

О ДВУХЭШЕЛОННЫХ СИСТЕМАХ УПРАВЛЕНИЯ ЗАПАСАМИ ПРИ ПУАССОНОВСКОМ СПРОСЕ И УПУЩЕННЫХ ПРОДАЖАХ

Си Ту Тант Син

Национальный исследовательский университет «МИЭТ», Зеленоград, Москва

Ключевые слова: управление запасами, цепи поставок, ресурсное обеспечение, математическое моделирование, имитационное моделирование, многоэшелонные системы, пуассоновский спрос, адаптивное управление.

Аннотация. В современной логистике и распределенном промышленном производстве мы рассматриваем двухэшелонную систему непрерывного пополнения запасов при пуассоновском спросе и политике пополнения «один к одному». Спрос не удовлетворяется, если на местном складе, центральном складе или на промежуточных этапах (в пути) нет нужных товаров. В данном исследовании мы предлагаем простой, быстрый и точный подход к оценке уровня обслуживания в этой сложной системе. В отличие от других существующих методов, нам не нужна сложная схема итеративного анализа. Наш метод отлично работает в широком спектре случаев, при этом отклонения от результатов точного моделирования составляют в среднем менее 0,1%, а в 95% всех тестовых случаев – менее 0,36%.

ABOUT TWO-ECHELON INVENTORY MANAGEMENT SYSTEMS WITH POISSON DEMAND AND LOST SALES

Si Thu Thant Sin

National Research University of Electronic Technology «MIET», Zelenograd, Moscow

Keywords: inventory management, supply chains, resource provisioning, mathematical modeling, simulation modeling, multi-echelon systems, poisson demand, adaptive management.

Abstract. In modern logistics and distributed industrial production, we consider a two-echelon continuous inventory replenishment system with Poisson demand and a 'one-for-one' replenishment policy. Demand is not met if the required goods are not available at the local warehouse, the central warehouse, or at intermediate stages (in transit). In this study, we propose a simple, fast, and accurate approach to evaluating the service level in this complex system. Unlike other existing methods, we do not require a complex iterative analysis scheme. Our method works perfectly across a wide range of cases, with deviations from exact simulation results averaging less than 0,1%, and in 95% of all test cases, it is less than 0,36%.

Научные основы многоэшелонного управления

Научные основы, модели и алгоритмы многоэшелонного автоматизированного управления потоками ресурсов распределенного промышленного производства требуют точной настройки логистических сетей. Для высокотехнологичных товаров, таких как сложное медицинское оборудование или системы оборонного назначения, простой может иметь серьезные последствия с точки зрения колоссальных производственных потерь и безопасности [1].

Для поддержания таких сложных систем в рабочем состоянии на протяжении всего их срока службы критически важна доступность запасных частей для оперативной замены вышедших из строя компонентов. Однако затраты на хранение соответствующих запасов могут быть огромными, поэтому исключительно важно иметь возможность оптимально размещать запасные части

в разных точках цепочки поставок. Локальные складские точки, расположенные в непосредственной близости от производственной базы, гарантируют быструю доставку, в то время как централизованное хранение позволяет поддерживать относительно низкий уровень страховых запасов за счет эффекта объединения рисков (risk pooling) [2].

В связи с этим возникает острая необходимость глубоко проанализировать эффективность двухэшелонной модели управления запасами, состоящей из различных локальных складов, которые непрерывно пополняются с центрального склада. Для дорогостоящих медленно движущихся деталей в научной практике принято моделировать спрос с помощью распределения Пуассона и применять строгую политику пополнения запасов по принципу «один к одному» (например, Muckstadt, 2005).

В этой статье мы проводим точный математический анализ частоты поступления заказов и их алгоритмической обработки на центральном складе.

Модель и допущения

Мы используем математический индекс $k = 0$ для обозначения склада-распределителя (центрального депо) и индексы $k = 0, \dots, K$ для обозначения местных локальных складов.

Спрос, поступающий на локальный склад k , удовлетворяется через двухэшелонную сеть (которую мы называем обычный канал), если нужный товар физически есть на этом складе, на центральном складе-распределителе или в транспортном логистическом конвейере между ними [3].

Для построения математической модели мы используем следующие строгие обозначения и допущения:

Запасы на складе в каждой точке $k = 0, \dots, K$ контролируются с помощью непрерывной политики управления запасами $(S_k - 1, S_k)$.

