

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Факультет комп'ютерних наук та кібернетики
Кафедра дослідження операцій

КУРСОВА РОБОТА

на тему:

Аналіз на ПК СМО $M/G/n/m$

студента 3 курсу

Байбула Кирило Аленовича

Науковий керівник:

професор, доктор фізико-математичних наук

Мацак Іван Каленикович

Київ — 2023

Реферат

Обсяг роботи 21 сторінок.

ДОСЛІДЖЕННЯ СИСТЕМ МАСОВОГО ОБСЛУГОВУВАННЯ ІЗ ОБМЕЖЕНОЮ ЧЕРГОЮ $M/G/n/m$.

Об'єктом дослідження є система масового обслуговування з обмеженою чергою $M/G/n/m$. Предметом роботи є розробка імітаційної моделі системи масового обслуговування з обмеженою чергою та аналіз результатів її роботи.

Досліджується вплив зміни параметрів системи на її продуктивність. Зокрема, аналіз того, як швидкість прибуття, швидкість обслуговування та місткість черги впливають на продуктивність системи.

Методи розроблення: комп'ютерне моделювання. Інструменти розроблення: мова програмування Rust та вільно поширюване інтегроване середовище розробки Emacs 28.2.

Зміст

1	Перелік умовних позначень, символів, скорочень	3
2	Вступ	4
3	Теоретичні відомості	6
3.1	СМО типу $M/M/c/K$	6
3.2	СМО типу $M/G/n/t$	9
4	Моделювання СМО	10
4.1	Імітаційна модель	10
4.2	Моделюючий алгоритм	10
5	Результати моделювання	13
5.1	$G(x)$ — експоненційний розподіл	13
6	Висновки	20

1 Перелік умовних позначень, символів, скорочень

СМО : система масового обслуговування.

$\Theta(t)$: довжина черги (стан) СМО в момент часу t ;

P_k : стаціонарна ймовірність того, що в системі є k запитів;

W_i : час очікування i -ого запиту у черзі;

2 Вступ

Курсова робота присвячена дослідженню систем масового обслуговування із обмеженою чергою. Подібні системи використовуються для моделювання різноманітних процесів, де можуть виникати черги, наприклад, у телекомунікаційних мережах, медичних установах, аеропортах, ресторанах, тощо.

Подібні системи ми будемо називати СМО із обмеженою чергою $M/G/n/m$, якщо:

1. Запити до СМО надходять як **Макрківський** (Пуассонівський) процес, що має параметр $\lambda > 0$ (тобто $\tau_{i+1} = t_{i+1} - t_i$ — час між надходженнями i -ого та $i + 1$ запиту має експоненційний розподіл з параметром λ);
2. Час обробки кожного запиту СМО визначен випадковою величиною ξ , що має довільний розподіл $G(x)$.
3. СМО має n приладів, що обробляють запити із черги;
4. СМО має чергу з довжиною не більше ніж m (тобто система не може тримати більше ніж m запитів у черзі, а максимальна кількість запитів, що може бути прийнята системою, дорівнює сумі кількості приладів і максимальної довжини черги, тобто — $n + m$);

Кожен новий запит, що надходить у систему, спочатку призначається для будь-якого незайнятого приладу обслуговування. Інакше, якщо всі прилади зайняті, запит стає у чергу для подальшого оброблення.

