Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»

Кафедра мережних технологій факультету інформатики

КЕРУВАННЯ В МЕРЕЖАХ ҐОРДОНА - НЬЮЕЛА
(CONTROL IN GORDON-NEWELL NETWORKS)

Текстова частина до курсової роботи
за спеціальністю „Системний аналіз” 6.050124

Керівник курсової роботи
к.т.н., доц. Чорней Р. К.
(прізвище та ініціали)

____
(підпис)
“10” травня 2021 р.

Виконав студент
Степанюк Р. А.
(прізвище та ініціали)
“10” травня 2021 р.

Київ 2021

Table of contents

List of Abbreviations

GNN – **G**ordon-**N**ewell **N**etwork;

FCFS – **F**irst **C**ome **F**irst **S**erve;

CTC – **C**openhagen **T**elephone **C**ompany;

MIT – **M**assachusetts **I**nstitute of **T**echnology;

UCLA – **U**niversity of **C**alifornia, **L**os **A**ngeles;

LDM – **L**ocal **D**ecision-**M**aker;

1. Introduction

1.1. Problem statement

Queues are frequently used in human-made systems. Some prominent examples are: going shopping, ordering food, or simply waiting for a request to proceed on a server. A queueing process is a model of waiting lines. It is constructed to predict queue length and waiting times. Connected queues make queueing networks in which these queues are linked by what is known as customer routing. The field of mathematical study known as *queueing theory* constructs models that help predict the waiting time, resolve issues and improve the efficiency of queueing processes. In 1967 W. J. Gordon and G. F. Newell published a paper that studied a particular case of queueing networks called closed queueing networks of exponential servers. The application of that study is used to this day.

1.2. Goals

As more and more companies are restructuring their business models to be more automized, the question of efficient usage of available capabilities arises. With the help of queuing networks, the only thing the customer has to do to get helpful information regarding the improvement of queueing processes in their company is to input the data that describes the system. As a result, they can be provided numbers, charts, and other data that will drastically improve the working process.

By implying strategy improvement procedure on local decision-makers, optimal local strategy can be derived, improving the working process flow in the network.

1.3. Real-life applications

Let us consider the following system that can be described as queueing network. There is a manufacturer factory that uses six robots when manufacturing products. The breakdown times for these robots are exponentially distributed with a mean of twenty-five hours. Two repairmen can fix the robots, and the times it takes to repair them are exponentially distributed with a mean of four hours.

Many valuable metrics can be obtained from such a factory's working process using the Gordon-Newell network. With two states of the system described as "working" and "broken," some valuable estimates can be calculated. For example, the average number of operational robots at any given time or percentage of time that all machines were working. Also, the Gordon-Newell theorem can help to find bottlenecks in the system and potentially improve its productivity.

2. Related work

2.1. Queueing models classification

In 1951, David George Kendall created the standard system that is being used to classify and describe a queueing model. It is called *Kendall's notation*. At first, this notation had three main factors:

- The arrival process (A) - denotes the time between arrivals to the queue, uses codes specified in *Table 1*.
- The service time distribution (S) - gives the distribution of time of the service of a customer, uses codes specified in *Table 1*.
- The number of servers (c) - specifies the number of service channels open at the node.

Later this notation was extended, and the following factors were added:

- The capacity of the queue (K) - the maximum number of customers allowed in the queue. Default is ∞;
- The calling population (N) - the size of the population of jobs to be served. Default is ∞;
- The queueing discipline (D) - the Service Discipline or Priority order that jobs in the queue, or waiting line, are served. Codes specified in *Table 2*. Default is *FCFS.*

