Building on belief

Replenishment of retail stores

TCS Quantum Computing Challenge
Challenge Sponsor: *M&S*

September 7, 2023

Problem Statement Overview

For any retailer with stores, it is not easy to service their stores with the right products at the right time.

Store replenishments are currently based on calculations made in supply chain software solutions. These calculations depend on several parameters and data sets such as:

- Store stock
- Store predicted sales (and returns, if applicable)
- In-transit stock to store
- Store stock targets (planograms)
- Distribution Centre (DC) available stock (for example, a ‘fair share’ approach may need to be applied if DC stock is low)

Store replenishments also depend on the upstream supply chain. The DCs that service the stores also have parameters and data sets that determine the level of stock they hold and when they are replenished:

- Safety stock, based on demand and demand variability, lead time and lead time variability, and desired service levels (for example, twice as much stock is required to support a 98% service level vs. an 84% service level)
- Cycle stock, based on supplier order quantities and frequencies (Efficient Order Quantities (EOQ) balance cost to purchase and carrying costs)
- Actual stock and in-transit stock
- DC capacity

However, the current supply chain systems do not take costs into account. For example,

- Inventory carrying costs for which there is a ‘fixed’ element (such as the allocation of administration costs) and an element that is 'variable’ by product (such as product value, handling costs, space requirements)
- Cost and service impact of rebalancing stock between warehouses vs. placing supplier orders
- Determining if the defined fixed store delivery frequencies are optimising Cost to Serve

It is hard, therefore, to optimise the balance between customer availability, working capital/inventory, and operating costs. This is challenging because:

a) The key metrics can be in opposition to one another. High customer availability, for example, can be helped by frequent replenishment (higher operating costs) and higher safety stock levels (higher working capital). Often, responsibility for these metrics sits within different teams, creating the potential for conflict of interests and silo mentalities. For example, different teams may manage the minimum quantities ordered for a product from a supplier, the lead times/delivery frequencies between DCs and stores, and the target stock level to be held for each product in each store.

b) There can be complexities within the metrics and the operating model. For example, variations between stores and products (rate of sale, range, delivery frequency), and variations of pack sizes. Often, this complexity is not fully understood or is only understood within individual areas without being pulled together to impartially assess trade-offs. The number of SKUs, and stores, and high SKU turnover (driven by fashion and seasonality) mean that manually, yet accurately adjusting the parameters based on real sales and lead times becomes difficult.

c) The parameters, their relationship, and their impact on the supply chain are not always frequently reviewed and regularly updated. This results in the system making decisions that are not aligned with the physical supply chain.

It is important, therefore, to clearly prioritise the metrics within a business (that is, to set a clear target), and then establish a way to optimise solutions that take the breadth and depth of the complexity into account.

Traditional analytics tools or techniques tend to focus on one-off solutions within individual metrics (that is, lowest operating cost) or work at an aggregate level (groups of products, rather than individual products). An ideal solution would use advanced analytics to take all factors into account and repeats the analysis in line with the business needs as well as determines alternative scenarios. There will always be one outcome that is theoretically the most profitable, but we often make decisions that are unprofitable for the sake of the brand (for example, selling beauty boxes at a loss over Christmas). Optimising for cost, availability, service, or capacity differently will result in different answers.

In summary,

1. The system parameters are extremely high-level and are not automatically optimised or frequently reviewed based on the actual supply chain performance. For example, delivery frequency, lead times and variability, pack sizes, and stock targets at DC and store based on product lifecycle.

2. Recommendations for store replenishment are based on parameters that do not consider true Cost to Serve, which tends to be used more for strategic rather than transactional decisions. In addition, such solutions cannot define multiple scenarios based on the same input parameters, considering the best cost, best service, and best capacity/stock, and simultaneously presenting these options with associated risks for assessment and implementation.

Business Value & Motivation

Why is this problem important for the industry?

This problem influences customer availability, which is a key business metric that affects customer experience and sales.

This problem also affects the Cost to Serve and working capital, which are key Supply Chain and Logistics metrics.

Critically, these metrics can trade off against one another. The industry does not yet have a standardised methodology for providing solutions to these trade-offs that encompasses all critical elements.

Why is this a good problem for exploring a Quantum solution? What metrics are sought to be improved (speed, accuracy, training time, etc.)?

Existing solutions do not fulfil the requirement to view multiple scenarios from the same input data, and target different solution parameters and trade-off network options as part of the business decision-making process.

In real-world applications, the problem becomes considerably large when dealing with many SKUs, stores, variability in demand, and other constraints, thus increasing the time complexity and difficulty to solve using current practices. In short, the problem becomes intractable in the classical world.

