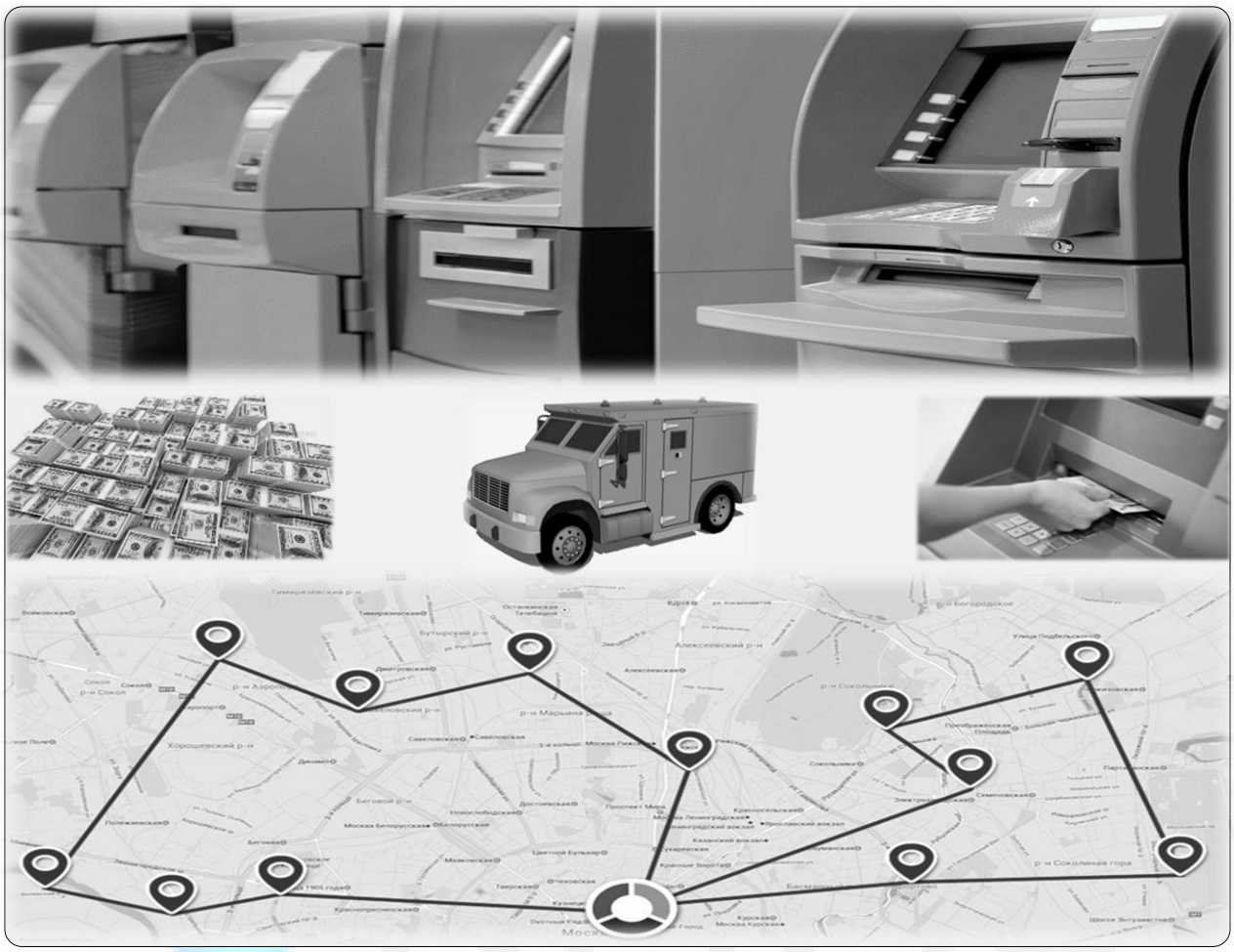


ATM REPLENISHMENT

AN OPTIMAL ATM CASH REPLENISHMENT STRATEGY USING DATA ANALYTICS & OPERATIONS RESEARCH

BY:
Dr Amit Garg, Chief Scientist, ORMAE
Priyadharshini S, BAI* – Indian Institute of Management, Bangalore
Dhananjaya BL, BAI* – Indian Institute of Management, Bangalore

* Business Analytics & Intelligence

Business Outline

There are more than 3 million Automated Teller Machines (ATMs) around the globe, which is only about 40 ATMs for 100,000 adults. There are 133 ATMs for 100,000 adults in US. Use of ATMs has grown exponentially over the past few years largely because of increased the functionality, proximity and customer adaption to technology. ATMs have become the main means of cash supply in UK – more than 70% of all cash withdrawals are now made by ATMs.