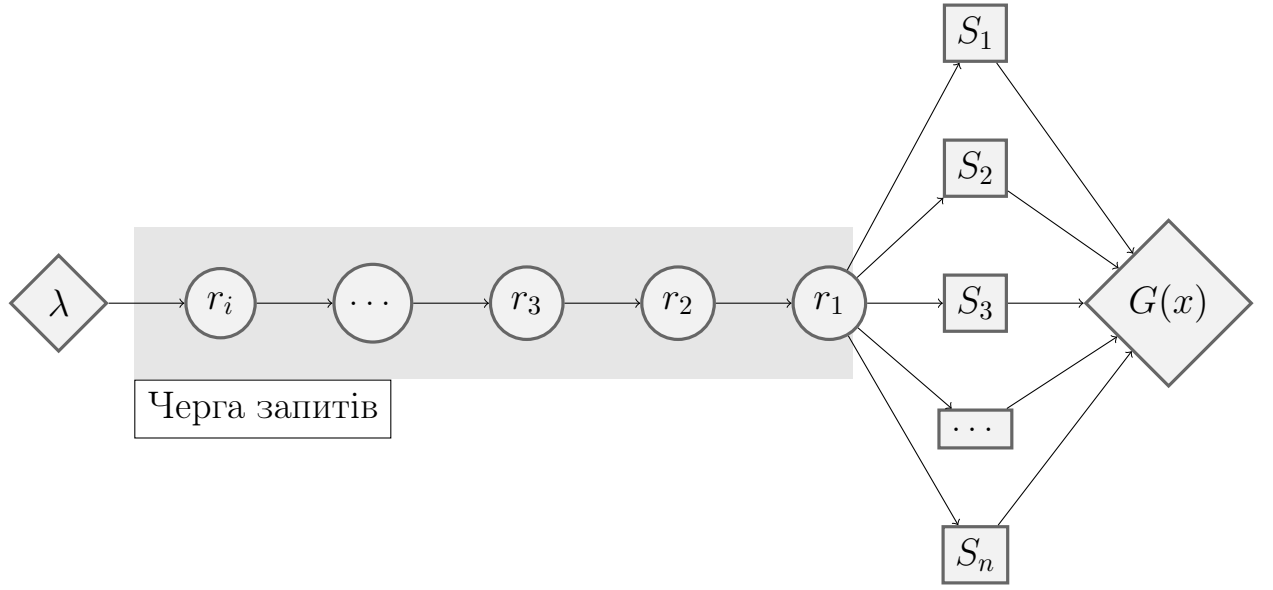


Рис. 1: Схема роботи $M/G/n/m$, де r_i — i -тий запит у черзі, S_n — n -тий прилад.

Головною метою данної роботи є знаходження стаціонарних характеристик СМО, таких як:

1. $P_k = P\{\Theta = k\} = \lim_{t \rightarrow \infty} P\{\Theta(t) = k\}, k = 0, 1, 2 \dots n + m;$
2. $M\Theta_k = \lim_{t \rightarrow \infty} M\Theta(t);$
3. $MW = \lim_{k \rightarrow \infty} W_k;$

Тобто, ми хочемо знайти стаціонарний розподіл кількості запитів у системі, середню кількість запитів у системі та середній час очікування запиту у черзі.

Через складність дослідження та аналізу подібної СМО, ми будемо використовувати метод Монте-Карло, тобто змоделюємо та запрограмуємо систему за допомогою комп'ютера та зробимо велику кількість експериментів, що дозволить нам знайти та проаналізувати бажані характеристики.

3 Теоретичні відомості

У цьому розділі будуть предоставлені короткі теоретичні відомості про системи масового обслуговування з переглядом їх основних видів, що надасть нам фундамент для дослідження та аналізу нашої СМО.

3.1 СМО типу $M/M/c/K$

Цікавим для огляду буде система масового обслуговування $M/M/c/K$, що має так само на вхід Марківський потік запитів та обмежену кількість приладів обслуговування кількістю c та скінчену чергу максимальною довжиною K , але, на відмінну від $M/G/n/m$, має експоненціальний час обслуговування із параметром μ . Тобто, $M/M/c/K$ — це частковий випадок $M/G/n/m$.

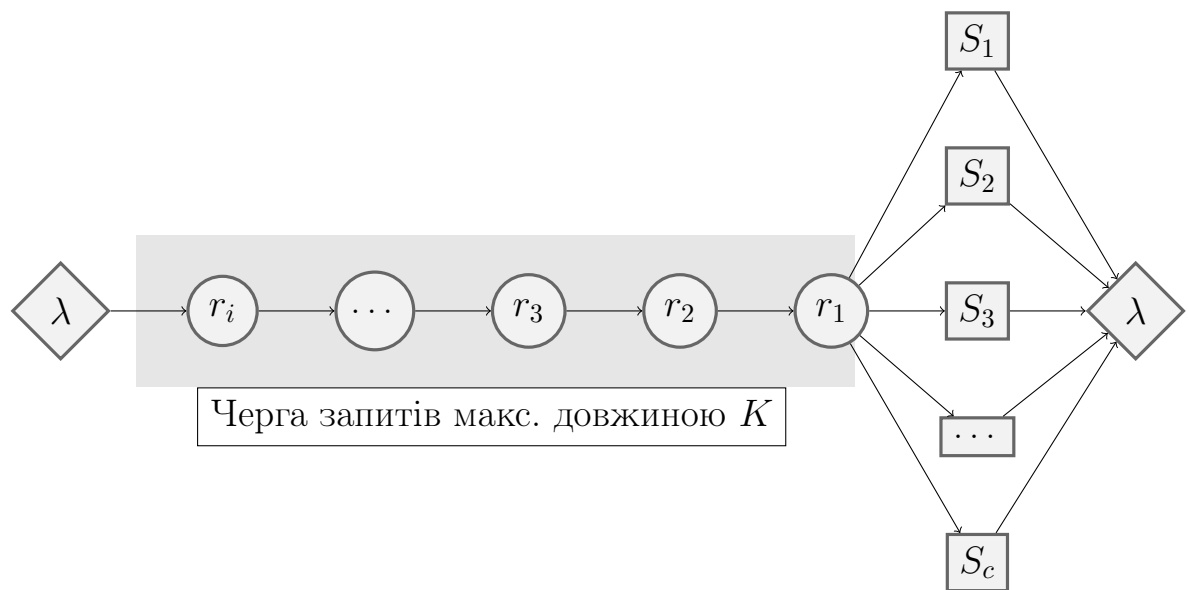


Рис. 2: Схема роботи $M/G/c/K$, де r_i — i -тий запит у черзі, S_c — c -тий прилад.