Поступающий спрос на местном складе k строго подчиняется закону распределения Пуассона со скоростью (интенсивностью) m_k .

Время пополнения запасов на центральном складе имеет экспоненциальное распределение со средним значением L_0 . Это предположение существенно упрощает последующий математический анализ с использованием цепей Маркова с непрерывным временем. Кроме того, как показывает практика, эффективность таких стохастических моделей с учетом упущенных продаж не сильно зависит от формы распределения времени выполнения заказа (см. Альфредссон и Верридт, 1999).

Время логистической транспортировки со склада-распределителя на локальный склад k является детерминированной величиной и строго равно L_k .

Время срочной аварийной доставки на местный склад k имеет произвольную функцию распределения вероятностей со средним ожидаемым значением T_k .

Нашими главными показателями эффективности (KPI) системы являются:

– Доля спроса на складе k , удовлетворяемого через систему (обычный канал), обозначаемая как α_k .

– Среднее ожидаемое время ожидания удовлетворения спроса через обычный канал $E[W_k]$ для $k = 1, \dots, K$.

Эти два показателя эффективности позволяют точно оценить среднее время простоя промышленного оборудования в ожидании замены критической детали (Downtime Waiting for Part), обозначаемое как $DTWP_k$ на местном складе k . Формула оценки выглядит следующим образом:

$$DTWP_k = \alpha_k E[W_k] + (1 - \alpha_k) T_k. \quad (1)$$

Математический анализ показателей системы

Мы находим среднее время ожидания $E[W_k]$ применяя закон Литтла. Согласно этому закону:

$$E[W_k] = \frac{E[BO_k]}{\alpha_k m_k}, \quad (2)$$

где BO_k обозначает количество отложенных заказов на локальном складе k .

В свою очередь, мы находим математическое ожидание $E[BO_k]$ и долю удовлетворенного спроса α_k по распределению количества отложенных заказов на центральном складе, предназначенных специально для местного склада k . Это является новизной предлагаемого алгоритмического анализа. Дефицит на локальном складе возникает исключительно тогда, когда склад отправляет запрос на пополнение своих запасов на центральный склад, где в этот момент нужного товара физически нет в наличии. Такое распределение позволяет точно определить распределение количества невыполненных заказов на складе k , что, в свою очередь, позволяет определить распределение величины BO_k по всем локальным складам [4].

Кроме того, можно напрямую получить показатель α_k : после того как на местном складе полностью заканчиваются запасы, по регулярному логистическому каналу можно удовлетворить не более S_k дополнительных поступающих запросов (либо напрямую из складских запасов центрального депо, либо из товаров, уже находящихся в транспортном конвейере).

Следовательно, если на складе уже накопилось S_k невыполненных заказов, любой дальнейший спрос на этом узле будет безвозвратно потерян для системы. Таким образом, показатель α_k (доля не потерянного спроса) математически определяется следующим образом:

$$\alpha_k = 1 - \Pr\{PI_k = S_k\}, \quad (3)$$

где PI_k — количество невыполненных заказов в каждом регионе.

Распределение количества невыполненных заказов на складе (PI_0)

В стандартных моделях при полном предварительном заказе можно легко распределить невыполненные заказы по локальным складам: при $y_0 > S_0$ невыполненных заказах условная вероятность наличия x_k невыполненных заказов на складе k просто подчиняется биномиальному распределению с $y_0 - S_0$ независимыми испытаниями и «вероятностью успеха», как показано в работе Грейвса (1985). Однако этот классический подход не относится к представленной модели упущенных продаж [5]. Кроме того, частота поступления новых заказов на пополнение запасов на центральном складе динамически зависит от текущего состояния всей системы: условная вероятность того, что на локальном складе k не

будет товара при условии $PI_0 = y_0$, монотонно увеличивается с ростом y_0 . Из-за этого эффекта частота поступления заказов на пополнение запасов на складе нелинейно уменьшается с ростом y_0 .

Сначала мы покажем, как производить вычисления абсолютно точно. Используя эти вычисленные условные вероятности, мы находим зависящие от состояния показатели частоты прибытия заказов на склад и строим распределение случайной величины PI_0 .