To overcome these challenges, the banking industry across the globe are embracing trends of self-service, 24 x 7 ATMs along with reduced operating cost and improved customer convenience. However, there is a trade-off: an ATM that's out of cash can negatively impact the reputation of the bank, on the other hand an ATM that is overfilled costs more and

Cost of ATM cash replenishment operation can be viewed in three parts. One: the cost of money in it; two: the logistic cost which accounts for transportation cost, handling cost, cash in transit cost, security cost; three: the fixed cost which includes vehicle depreciation cost, driver wages, order cost and insurance cost.

- **80% of ATMs & branches hold more cash than needed. - NCR Cash management consultancy**
- **According to ATM industry estimates, cash-related costs can account for around 30% of the cost of running an ATM network.**
- **North American banks often maintain as much as 40% more cash than necessary at their ATMs due to ineffective cash forecasting.**
- **Cash handling can represent 5–10% of a bank's operating costs and is the equivalent of 1–2% of total retail sales.**

Scope of the Study

This study focuses on establishing a conceptual frame work for **optimizing the cost of cash replenishment for networks of ATMs** and **minimizing the cash-dry instances at individual ATM**.

The objective is to **minimize the total cost of logistics, inventory, and improving customer service satisfaction**. This approach considers the combination of following techniques:

- **Time series forecasting** accounting for seasonality, trend and local calendar events
- **Inventory theory** – determining frequency and quantity of cash to be replenished for each ATM
- **Transportation problem** – clustering & single period vehicle routing problem to determine optimal route and number of vehicles needed.

Business Operations

Financial Institutions (FIs) and the Independent ATM Deployers (IADs) outsource their ATM operations to ATM Managed Services (ATMMS) enabling the FIs to focus on the core business development that increases their revenue.

ATMs is a turnkey service that includes managing vault cash, optimizing load amounts and schedules, managing armoured carriers, replenishing ATMs, providing first-line maintenance as well as insurance coverage.

Typically, the ATMMS collects cash from the central vault location and delivers it to various bank branches and ATMs (both onsite and offsite locations) for replenishment. The routing of the vehicle fleet happens from a single hub to the various locations.

Monitoring

ATMMS monitor their ATM cash levels to determine whether the machines need to be restocked. If yes, IADs decide how much cash should be restocked and in what denomination(s). They may choose to replenish ATM cash on a fixed schedule or as needed, based on the policy set. For instance, every Monday and Friday morning, or when there is low cash inventory left in the ATM.

Sources of cash

The sources of cash for IADs are typically their own working capital or from vault cash suppliers, which are usually owned by banks. Banks obtain their cash either from their country's central bank, from other banks, or from business customers' deposits. Once the cash arrives, banks store it in their branches and vaults.

Ordering, Sorting & Transportation

The cash is either ordered from the funding bank or third-party cash supplier like Cash Connect or Elan Financial Services (in US). From the bank the cash is sent to a money centre near the ATM. The cash provider coordinates with the armoured carrier to pick up the physical currency at the money centre. The third-party cash provider recycles cash in armoured carrier inventories, from which the IAD uses the existing cash.

Banknotes are counted and sorted at a cash processing centre or bank branch before being packaged to load into the ATMs. Damaged or soiled notes are taken out and are sold to the central bank at face value. Cash centres use bulk cash acceptors to process incoming notes. These machines simultaneously count and authenticate notes, separating them by denomination into individual containers.

FIs and IADs have two cash-packaging options. They can place the notes either in a tamper-proof ATM cash cassette or in a sealed bag that is then taken to an ATM. The advantage of sealed bags is that they are more space-efficient, easy to handle, and require less time to complete the replenishment process when compared to cassettes. The transportation of the money happens in an armoured carriers which typically have a driver and a security guard.

Methodology

Data science and Optimization techniques can add value to the services benefiting both the ATMMS and FIs. Given the past data, we obtain useful insights that can be turned into actionable outputs leading to reduction in the cost of the replenishment operation at the same time improving customer convenience.

Cash replenishment planning is a formulation of an optimized plan to maintain adequate amount of cash in ATMs by monitoring the amount of remaining cash, forecasting future demand, and appropriately timing the replenishment.