Ця модель може бути представлена у вигляді ланцюга Марківських станів, із матрицею переходів:

$$Q = \begin{pmatrix} -\lambda & \lambda & & & & & \\ \mu & -(\lambda + \mu) & \lambda & & & & \\ & 2\mu & -(\lambda + 2\mu) & \lambda & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & c\mu & -(\lambda + c\mu) & \lambda & \\ & & & & \ddots & \ddots & \ddots \\ & & & & & c\mu & -c\mu \end{pmatrix}$$

на просторі станів $\{0, 1, 2, \dots, K\}$. Подібний ланцюг називають процесом “смерті-народження” (birth-death process), де переходи бувають лише двох видів: *народження*, що збільшує змінну стану на одиницю, та *смерть*, що зменшує її на одиницю [3]. Саме тому, простіше представити матрицю переходів у вигляді машини станів:

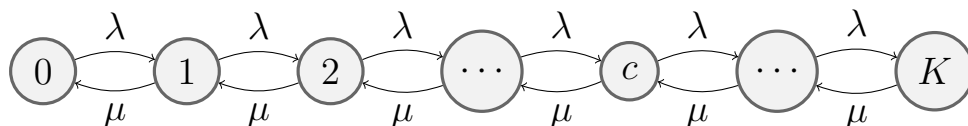


Рис. 3: Машина станів $M/M/c/K$

Також важливою характеристикою СМО, є **утілізація**. Утілізація — це відношення часу, коли пристрій зайнятий, до загального часу. Для $M/M/c/K$ утілізація має вигляд:

$$\rho = \frac{\lambda}{c\mu} \quad (1)$$

При цьому ж, ρ має обмеження $\rho < 1$, що відповідає тому, що система не може обслуговувати запити швидше, ніж вони надходять. Інакше, система буде постійно переповнювати чергу і губити запити.

Стационарний розподіл кількості запитів у системі P_k для $M/M/c/K$ був наведений ще у 1981 році [1] та має вигляд:

$$P_0 = \left[\sum_{k=0}^c \frac{\lambda^k}{\mu^k k!} + \frac{\lambda^c}{\mu^c c!} \sum_{k=c+1}^K \frac{\lambda^{k-c}}{\mu^{k-c} c^{k-c}} \right]^{-1}$$

$$P_k = \begin{cases} \frac{\lambda^k}{\mu^k k!} P_0, & k = 1, 2, \dots, c \\ \frac{\lambda^k}{\mu^k c^{k-c} c!} P_0, & k = c+1, c+2, \dots, K \end{cases} \quad (2)$$

Середня кількість часу проведеного запитом у системі:

$$W = \frac{1}{\mu} + P_0 \frac{\rho(c\rho)^c}{\lambda(1-\rho)^2 c!} \quad (3)$$

Середня кількість часу очікування запиту у черзі:

$$W_q = P_0 \frac{\rho(c\rho)^c}{\lambda(1-\rho)^2 c!} \quad (4)$$

І звідси ж, за формулою Літтла, середня кількість запитів у системі:

$$\mathbb{M}\Theta = \lambda W = \frac{\lambda}{\mu} + P_0 \frac{\rho(c\rho)^c}{(1-\rho)^2 c!} \quad (5)$$

3.2 СМО типу $M/G/n/t$

Як було зазначено вище, СМО типу $M/G/n/t$ є більш загальним випадком СМО типу $M/M/c/K$ (див 3.1), де G — це загальний розподіл часу обслуговування запиту (також див. 2). Саме тому, $M/G/n/t$ має багато спільних властивостей з $M/M/c/K$, але, через невизначеність, загальність G , $M/G/n/t$ — є важким для аналітичного дослідження.

Tijms та ін. вважають, що *"малоймовірно, що можна розробити методи для обчислення точних числових значень стаціонарної ймовірності в черзі $M/G/k$ "* [4].

Саме тому, для $M/G/n/t$ спробуємо використати інший підхід, а саме — моделювання.

4 Моделювання СМО

У даному розділі буде розглянуто підходи до моделювання СМО, що дозволить нам вибрати найбільш оптимальний підхід для нашої задачі. Починаючи з огляду архітектури системи, закінчуючи описом алгоритму моделювання.