So, this problem is a good candidate for exploration using Quantum computing as quantum algorithms can solve this problem faster and more efficiently than classical algorithms.

Quantum solution shall target improvement in metrics such as reduction in inventory and increase in sales.

Which Business KPIs/Measures could possibly be impacted as a result of a Quantum solution?

Business KPIs/Measures that require improvement are listed as follows:

- Logistics Cost to Serve including transport and warehousing costs
- Working capital/Inventory
- Warehouse Capacity utilisation
- Inventory turns

To validate the performance of the solutions, the following metrics shall be considered such as average inventory coverage which measures how long the end of the day's closing stock can fulfil the daily demand into the future, or other relevant metrics in practice today.

Current Solution

How is the problem currently being addressed?

The current solution uses a rule-based approach instead of a Machine-Learning based approach. Currently, items are grouped based on predicted sales rates, however, the groups can contain thousands of items. A generic day of cover of inventory is calculated for the group. However, within a group, there is a level of variability that should be used to fine-tune the number of days cover at an item level.

The classical approaches that the Retail Industry uses to solve the problem are detailed as follows. Please use the 'References' section of this document for the research papers cited.

- The traditional approach to inventory classification is ABC analysis (Syntetos et al. 2009). This involves classifying SKUs; the selection of classes is based on experience not accounting for variable demands and value, which is a major limitation. This led to the evolution of the Multiple Criteria Inventory Classification (MCIC). This method uses multiple criteria such as demand variability, lead times, and carrying cost to classify SKUs into different categories (Ramanathan 2006, Hadi-Vencheh 2010). Some challenges with MCIC methods are the selection of appropriate criteria and the determination of weights assigned to each criterion; these choices can be subjective and dependent on the judgment of the decision-maker.
- Most methods focus on the development of pure SKU ranking methods, as opposed to optimizing inventory performance. The integrated inventory management approach considers inventory classification and control decisions in an integrated manner. (Teunter et al. 2010) (Yang et al. 2017)
- Another approach is to maximize profits while classifying SKUs into respective cycle service levels. The results show superior inventory performance compared to the traditional methods. Further work for inventory management using deep reinforcement learning (Meisheri et. al 2021) (Ganesan et al., 2020) is done and being explored.
- Optimal replenishment decisions and inventory levels for efficient supply chain operations in a network-level model are yet to be explored, which will provide us the opportunity to explore more models and methods in this domain.

What are the results of the current solution?

The generic approach (in terms of grouping thousands of items based on predicted sales rate and inventory calculation at the group level) results in excess stock for some items and risk of too little stock for other items. This can lead to a missed sale and disappointed customer. A more granular approach should be possible with Quantum, delivering improved sales at lower cost through inventory classification and replenishment at SKU levels at each store.

Are the current results satisfactory? What are the limitations of the current solution?

The current results are not satisfactory for the grouping of thousands of items with generic days of cover of inventory calculated at the group level. Hence, this limitation must be addressed through a level of variability that should be used to fine-tune the number of days of cover at an item level.

Availability measures are good, however following KPIs need improvement:

- High delivery frequency and cost to serve
- High store inventory/working capital
- Singles per order line drives underutilised warehouse capacity and higher cost

Problem Definition for the Quantum Challenge

Redefined scope of the problem for the purpose of the Challenge

Objective

Optimize the inventory replenishment for retail stores from the Distribution Centre (DC) to:

- maximise profitability
- reduce the working capital
- increase the service level fulfilments or commitments

Considerations

- Typical retail chains sell about 200K to 300K SKUs in larger outlets.
- Retail stores are defined into multiple types (even under one ownership/name across its retail chain of stores) based on store positioning, branding, and products sold.
- Each store may have multiple departments, 100 to 150 approximately.
- Each department may have multiple categories of products with many brands.
- Replenishment Problem
    - Stores need to replenish SKUs from Distribution Centres using the following information:
        - Sales forecasting considering seasonal effects with given levels of forecast accuracy, exceptions, and outlier definitions.
        - Replenishment options considering Store types, Product Categories, SKU service levels, safety stocks, and supply lead times with variability/delivery schedule options.
- Formulate and model the problem as an optimisation problem either only at the store level (or DC-store cluster level) to effectively manage inventory level at stores with the following scope:
    - Considering SKU demand volumes, minimal/maximal order quantities, replenishment frequencies, and service levels.
    - Forecasting accuracy, promotional information/external factors, and stock availability at DCs.
    - Objectives of minimizing working capital invested in safety stocks and maximizing profits/margins across product SKUs in a planning horizon.