Далее находим безусловное распределение следующим образом:

$$\Pr\{PI_k = y\} = \sum_{y_0=0}^{S^{tot}} \Pr\{PI_k = y | PI_0 = y_0\} \Pr\{PI_0 = y_0\}. \quad (4)$$

В данной формуле переменная S^{tot} – это сокращенное математическое обозначение для суммы базисных уровней запасов всех локальных складов, то есть $S^{tot} = \sum_{k=1}^K S_k$.

Важно отметить, что переменная PI_0 не может превышать это граничное значение: на центральном складе может физически находиться не более S_0 невыполненных заказов на собственное пополнение запасов. После того как на центральном складе полностью закончатся запасы, дальнейшие невыполненные заказы от локальных складов могут быть выполнены только в том случае, если на самих складах еще есть товар, что постепенно приводит к возникновению отложенных заказов уже на локальном уровне. Поскольку количество отложенных заказов на любом локальном складе k по определению не может превышать его базовый запас S_k , мы получаем строгую верхнюю границу PI_0 единицы измерения.

Шаг 1: Вычисление условных вероятностей и распределения

При условии $y_0 > S_0$ на центральном складе в общей сложности имеется ровно $y_0 - S_0$ невыполненных заказов, которые алгоритм должен распределить по K локальным складам. Поскольку на каждом конкретном складе k может быть накоплено не более S_k невыполненных заказов, возможности для комбинаторного распределения строго ограничены, и стандартное биномиальное распределение здесь абсолютно неприменимо.

Для наглядности приведем пример: если $y_0 - S_0 = 5$ невыполненных заказов на центральном складе нужно распределить между тремя локальными складами, при этом ограничение составляет $S_k = 2$ для каждого склада, то возможны только следующие комбинаторные варианты: $\{(1, 2, 2), (2, 1, 2), (2, 2, 1)\}$.

Если бы мы предположили бесконечную вместимость, то есть $S_k \rightarrow \infty \forall k$, то совместное распределение вероятностей при условии $PI_0 = y_0$ представляло бы собой K -категорийное мультиномиальное распределение с $y_0 - S_0$ испытаниями. Обозначим это гипотетическое распределение вероятностей как P_{multi} .

Однако, очевидно, что реальное совместное распределение вероятностей для конечных значений S_k имеет форму усеченного мультиномиального распределения, поскольку мы вводим жесткое условие на верхние границы S_k для количества отложенных заказов депо для каждого склада k :

$$p(x_1, \dots, x_K | PI_0 = y_0) = \frac{P_{multi}(x_1, \dots, x_K)}{C_{norm}}, \quad (5)$$

где C_{norm} – нормировочная константа.

Маргинальные функции плотности вероятности можно выразить через эти совместные вероятности. Чтобы быстро и эффективно численно оценивать $p(x_1, \dots, x_K | PI_0 = y_0)$ без полного перебора, мы внедряем элегантный метод, предложенный Левиным (Levin, 1981).

Пусть $n = y_0 - S_0$ обозначает общее количество отложенных заказов депо. Метод Левина (1981) доказывает следующую зависимость:

$$\Pr\{X_1 = x_1, \dots, X_K = x_K | W = n\} = \frac{\Pr\{W = n\}^{-1} e^{-v} v^n}{n!} \prod_{k=1}^K \Pr\{Z_k = x_k\}. \quad (6)$$

В этой формуле параметр v может быть абсолютно любым действительным числом. Переменные X_K – это независимые случайные величины, распределенные по классическому закону Пуассона, а W – это сумма независимых усеченных случайных величин, также распределенных по закону Пуассона. То есть Z_k подчиняется усеченному распределению Пуассона с верхней границей x_k .

Как утверждает Левин (1981), мы должны рассматривать v как параметр настройки, который может быть свободно выбран для обеспечения максимального удобства и численной стабильности алгоритма. Он настоятельно предлагает установить $v = n$, потому что при таком выборе можно использовать приближение Стирлинга для быстрого вычисления первого члена уравнения. Наконец, вероятность $\Pr\{W = n\}$ можно точно вычислить либо с помощью алгоритма явных сверток (как мы делаем в наших численных экспериментах для максимальной точности), либо с помощью нормального приближения для ускорения.