4.1 Імітаційна модель

Одним із підходів, зазначених у [6], є побудова моделі, що буде працювати з імітованим, дискретним часом, де кожен крок системи є фіксованим інтервалом. Неперервні розподіли, таким чином, стають дискретними, що призводить до округлення помилок. І звідси, єдиний спосіб зменшити розмір помилки — зменшити крок часу. З іншого боку, з цього випливає і збільшення часу роботи програмної моделі.

Саме тому, у [5] було запропоновано імітаційну модель, що працює зі змінними кроками часу, що базуються на подіях, що виникають у самій системі. Тобто кожен наступний крок — стрибок у часі від останньої події до найближчої наступної. Тим самим ми пришвидшуємо роботу імітаційної моделі у декілька разів та позбавляємося від дискретності у часі.

4.2 Моделюючий алгоритм

Параметрами нашої моделі будуть:

1. n — кількість приладів обслуговування;
2. m — довжина черги;

3. λ — інтенсивність надходження запитів;
4. $G(x)$ — розподіл часу обслуговування запитів. На практиці буде або експоненційним (завдяки чому система буде перетворюватися у СМО типу $M/M/c$ [3]), або рівномірним;
5. T — час життя моделі.

Також головними об'єктами моделювання є:

1. t — теперішній час моделі;
2. Q — черга запитів;
3. E — відсортована по часу черга подій. Елементи цієї черги бувають двох видів: запит, що прибув у систему, та запит, що вийшов з системи;
4. S — множина приладів обслуговування;
5. $r_i(t)$ — i -тий запит, що має параметр t — час потрібний приладу для обслуговування запитів;
6. $Rand_P$ — підпрограма, що генерує випадкові числа з експоненційним розподілом, що визначають наступний час прибуття запиту;
7. $Rand_G$ — підпрограма, що генерує випадкові числа з розподілом $G(x)$, що визначають час обслуговування запиту.

Таким чином, все життя системи буде складатися з послідовного ітерування по подіям із черги E . Якщо подія — прибуття запиту, то він додається до черги Q , і, за допомогою підпрограми $Rand_P$,

генерується час наступного прибуття запиту. що також додається до черги E . Якщо подія — виходження запиту, то запит позначається як виконаний, і, якщо черга Q не пуста, то з неї видаляється наступний запит, а у чергу E додається подія виходження цього запиту, час котрого визначається за допомогою підпрограми $Rand_G$.

Приклад імплементації можна побачити у репозиторії [2].

5 Результати моделювання

Провівши декілька експериментів на імітаційній моделі із фіксованими параметрами, можемо розглянути результати.

5.1 $G(x)$ — експоненційний розподіл

Розглянемо випадок, коли $G(x)$ — експоненційний розподіл. Фактично, це є частковим випадком СМО $M/M/c/K$, що описуна у 3.1.

Змоделюємо ситуацію з $n = 5$ приладами обслуговування, довжиною черги $m = 30$, очікуваним часом обслуговування $1/\lambda = 26$ секунд, з очікуваним періодом надходження запитів $1/\mu = 4$ секунд.

При $T = 10,000$ секунд, розглянемо кількість запитів в залежності від часу (рис. 4). Бачимо, що кількість запитів у системі коливається від 25 до 35 із частими піками саме до 35, що свідчить про втрати, пов'язані з переповненням черги. Що і не дивно, враховуючи утилізацію $M/M/c/k$ (3.1) :

$$\rho = \frac{\lambda}{c\mu} = \frac{26}{5 \cdot 4} = 1.3 \not< 1$$

Подібну систему не можна вважати ефективною, саме тому спробуймо зменшити очікуваний час обслуговування так, щоб враховувалася умова 3.1.

При $1/\lambda = 60$, $n = 3$, $1/\mu = 50$, $m = 8$, $T = 10,000$ секунд, утилізація становитиме: $\rho = 0.27 < 1$. З рис. 5 бачимо графік станів із піками, що не перевищують максимальну кількість запитів у системі, тобто система не втрачає запити, а отже є ефективною.