The redefined scope of the problem is divided into two parts as follows:

1. Inventory Replenishment at Store Level ***(Level 1)***
   Maximize the profits/margins and minimize the working capital with respect to store-level constraints
2. Inventory Replenishment at DC + Store Level ***(Level 2)***
   Maximize the profits/margins and minimize the working capital with respect to store and DC-level constraints

Inventory Replenishment at Store Level (Level 1)

The problem is to replenish SKUs in a store using the following information:

I. Sales forecasting
II. Replenishment based on the following:
    a. Store types
    b. Product Categories
    c. SKU service levels
    d. Supply lead time with delivery schedule

Objective Function

Maximize profits/margins across products =

(Expected gross profit X expected sales volume) – (Inventory holding cost)

Input Parameters

- Forecasted demand of each SKU in every time period
- Standard deviation of the forecast demands each SKU in all time periods
- Unit gross profit of all SKUs
- Unit cost of SKUs
- Inventory holding cost (% of the cost of SKU) per time unit
- Standard normal value (z-value) associated with CSL
- Maximum available inventory capital
- Derived parameters based on replenishment lead time and review interval

Decision Variables

- Classify every SKU for a particular time period in a CSL class
- Capture on-hand inventory of all SKUs at the end of every time period
- Capture the demand being satisfied for all SKUs for all time periods

Constraints

1. An SKU should be assigned exactly one inventory class in any given time period
2. Expected demand to be satisfied in the time period, t <= forecasted demand
3. On hand inventory of preceding time period <= On hand inventory after current time period + expected demand to be satisfied, provided inventory replenishment order is expected arrived or not
4. On hand inventory + expected demand to be satisfied on time period, t >= forecasted demand + safety stock
5. The total invested inventory value in any given time period <= Total available inventory capital

Datasets

1. Store demand by product, by time period
2. Store stock files (historical sales, future forecast, SKU segments, previous orders)
3. Product lead time information
4. Unit cost price of SKU

5. Unit gross profit for SKU
6. Holding cost for each SKU (% of CP of SKU)
7. Total budget amount (Available inventory capital)
8. Product categories: (Combination of below categories)
    i. on the basis of margin (high, low, moderate)
    ii. on the basis of sales (fast moving, slow moving, moderate moving, non-moving)
    iii. seasonality

Inventory Replenishment at DC + Store Level (Level 2)

This optimisation problem deals with the most efficient way to distribute goods from DCs to store and optimizing inventory in all the stores with the following considerations:

- There would be multiple stores and multiple DCs
- Inventory classification should be handled at store level
- Safety stock should be handled at store level
- Flow balancing for both DC and store level

Objective Function

Maximize profits across products =

Expected gross profit * expected sales volume from stores
– the inventory holding cost at stores
– product cost
– penalty for backorders at stores
– penalty for backorders at DCs

Input Parameters

- Maximum available inventory capital for every store
- Maximum available inventory capital for every DC
- Forecasted demand of each SKU in every time period for every store
- Standard deviation of the forecast demands of each SKU in all time periods for all stores
- Unit gross profit of all SKUs
- Unit cost of SKUs
- Inventory holding cost (% of the cost of SKU) per time unit
- Standard normal value (z-value) associated with CSL for all stores

Decision Variables

- Classify every SKU for a particular time period in a CSL class for every store
- Capture on-hand inventory of all SKUs at the end of every time period for all stores
- Capture the demand being satisfied for all SKUs for all time periods for all stores
- Capture on-hand inventory of all SKUs at the end of every time period for all DCs
- Capture the demand being satisfied for an SKU for a time period from a DC to a store
- Capture the backorder quantity for all SKUs for all time periods for every store

- Capture the backorder quantity for all SKUs for all time periods for every DC
- Capture forecasted demand for DC for all time periods for every SKU

Constraints

1. An SKU should be assigned to exactly one inventory class in any given time period for a given store
2. Forecasted demand for a given store to be satisfied in the time period, t = backorder for that store + expected satisfied demand for the store
3. On hand inventory of the preceding time period of a given store + expected demand satisfied by all DCs for that store = Current on hand inventory of the store + expected demand to be satisfied by that store
4. Expected demand satisfied by one DC for all stores + Backorder of that DC = forecasted demand for that DC (variable)