The solution thus has two parts:

1. Data analysis: Forecast demand based on the past data and plan inventory.
2. Optimization: To create optimized distribution plans from vaults to ATMs.

Typically, an ATM machine has a capacity of $200,000 (4 cassettes each with 1000 bill capacity), each cassette with varying denominations.

Some of the offsite location ATMs may have a smaller capacity of $100,000 with only 2 cassettes.

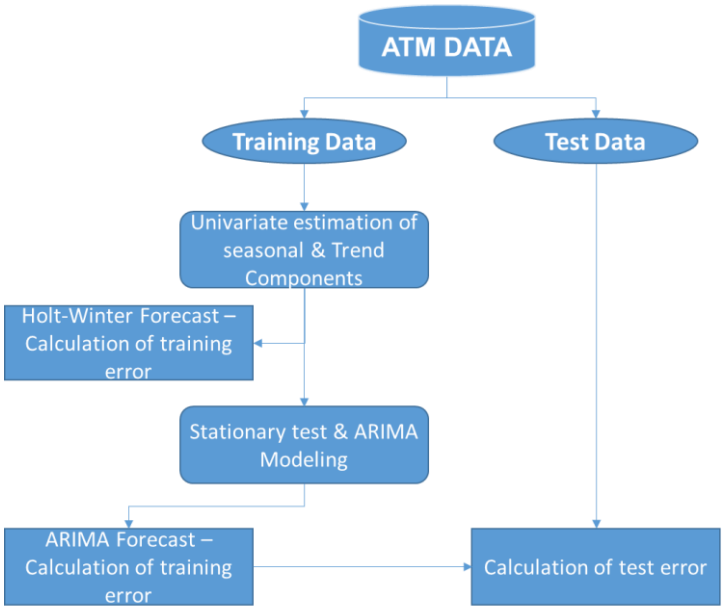
$10,000 is the average amount that would be held in ATMs that are remotely located and sometimes even less during non-peak hours.

The withdrawal limits vary for a customer and a non-customer. A customer of the bank will have $500 as the limit at a time and $3000 as the daily limit while a non-customer will have $1000 as the maximum daily limit.

The maximum deposits of cash at a time is 50 bills and the daily maximum deposit limit is $5000.

Data Analysis – Demand Forecasting

- Demand for the ATM cash reserve is forecasted on a daily/weekly/monthly basis. The outflow amount is predicted for each day of the following month
- Demand on ATMs is highly correlated to locale, localized economics and local events
- Cash demand has seasonality on weekly and monthly basis, weekend demands are 1.5 to 2 times higher than weekday demand. Demands are higher during first week of the month, attributed to salary deposits.
- Cash Demand peaks during holidays and festive seasons like Christmas, and can be 40 percent higher than at normal operating times
- Major factors affecting the demand:
  - ATM location
  - Festive / Holiday seasons
  - Special events like any local events - sports, bonus payouts, beginning of the month, end of the year
  - Characteristics of the financial institution
  - Required periodicity of the replenishment service at the ATM and the time of service
  - Extreme weather conditions like hurricanes, blizzards
- Time series methods like simple moving average, exponential smoothing and univariate ARIMA models can be employed after establishing stationarity assumption.
- The model with the least Mean Absolute Percentage Error can be used to make most accurate forecasting
- This forecasted time series data is later adjusted for special local events at the ATM locations
- Using the outflow predictions changes in the remaining ATM cash are calculated



- Finalize replenishment plan - the replenishment frequency and amount is finalized based on :
  - ATM's remaining cash data
  - Forecasted demand
  - Maximum capacity of the ATM machine
  - Daily maximum withdrawal limit
  - Daily maximum cash limit that can be deposited in the ATM
  - Maintenance shutdown, if any
  - Peak time
  - Non-peak time
  - Location of the ATM
- The actual record is compared with the plan on a regular basis (monthly) to check if the prediction is accurate.
- Varying cash denominations is also predicted for the ATM to balance the cash denominations based on the past data and how much of each of the currency notes can be stored in the ATM machines.

Routing Optimization

The optimal strategy is to focus on the reduction of cash-related expenses while safeguarding that ATMs do not run out of cash.

The proposed Single Periodic Vehicle Routing Problem (VRP) model will be able to solve single replenishments with simultaneous requirements of many ATMs daily in which a set of routes is planned starting from a single hub.

It's a 2-Phase algorithm: The problem is decomposed into Clustering of ATM locations into feasible routes and the Actual route construction on a daily basis.

