Approximate dynamic programming for lateral transshipment problems in multi-location inventory systems

Joern Meissner, Olga Rusyaeva

Many companies allocate their inventories across multiple regions based on historical sales rates. However, the random fluctuation in customer needs, caused, for example, by weather conditions and other external factors, might cause significant deviations from the expected demand, leading to overstock in some regions and stock outs in others regions. To fix the mismatch between actual customer demand and the available stock in multiple locations in the absence of a replenishment opportunity from a central warehouse, companies often turn to lateral transshipments, e.g. the movement of stock between locations of the same echelon.

So far the use of lateral transshipments has been observed in the management of spare parts. As slow-moving expensive items, spare parts require a quick respond to an infrequent demand. Lateral transshipments allow an additional source of procurement from a close-by location once a stockout of spare parts is occurred, *reactive transshipments*. This leads to the cost reduction and service improvements compare to no-transshipment networks. Up until recently, the main limitation in applying this strategy for final products and consumable goods has been insufficient and unreliable information about current inventories within supply chains. Thereby upcoming developments in computer technologies and electronics give the opportunity to exchange real time data about inventories. This information might be used by lateral transshipments not only to satisfy the demand of customers that are willing to wait, but, furthermore, to predict stockouts, *proactive transshipments*.

In our paper we focus on the items with a short selling season and a highly uncertain demand. As a case of the use of lateral transshipments, we refer to the following problem.

Consider a multi-location distribution network with one central warehouse in Bavaria, Germany, and 15 retail outlets in different states of Germany. Far in advance, the central warehouse makes an order of a product to a supplier incorporating a long lead time of the production and the delivery. Once the order is arrived to the central warehouse, he makes the decision about the amount of goods, that should be delivered to retailers. Assume that a replenishment decision for each retailer is implemented at the beginning of each month during the selling season. Each day customers are coming to retailers in a hope to find full shelves of products that satisfy their expectations. That runs down stocks of retailers, and in some states it happens faster than in others depending on customer's needs of the product. If a customer can not find the desired product because of the stockout, we assume that the customer is lost for us together with the profit that we could gain. In the beginning of each day of the month the retailers have enough information about their current inventories. If the amount of some goods is not enough to meet the demand of customers in upcoming days, these products can be transferred from other outlets of the distribution network. Depending on the length of the selling season, there might be other replenishments from the central warehouse as well as a global procurement from the supplier.

In our work we develop a model for this particular situation, that can be classified as a multi-location multi-period transshipment problem with lost sales in case of stockouts.

We present a dynamic programming approach to find an optimal transshipment policy. Each period of time the policy decides if the transshipment should be performed, and how many units and between which locations should be shipped to maximize the revenue of the whole network over all time periods. As the size of the state space and decision space is large for real-size problems and, therefore, causes the problem of the *curse of dimensionality* for the dynamic programming, we derive a near-optimal transshipment policy by applying a forward approximate dynamic approach.

We approximate the value function by a sum of separable, piecewise-linear, concave functions. Chosen approximation allows us to reformulate the maximization problem as a minimum-cost flow problem, which significantly simplifies further calculations of solutions. We apply the Concave Adaptive Value Estimation (CAVE) algorithm to learn the value function. The algorithm updates the concave, separable, piecewise linear function using the stochastic gradient estimates of this function at different states of the state space in a way that preserves the concavity of the approximation.

Finally, we compare the performance of our transshipment policy to other proactive policies. Numerical experiments show that our proposed algorithm (ADP) achieves a competitive performance against the state-of-the-art methods in the literature. For a small size instances, where the optimal solution can be found, there is no significant difference between an optimal solution and the solution found by ADP in 99 % of scenarios. For a real-sized problems, ADP gives high-quality solutions, that achieve 92%-99% of the upper bound, obtained through the deterministic demand problem. The observed performance of the policy for a multi-location problems with different network configurations demonstrates the capabilities of ADP and suggests the use of it for practical problems.