Datasets

1. Store demand by product, by time period
2. Store stock files (historical sales, future forecast, SKU segments, previous orders)
3. Product lead time information
4. Unit cost price of SKU
5. Unit gross profit for SKU
6. Holding cost for each SKU (% of CP of SKU)
7. Max available inventory capital for each store and each DC
8. Warehouse stock files
9. SKU mapping from DCs to stores
10. Allocation priority order of stores from DCs (which store is to be served on the basis of priority from a DC) to fulfil the requirement
11. DCs/stores capacity and available inventory capital
12. Product categories: (Combination of below categories)
    i. on the basis of margin (high, low, moderate)
    ii. on the basis of sales (fast moving, slow moving, moderate moving, non-moving)
    iii. seasonality

Challenge Evaluation Criteria

The entries will be evaluated by a Jury panel consisting of experts from the Industry and Academia. A representative list of criteria that would be considered by the Jury while evaluating the Phase-I and Phase-II submissions is given below. Please note that there are no specific weightages assigned to any of the criteria, and this should not be interpreted as a comprehensive list. The criteria listed below are indicative, and the Jury panel would be free to use their expertise, experience, and judgement to evaluate the entries.

Phase I:

- **Background work done:** Quality of background work done to understand the use case and the state of the art.
- **Innovation Quotient:** The overall innovation quotient in terms of originality and novelty seen in the approach, concept, and the algorithm.
- **Comprehensiveness:** Level of detail, coverage of the tasks towards the targeted goal of the challenge.
- **Promise:** As reflected through the results from any early experimentation done. Promise of the planned approach for Phase II.
- **Technical soundness:** Ability of the team to defend their work during presentation/ interaction with the Jury. Capability of the team to carry forward the work.

Phase II:

- **Innovation Quotient:** The overall innovation quotient in terms of originality and novelty seen in the approach, concept, and the algorithm.
- **Comprehensiveness:** Level of detail, coverage of the tasks towards the targeted goal of the challenge.
- **Technical completeness:** Quality and completeness of the code/solution and its ability to execute and produce results as documented.
- **Comparison with internal benchmarks:** How well the solution compares with any benchmark results that may have been achieved by the organisers.
- **Impact:** Value impact of the solution on current quantum hardware (any benefits shown in terms of improved optimisation, speed, and accuracy vis-à-vis classical approaches).
- **Extensibility:** Ease of adapting/modifying the solution to address variants of the use case (especially additional complexities).
- **Resource requirements & Scalability:** Any resource estimation provided for the solution to indicate the potential future scaling of the solution.
- **Technical soundness:** Ability of the team to defend their work during presentation/ interaction with the Jury. Capability of the team to carry forward the work.
- **Pitch:** Clarity, demonstration, and structure of the overall pitch.

References

1. **Adaptive Inventory Replenishment for Dynamic Supply Chains with Uncertain Market Demand**
   Viswanath kumar Ganesan, David Sundararaj, Ananda Srinivas
   https://www.researchgate.net/publication/340088950_Adaptive_Inventory_Replenishment_for_Dynamic_Supply_Chains_with_Uncertain_Market_Demand

2. **Mitigating Uncertainty in Supply Chain**
   Viswanath kumar Ganesan, Ananda Srinivas
   https://www.researchgate.net/publication/370266892_Mitigating_Uncertainty_in_Supply_Chains

3. **A Learning Based Framework for Handling Uncertain Lead Times in Multi-Product Inventory Management**
   H Meisheri, S Nath, M Baranwal, H Khadilkar
   https://arxiv.org/pdf/2203.00885.pdf

4. **Scalable multi-product inventory control with lead time constraints using reinforcement learning**
   Hardik Meisheri, Nazneen N. Sultana, Mayank Baranwal, Vinita Baniwal, Somjit Nath, Satyam Verma, Balaraman Ravindran & Harshad Khadilkar
   https://link.springer.com/article/10.1007/s00521-021-06129-w

5. **Reinforcement Learning for Multi-Product Multi-Node Inventory Management in Supply Chains**
   Nazneen N Sultana, Hardik Meisheri, Vinita Baniwal, Somjit Nath, Balaraman Ravindran, Harshad Khadilkar
   https://arxiv.org/abs/2006.04037

6. **ABC Inventory Classification with multiple criteria using weighted linear optimization. Computers & Operations Research 33:695-700**
   Ramanathan, R. 2006.
   https://linkinghub.elsevier.com/retrieve/pii/S0305054804001790

7. **ABC Classification: Service Levels and Inventory Costs. PRODUCTION AND OPERATIONS MANAGEMENT 19:343-352**
   Teunter, R.H., Babai, M.Z., Syntetos, A.A. 2010
   https://onlinelibrary.wiley.com/doi/10.1111/j.1937-5956.2009.01098.x