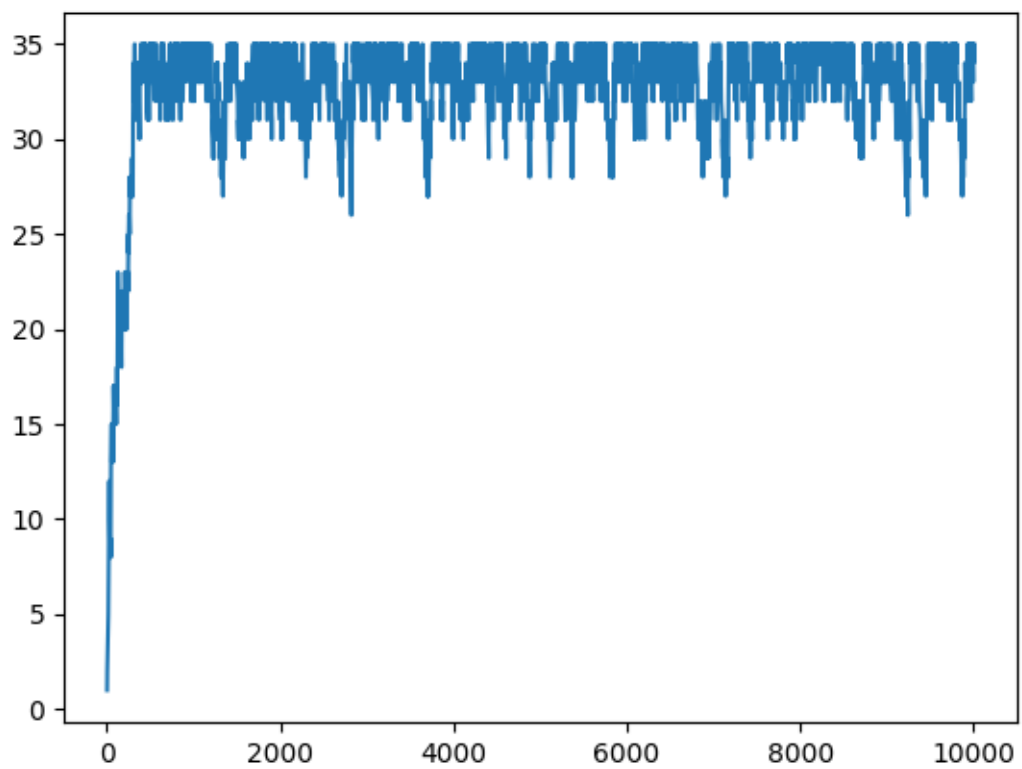


Рис. 4: Залежність кількості запитів від часу при $1/\lambda = 17$, $n = 5$, $1/\mu = 4$

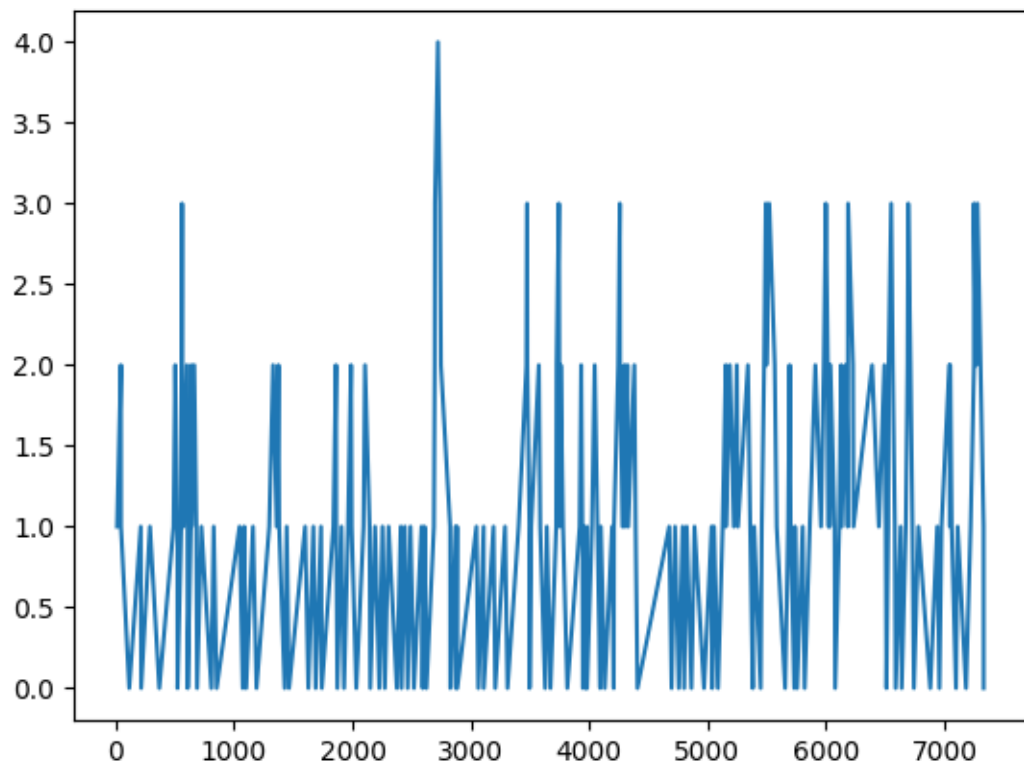


Рис. 5: Залежність кількості запитів від часу при $1/\lambda = 60$, $n = 3$, $1/\mu = 50$

Розглянемо середній час очікування запиту у черзі (рис. 6). Бачимо, невизначені коливання.