Symbol	Name	Description (arrival process)	Description (service time distribution)
M	Markovian	Poisson process or random arrival process	Exponential service time
D	Degenerate distribution	A deterministic or fixed inter-arrival time	A deterministic or fixed service time
$MMPP$	Markov modulated poisson process	Poisson process where arrivals are in "clusters"	Exponential service time distributions, where the rate parameter is controlled by a Markov chain
G	General distribution	General distribution	General distribution
E_k	Erlang distribution	An Erlang distribution with k as the shape parameter	An Erlang distribution with k as the shape parameter
PH	Phase-type distribution	Some of the above distributions are special cases of the phase-type	Some of the above distributions are special cases of the phase-type

Table 1: Basic codes for queueing models

Symbol	Name	Description
FCFS	First Come First Served	The customers are served in the order they arrived in
LCFS	Last Come First Served	The customers are served in the reverse order to the order they arrived in
SIRO	Service In Random Order	The customers are served in a random order with no regard to arrival order
PQ	Priority Queuing	There are several options: Preemptive Priority Queuing, Non Preemptive Queuing, Class Based Weighted Fair Queuing, Weighted Fair Queuing

PS	Processor Sharing	The customers are served in the determine order with no regard of arrival order

Table 2: The queue's disciplines

These notations are used to describe queueing models. The main focus of this paper will be M/M/c/k queue. It represents the queue where a Poisson process determines arrivals, job service times have an exponential distribution, and the system has *c* servers. Such a model is one of the most elementary queueing models and has many real-life applications, making it very attractive to study in the field of queueing theory.

Figure 1: M/M/1 queueing model

In M/M/c/k/0/FCFS, queue arrivals occur according to a Poisson process at rate λ and move the process from node *i* to node *i + 1* based on the number of services *c*. Rate parameter μ is used to determine service times using exponential distribution, which will have an expected service time equal to 1/μ. The number of customers is equal to *k*. The pool of arriving customers is equal to zero. *FCFS* discipline, also known as "First in - first out," states that the customers that arrived at the node first will leave the node sooner than those who arrived later.

2.2. Development of queueing theory

Agner Krarup Erlang is considered to be the father of queueing theory. Despite his early death at the age of 51 in 1929, this danish mathematician published numerous significant papers, some of which include:

- 1909 – "The Theory of Probabilities and Telephone Conversations," which proves that the Poisson distribution applies to random telephone traffic;
- 1917 – "Solution of some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges," which contains his classic formulae for call loss and waiting time;
- 1920 - "Telephone waiting times," which is Erlang's principal work on waiting times, assuming constant holding times.

The main goal of his studies was to improve the working process at telephone traffic since, for a significant part of his life, he was working for the CTC (from 1908 to 1929).

Some other well-known names in queueing theory are Felix Pollaczek, who solved M/G/1 queue. Later with Aleksandr Khinchin, he created Pollaczek–Khinchine formula. It states a relationship between the queue length and service time distribution Laplace transforms for an M/G/1 queue. Later, Pollaczek also studied the GI/G/1 queue using an integral equation.

David George Kendall, known for introducing queueing notations, solved the GI/M/k queue in 1957. From 1962 till his retirement in 1985, he was the Professor of Mathematical Statistics in the Statistical Laboratory, University of Cambridge.

Kingman's formula, published in 1961 by John Kingman, gave a formula for the mean waiting time in a G/G/1 queue.

An Institute Professor at MIT (the Massachusetts Institute of Technology), John Little, developed in 1961 Little's law:

$$L = \lambda W$$

It states: "The average number of customers in a system over some interval(L) is equal to their average arrival rate(λ), multiplied by their average time in the system(W)." Little's law is one of the fundamental laws in queueing theory.

Leonard Kleinrock is often called the modern father figure of queueing theory. He is a professor at UCLA's Henry Samueli School of Engineering and Applied Science. His initial contribution to the field of queueing theory was his doctoral thesis proposal in 1961 and led to a doctoral thesis at MIT in 1962, which was later published in book form in 1964. In this work, he analyzed queueing delays in Plan 55-A, a message switching system operated by Western Union for processing telegrams. His theoretical work, which was published in the early 1970s, underpinned packet switching in the ARPANET, a forerunner to the Internet. Kleinrock later published several more of the standard works on the subject.

