. 7000 on first day in Table 4) and should a shortage be expected (in Table 4 because of withdrawals of 9000 on the next day) the optimization algorithm can be ran again.

4 Decision Making

Given precise predictions on cash demands, in order to determine optimal filling of ATMs with cash, it suffices to use the described optimization procedure. However, predicting cash demand is a rather tough task, especially several days in advance. Namely, one of the statistics used — average of cash demand in last several days is not available for the future since demand history is not available for the given or future days. For future days, that statistic can only be approximated by the average corresponding to the day on which the prediction is made. Therefore, for better decision making, it would be better to perform optimization each day as soon as new demand data becomes available, and that way correct potentially bad decisions caused by bad prediction. For example, if one predicts on day one that an ATM will need refilling in 5 days, surprisingly high cash demand on that same day might change the prediction on the second day so that the refilling period is shortened.

5 Conslusion

Aim of the study group team was to minimize total costs of the Credit Agricole Bank in Serbia when supplying ATMs with cash. Main task was to find an optimal balance between two opposite approaches such that ATMs are supplied with sufficient amount of money for satisfying customers' needs. Philosophy of the first one is that ATMs should be rarely refilled with large amounts of cash which brings low transport cost, high freezing and high insurance costs. On the contrary, other approach is based on frequently supplying ATMs with small quantities of cash which reduces freezing and insurance cost and increases transport cost. Numerical experiment was focused only on transactions at external ATMs. Time series analysis and linear models were employed for obtaining predictions of daily cash demand at each ATM. Obtained results were later incorporated in optimization model of filling-in schedule. Final outcome of optimization algorithm gives information on days when the filling-in should occur and about amount of money should be filled in. The algorithm checks every solution and it results with the global minimum cost for the schedule optimized for these days.

References

[1] R. Hyndman, Y. Khandakar, *Automatic Time Series Forecasting: The forecast Package for R*, Journal of Statistical Software, 2008.

[2] J. Garca Cabello, *An efficient Liquidity Management For ATMs*, Aestimatio, the IEB international journal of finance, 2013.

[3] V. Kamini,V. Ravi, A. Prinzie, D. Van den Poel, *Cash Demand Forecasting in ATMs by Clustering and Neural Networks*, Working paper-University of Ghent, 2013.

[4] P. Soares Esteves, P. M. M. Rodrigues *Calendar Effects in daily ATM withdrawals*, Working paper - Banco de Portugal, 2010.