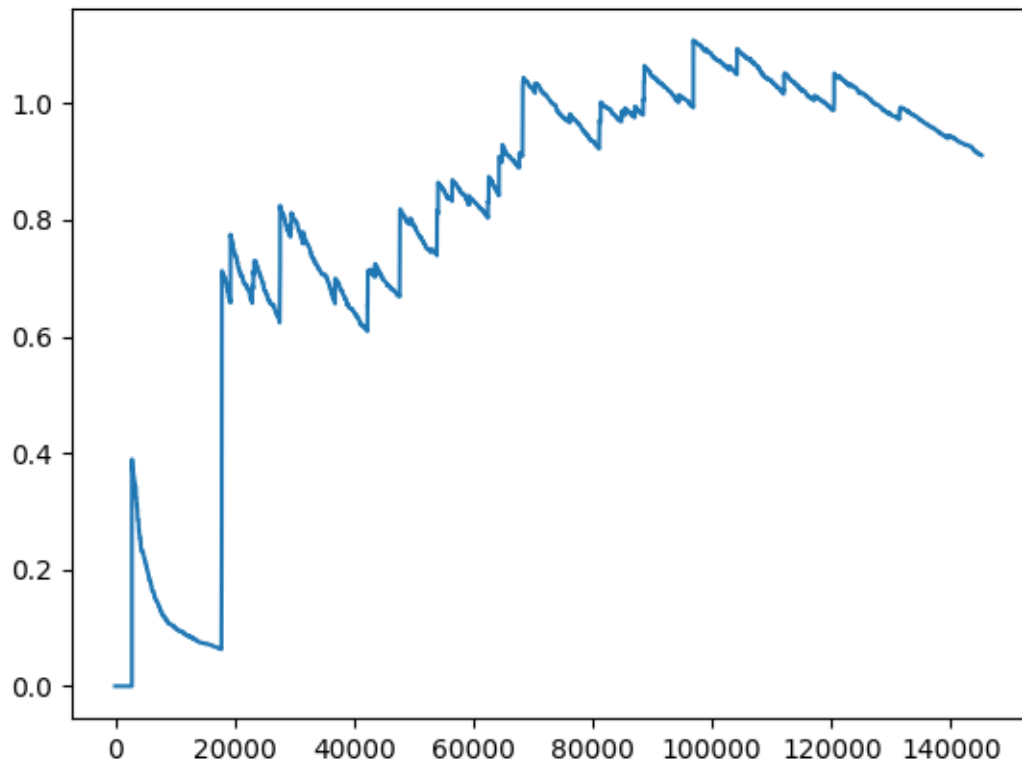


Рис. 6: Залежність середньої кількості часу у черзі від часу при $1/\lambda = 60$, $n = 3$, $1/\mu = 50$ та $T = 140,000$

Так як, нам потрібно розглянути стаціонарні характеристики при $t \rightarrow \infty$, спробуємо смодельовати систему при $T = 100,000,000$ (рис. 7). Звідси, бачимо, кумулятивне середнє при $t \rightarrow \infty$ наближається до фіксованого значення, а отже границя $\mathbb{M}W = \lim_{k \rightarrow \infty} W_k$ існує.

Нам відомо, що при $G(x)$ — експоненційний розподіл $M/G/n/m$ — це $M/M/c/K$ (див. 3.1), спробуємо знайти W_q за формулою (3.1)

(розрахунки див. [2] у директорії *notebooks*). Звідси:

$$W_q = 1.331806030711408$$

що позначено червоною лінією на рис. 8.