3. Queueing networks

3.1. Categories of queueing networks

There are several categories of queueing networks. These include open, closed, and mixed queueing networks.

- Open networks are where new customers can enter the system, and under certain conditions, it is possible to leave the network. These types of systems are the most useful due to their real-life applications.
- Closed networks have a static number of customers, which can not change. Customers constantly are moving from one node to another over time.
- The network is mixed if it is open for some classes of customers and closed for others.

3.2. Jacksonian Networks

While at UCLA, James R. Jackson developed a paper about so-called Jackson's (or Jacksonian) networks. These are queueing networks where it is particularly simple to compute the equilibrium distribution as the network has a product-form solution. For a queueing network to be considered Jacksonian, it has to be:

1. open, meaning any external arrivals of customers from a Poisson process;
2. service discipline is FCFS;
3. all service times are exponentially distributed;
4. after completing the process in a node, each customer will either move to another one or leave the system altogether;
5. the utilization of all of the queues must be less than one.

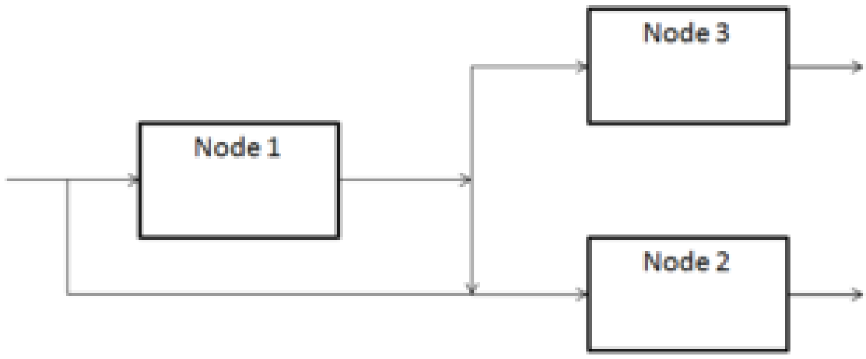


Figure 2: A three-node open Jackson network

3.3. Gordon-Newell Networks

Gordon-Newell theorem was published in 1967 by W. J. Gordon and G. F. Newell in their paper called "Closed queueing networks with exponential servers." Since Jackson's theorem could be applied only to open networks, Gordon and Newell extended it to closed queueing networks of exponential servers, meaning that after the open network solution has been calculated, the infeasible states are eliminated by renormalizing the probabilities.

Gordon-Newell network is a closed queued interconnected network with J nodes and a total population of K circulating individuals. It also uses the FCFS discipline. The state of such a network can be described by a state space $S(K, J) = (n_1, n_2, ..., n_J)$, where each n_j is the number of customers in a queue in server *j*. This means that the sum of individuals in queues over all nodes at a given time *t* is equal to the number of customers in the system:

$$\sum_{j=1}^{J} n_j = K$$

Time which customers spend at nodes is random and based on exponential distribution with a mean $\frac{1}{\lambda}$. Let $\xi = \{\xi^t, t = 0, 1,...\}$be a discrete time Markov process. Also, the probability that the customer that completed the process in given node i will proceed to node j is equal to r_{ij}. This forms a transition Markov matrix, which will have the following form:

$P = (p_{ij}: i, j = 1, \ldots, J)$.

In the case of using the M/M/J/K/0/FCFS model, the customer will only be able to move from node j to node $j + 1$. The probability of this is equal to $p_{j\,j+1}$, which means that with probability $1 - p_{j\,j+1}$ the customer will not move to the next node and instead will stay at the current node for at least one more time slot. If the individual reached the last node and successfully left it, it moves to the first one.