Шаг 2: Вычисление вероятностей состояний системы $\Pr\{PI_0 = y_0\}$

Учитывая наличие y_0 невыполненных заказов на центральном складе, мы определяем интенсивность (скорость) поступления новых заказов $M(y_0)$ следующим образом. Новый спрос на локальном складе k может быть удовлетворен через обычный канал с определенной вероятностью. В этом случае центральный склад получает новые запросы со склада k со скоростью m_k . Таким образом, мы получаем математическую модель скорости:

$$M(y_0) = \sum_{k=1}^K m_k (1 - \Pr\{PI_k = S_k | PI_0 = y_0\}). \quad (7)$$

Скорость максимальна, когда $y_0 < S_0$, поскольку в этом состоянии на складе еще есть физический товар. Приближая этот сложный процесс поступления к стандартному процессу Пуассона, мы можем графически и математически представить эволюцию переменной PI_0 в виде цепи Маркова с непрерывным временем. На этой диаграмме состояний параметр μ обозначает регулярную скорость пополнения системы (т. е. $\mu = 1/L_0$).

Ожидаемое количество отложенных заказов ($E[BO_k]$)

В итоге мы находим искомую метрику $E[BO_k]$ по полученному распределению общего количества невыполненных заказов PI_k на локальном складе k . Величина PI_k структурно состоит из двух компонентов: (i) товаров, заказанных в центральном депо специально для склада k , и (ii) товаров, уже физически находящихся в пути (в конвейере) со склада-депо на склад k , количество которых мы обозначим как случайную величину Q_k .

В упрощенных моделях при полном обратном заказе эти два элемента считаются абсолютно независимыми (Грейвс, 1985), и PI_k можно вычислить как простую математическую свертку Q_k .

Тем не менее, мы можем достаточно точно аппроксимировать PI_k , обоснованно предположив, что Q_k имеет распределение Пуассона со средним значением $m_k L_k$, и что эти компоненты взаимно независимы в допустимых пределах. Когда склад k достигает предела, он не получит новых заказов, и Q_k будет равно 0. Таким образом, мы получаем выражения для вероятностей и ожиданий:

$$\Pr\{PI_k = y_k\} = \sum_{x_k} \Pr\{PI_0^k = x_k\} \cdot \Pr\{Q_k = y_k - x_k\}, \quad (8)$$

$$E[BO_k] = \sum_{y_k=S_k+1}^{S_k+Q_k^{\max}} (y_k - S_k) \Pr\{PI_k = y_k\}. \quad (9)$$

Чтобы оптимизировать алгоритм и сократить сумму при вычислениях, мы устанавливаем верхнюю границу для PI_k , где $Q_k^{\max}$ является точной верхней границей для товаров в пути.

Выводы и перспективы развития

В ходе данного исследования мы успешно разработали простую, вычислительно легкую и высокоточную математическую аппроксимацию для двухэшелонной модели управления сложными запасами с пуассоновским спросом и упущенными продажами.

Фундаментальное отличие нашего подхода от существующих классических методов заключается в том, что мы не делаем строгих допущений относительно поведения сети при дефиците и не используем ресурсоемкую процедуру, которая обычно применяется в научной литературе.

Список литературы

1. Graves S.C. A multi-echelon inventory model for a repairable item with one-for-one replenishment // Management Science. 1985, vol. 31(10), pp. 1247-1256.
2. Si Thu Thant Sin. Development of a Model and Algorithms for Trigger Control of Technological Processes of Resource Provision // Advances in Automation VI. RusAutoCon 2024. Lecture Notes in Electrical Engineering, vol. 1324. Springer, Cham, 2025. doi.org/10.1007/978-3-031-82494-4_13.
3. Alfredsson P., Verrijdtz J. Modeling emergency supply in a two-echelon inventory system // Management Science. 1999, vol. 45(10), pp. 1416-1431.

4. Si Thu Thant Sin, Portnov E.M., Bain A.M. Research and Development of Algorithms for Dispatching Tasks in Distributed Computing Systems // International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russian Federation, 2024, pp. 700-704.
5. Сумец А.М. Выбор стратегии управления товарными запасами с учетом фактора сезонности // Вестник Полоцкого государственного университета. Серия D: Экономические и юридические науки. – 2010. – №4. – С. 163-170.

Сведения об авторе:

Си Ту Тант Син – к.т.н., докторант.