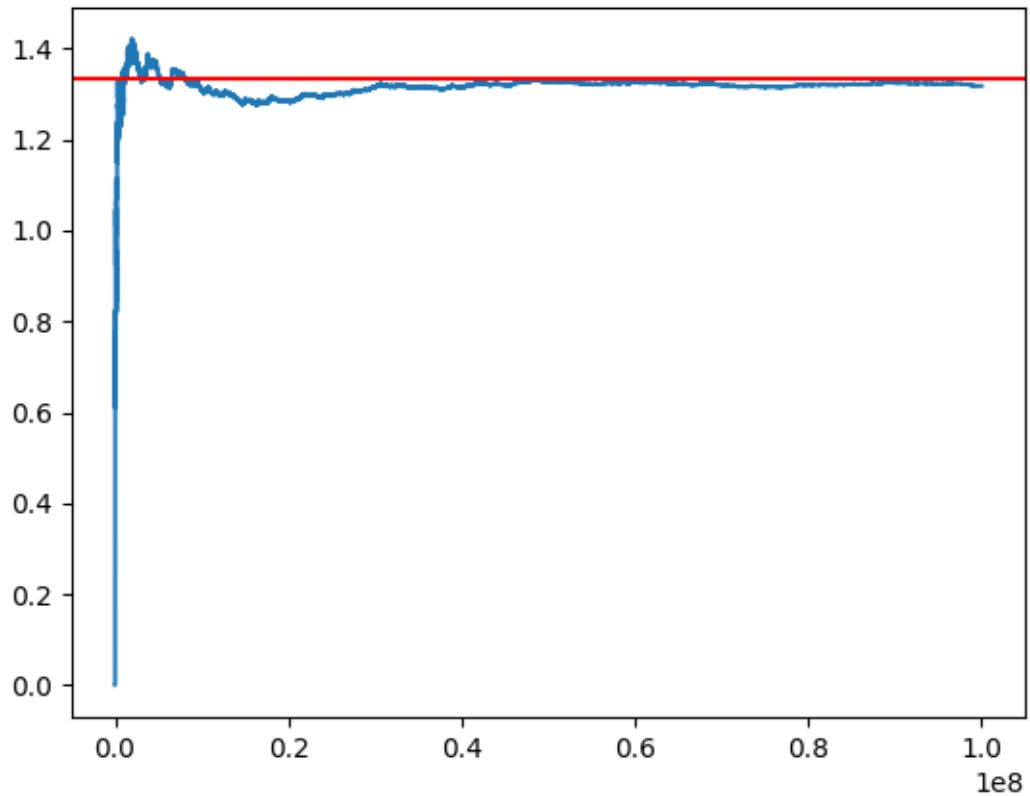


Рис. 7: Залежність середньої кількості часу у черзі від часу при $1/\lambda = 60$, $n = 3$, $1/\mu = 50$ та $T = 100,000,000$. Червона лінія $y = W_q$

Аналогічні результати для середньої кількості запитів у системі (рис. 8) і також, для стаціонарного розподілу P_n (див. 9).

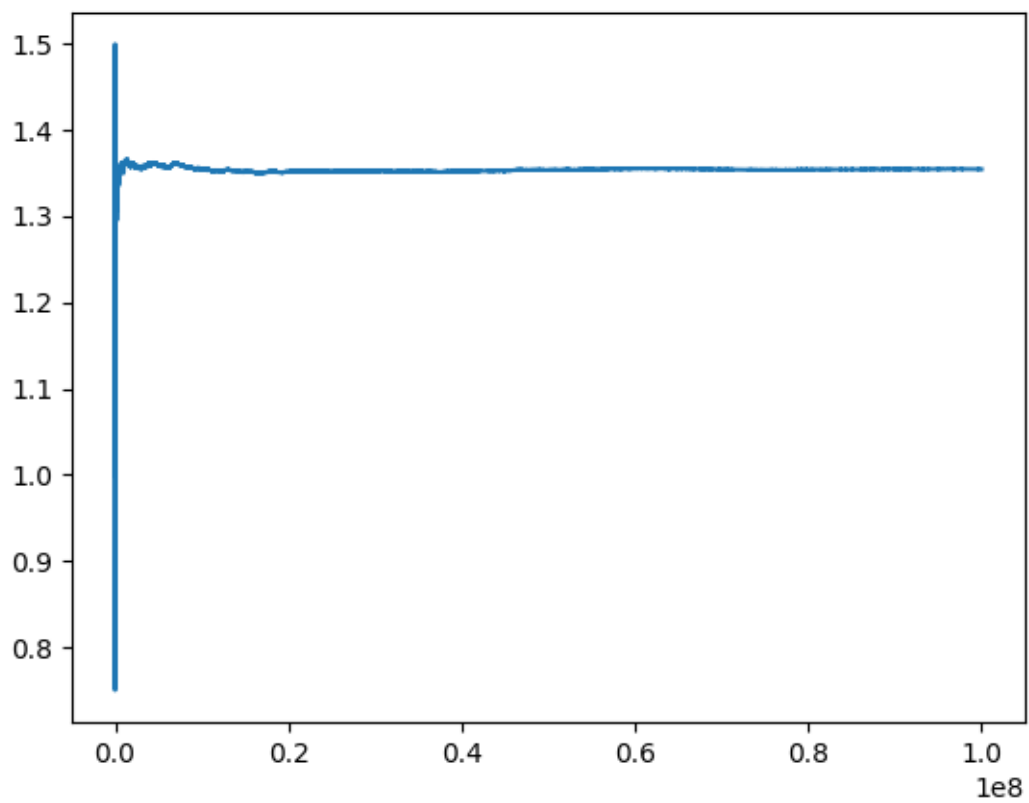


Рис. 8: Залежність кумулятивної середньої кількості запитів у черзі від часу при $1/\lambda = 17$, $n = 5$, $1/\mu = 4$ та $T = 100,000,000$