Objective function:

To minimize the total travel cost and the inventory-related cost.

Decision Variables:

Service days for each ATM and the vehicle routes for a service provider on each day.

Constraints:

- Each vehicle makes only one route per day.
- All the routes start in the same depot/hub.
- To keep the fleet of vehicles as low as possible.
- SLAs with the armoured carriers /ATM managed services.
- Typically, the armoured carrier service time is 9 hours a day.
- Typical Fixed costs involve:
  - Leasing cost of the ATM machine which typically ranges from $500-$2000 annually. This depends on ability of ATM machine to discharge multiple currencies vs single currency bill. It also depends on the technology of the ATM.
  - Cost of cash handling on an average is $100/ATM/month.
  - Armoured vehicle leasing cost ~ $4000-10000 per year .On an average,$7000 per year or $20 per day taken.
  - Armoured vehicle per mile cost ~ $0.40.
  - Driver or Security person or vault person salary per hour: $15
  - Leasing cost of vault location ranges from $1000-$3000 /month, on an average of $2000/month. An employee at a vault earns $2500 per month.
    Hence the total monthly cost per vault = $4500 per month.
- Capacity of the vehicle – Cash amounting to millions of dollars
- Each ATM is visited in the cluster on the same day
- Average Replenishment time = 20 minutes =1/3 rd an hour
- During the T-days period each ATM must be visited at least once.
- Generally the replenishment is done weekly once for an ATM and hence T=7 days.

Algorithm flow:

- Planning is done for T-day horizon.
- Each ATM is assigned a day over the T-day period.
- Once a decision is taken as to which ATMs are to be replenished for the day based on the forecast, limited set of clusters are formed based on distance between the ATMs.
- Clustering is the task of grouping a set of objects in such a way that objects in the same group are similar to each other than those in other clusters.
- K-means clustering is a method that aims to partition n observations into k clusters (k<=n) in which each observation belongs to the cluster with the nearest mean.
- Given a set of ATMs (x1,x2,…….xn) where each ATM is a d-dimensional real vector, k-means aims to partition the n ATMs into k(<=n) sets S = {S1,S2, …….Sk) so as to minimize the within-cluster sum of squares which is equivalent to assigning by the smallest Euclidean distance.
- This algorithm is an iterative refinement technique.
- It starts with an initial set of k means and an assignment step where each ATM is assigned to a cluster whose mean yields the least squared Euclidean distance. Then the new means to the centroids of the ATMs in the new clusters are calculated.
- The algorithm has converged when the assignments no longer change. Since both the steps optimize the within-cluster sum of squares objective, there exists only a finite number of such partitions, the algorithm converges to a local optimum.
- Once the clusters are formed, route optimization is done for each of the clusters using the travelling salesman problem (TSP). There are many advanced implementations of TSP available.
- One of the approach for TSP problem is, the Nearest Neighbour (NN) algorithm, which can be applied to get initial solution of TSP. Vehicle starts at the hub and repeatedly visits the nearest ATM until all of them have been visited in the cluster.
- After initial TSP solution via NN, interchange heuristics can be applied to improve it. It reconnects nodes in such a way that it disconnects a node a* from a1-a*-a2 by making connection direct a1-a2 & it inserts this node a* in another arc b1 – b2 by making it b1-a*-b2 where a1,a2,a*,b1& b2 are nodes. It selects

disconnection node & reconnection arc in a way to maximize savings. It keeps repeating this process till no additional savings can be achieved.

Scenarios:

Single Vault Scenario:

Considering approximately 1000 ATMs in a downtown location of a major city, most of the ATMs are located within a mile of distance.

Service time of the carrier = 9 hours /day

With single cluster, if the round trip time to and from a cluster to vault and time taken within the cluster is around 3 hours, then remaining 6 hours are for replenishment.

Replenishment time = 1/3rd an hour and hence within 6 hours 18 ATMs can be replenished.

Annual cost of 1 vehicle running once a week = $320*52 = $16,640 based on the following cost factors:

1. The driver and the security wages & vehicle leasing cost to average to $15*18 + $20 = $290 per day.
2. Mileage cost = $30 depending on the total distance.

Thus, with n number of ATMS that have to be replenished for the day, k clusters are formed (k<=n) based on the Euclidean distance between the ATMs with each cluster consisting around 18 ATMs. Therefore, about 1000/18 ~ 55.6 clusters are formed. Fractional cluster means a cluster with smaller number of ATMs in it than needed for whole day route.