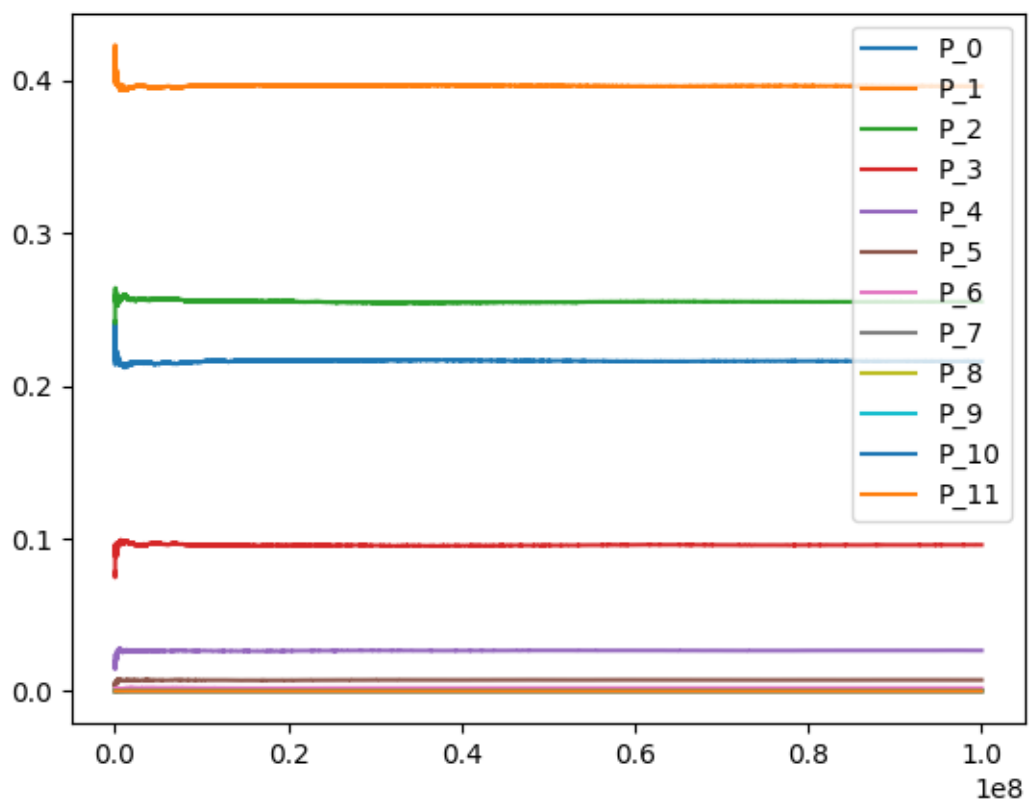


Рис. 9: Графі ймовіріості від часу при $1/\lambda = 17$, $n = 5$, $1/\mu = 4$ та $T = 100,000,000$

6 Висновки

Аналізуючи дані експерименту з моделлю $M/G/n/t$ чергової системи, ми дійшли до висновку, що через загальну природу СМО, з невизначеним розподілом $G(x)$, для ефективного аналізу краще використовувати наступні підходи:

1. Розглянути всі можливі випадки розподілу $G(x)$. Проаналізувати кожен окремо, як одну з вже існуючих моделей черг, по типу: $M/M/c$, $M/D/c$ і тому подібне.
2. Побудувати імітаційну модель СМО, яка відповідає нашій системі. І використовувати її для аналізу.

Імітаційне моделювання для таких складних випадків як СМО $M/G/n/t$ досі є ефективним і майже чи не єдиним інструментом для дослідження. Саме тому, для даної курсової роботи, ми використали саме його.

Результати написаної програми, яка моделює СМО $M/G/n/t$, показали, що на характеристики системи впливає низка факторів, таких як: інтенсивність вхідного потоку, інтенсивність обслуговування, кількість місць у черзі та кількість обслуговуючих приладів.

Стационарні характеристики із часом збігаються з теоретичними, що підтверджує правильність роботи програми.

Література

- [1] Allen. *Probability, Statistics and Queueing Theory, with Computer Science Applications*. 1981, c. 679—680. DOI: 10.2307/2287863.
- [2] Kyrylo Baybula. *Analyze on PC M/G/n/m*. <https://github.com/Velnbur/mgnm-queue-modeling>. Вер. 0.1.0. Трав. 2023.
- [3] Willy Feller. *Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung*. Т. 5. 1939, c. 11—40. DOI: 10.1007/bf01602932.
- [4] M. H. Van Hoorn H. C. Tijms та A. Federgruen. “Approximations for the Steady-State Probabilities in the M/G/c Queue”. В: *Advances in Applied Probability* 13 (1981), c. 186—206. DOI: 10.2307/1426474.
- [5] M. van Tits та H. van der Veeken. “Simulation of a queueing problem with balking”. В: *SIMULATION: Transactions of the Society for Modeling and Simulation* 35 (1980), c. 89. DOI: 10.1177/003754978003500303.
- [6] John Wiley та Sons. “Computer Simulation Techniques”. В: *Journal of the Franklin Institute* 288.10 (1968). DOI: 10.1016/0016-0032(69)00211-0.