Weekly routing cost = 55.6*16640 = **$925,184**

Annual vault rental cost = **$54,000**

If single currency ATM leasing cost average to $800 per annum, then ATM dispenser leasing cost for all ATMs = **$800,000**

Total cost = 925,184+54,000+800,000 = **$1,779,184**

Multiple Vault Scenario:

Considering a scenario of having more than one hub, i.e., more than one cash vault location from where vehicles load the cash enrouted to the ATMs. Assuming that a vehicle that starts at one vault returns to the same vault.

To start with, the vault location (starting hub) is decided for each of the ATMs .Then clusters for each of the hubs are formed.

The cost of extra vault location= $54,000 per year.

It will be profitable if the number of clusters decreases by 54,000/16,640 ~3.2 on adding an extra vault.

To do that, each cluster should be able to handle 1000/(55.6-3.2) = 19.1 ATMs.

It will be beneficial if adding extra vault **decreases average time** from vault to cluster & back by more than 20*1.1 or **22 minutes** as time within the cluster may not be significantly affected.

Multiple Bill Currency Dispenser:

Considering ATMs with multiple bill currency dispenser. Generally, ATMs dispense $20 and $10 bills. The ATM machine leasing cost may vary depending on this and the technology involved, which would range between $500-2000.

Assuming multiple bill dispensing ability might cost an ATM around $1000 per year.

Considering the teller time taken by a customer at a branch because of a requirement of currency of other type = 5 minutes.

Customer inconvenience cost ~ $5 per occurrence or more if time spent by customer = 20 minutes or so. Another inconvenience factor is customer will have to wait for next business day for getting multiple currency bill from branch if the time has crossed the regular hours of the branch.

Teller's time cost per customer ~ 1/12 *15 = $1.25

Number of customer trips for non ATM dispensed currency required to justify multiple currency ATM:

1. 800, if teller's cost alone is taken into account.
2. 160, if both teller's cost and customer inconvenience cost are taken into account.

But increasing customer's cost leads to customer churn.

Single period vs Multi period vehicle routing problem:

Considering an ATM whose cash is consumed at a rate that needs replenishment more than once a week, i.e., the total withdrawal per week exceeds $200k, which means more than 400 withdrawals happen per week assuming the average withdrawal amount is $500 per transaction. If this happens throughout the year adding another ATM is the solution for which the cost is :

Annual total cost of adding another ATM = Annual leasing cost of one ATM + Annual routing cost per ATM

Total annual cost = 800 + 925 = **$1725**

Conclusion

Multiple Vault vs Single Vault: Having more than one vault decreases the time taken to finish the replenishment. Reduction in time savings can be observed in progression with additional vaults. By balancing replenishment time cost i.e. more ATMs per cluster with cost of adding additional vaults, optimal number of vaults can be estimated.

Single period vs Multi period vehicle routing problem: Adding another ATM for the scenario addresses the problems of ATM running out of cash and long queues outside the ATM. The same may not be justified when the ATM withdrawal peaks seasonally or only during few times of the year. In that case, a visit to the ATM more than once a week will help.

Multiple Bill Currency Dispenser vs Single bill currency dispenser: The teller time and cost is saved, the customer inconvenience can be avoided .The customer convenience is increased as the ATM is available 24/7. Therefore, Multiple Bill currency dispenser is strongly encouraged for busy ATMs.

Benefits of this Project

- Through cash replenishment planning service, ATM uptime ratio can be improved while reducing the ATM operational cost
- Reduced logistics and inventory management costs by 2-5%
- Optimizing ATM cash management frees idle cash for more profitable revenue-generating opportunities
- Customer convenience is improved by eliminating ATM downtime due to lack of cash and by availability of appropriate denominations in the ATMs
- Minimizes emergency cash requests
- Improves operational efficiency
- Forecasting helps ATM managed service issue pre-emptive cash outage alerts

References:

1. **ATM Cash Management 101**
   https://www.atmmarketplace.com/whitepapers/atm-cash-management-101/

2. **Transforming Cash Supply Chain Management**
   https://www.atmmarketplace.com/whitepapers/transforming-cash-supply-chain-management/

3. **Using Vault Cash service Helped IAD More than Double ATM Business**
   https://www.atmmarketplace.com/whitepapers/using-vault-cash-service-helped-iad-more-than-double-atm-business/

Disclaimer:

Facts and figures quoted in this document are from trusted and established sources.

For calculation purpose, an average from the range of figures is taken. The result may vary depending on the actual figures.