Now, following the Gordon-Newell theorem, the equilibrium distribution of customers in the network is given by:

$$\pi(n_1, \ldots, n_J) = G(K)^{-1} \prod_{j=1}^{J} \left(\frac{e_j}{\mu_j}\right)^{n_j} \qquad (1)$$

Here e_i is the visit ratio. $(e_1, e_2, \ldots, e_J)$ is a real positive solution to the eigenvector-like equations:

$$e_i = \sum_{j=1}^{J} e_j p_{ji}, \text{ for } 1 \le i \le J \qquad (2)$$

Furthermore, $G(K)$ is the normalizing constant, which is defined so that all the $\pi(n_1, \ldots, n_j)$sum to one.

$$G(K) = \sum_{n \in S(K,J)} \prod_{j=1}^{J} \left(\frac{e_j}{\mu_j}\right)^{n_j} \qquad (3)$$

It is important to note that the summation in the above formula (3) is taken over all C_{J+K-1}^{K} possible system states $(n_1, n_2, \ldots, n_J)$.

The solution presented in equation (1) is actually a particular case of the results obtained by Newell and Gordon since it is assumed in (1) that a node's mean service time is independent of the number of customers present. Networks containing facilities with load-dependent service times have somewhat different computational aspects and are treated differently.

3.4. Derived distributions of GNN

Now, if it is necessary to calculate the probability that at the facility i there are precisely k customers, the following formula is used:

$$\pi(n_i = k) = \sum_{n \in S(K,J)\ \&\&\ n_i = k} \pi(n_1,\dots, n_J) \qquad (4)$$

According to Jeffrey P. Buzen, a famous contributor to queueing theory, rather than calculating (4) directly, it is better to calculate the following (5) first:

$$\pi(n_i \geq k) = \sum_{n \in S(K,J)\ \&\&\ n_i \geq k} \pi(n_1,\dots, n_J) = \sum_{n \in S(K,J)\ \&\&\ n_i \geq k} G(K)^{-1} \prod_{j=1}^{J} (\frac{e_j}{\mu_j})^{n_j} =$$

$$= G(K)^{-1} \prod_{j=1}^{J} (\frac{e_i}{\mu_i})^{k} \sum_{n \in S(K-k,J)} \prod_{j=1}^{J} (\frac{e_j}{\mu_j})^{n_j} = (\frac{e_i}{\mu_i})^{k} \frac{G(K-k)}{G(K)} \qquad (5)\text{, then:}$$

$$\pi(n_i = k) = G(K)^{-1} (\frac{e_i}{\mu_i})^{n_k} [G(K - k) - \frac{e_i}{\mu_i} G(K - k - 1)] \qquad (6)$$

$$G(n) = 0\ if\ n < 0$$

From equation (6):

$$E(n_i) = \sum_{k=1}^{J} (\frac{e_i}{\mu_i})^{k} \frac{G(K-k)}{G(K)} \qquad (7)$$

As soon as $G(j)\ for\ 1 < j < J$ are calculated, it is possible to efficiently compute a number of potentially helpful network characteristics. This can be done using equations (1), (5), (6), and (7).

3.5. Controlled processes in GNN

If local decision-makers present at the nodes, they can adapt service probabilities according to the loading of their node and the adjacent ones. By using deterministic policies, it is possible to obtain the optimal behavior of the LDMs. U denotes local decision spaces as:

$$U = \{u^1, u^2, \ldots, u^n\}$$

These spaces are independent of time and finite.

Optimal local strategy δ^* can be found in the class of admissible deterministic stationary local Markov strategies with time-invariant restriction sets. Local Markov strategy, denoted by $\delta = (\delta_1, \delta_2, \ldots \delta_J)$ where δ_j are local decision strategies of history-dependent decisions. They consist of Δ_j, which are the general unrestricted history dependent decisions.

One-step cost $r(x^t, u^t) \geq 0$ is incurred in the network if the decision u^t is made. R_y^{δ} is asymptotic average expected costs. The problem is to find strategy that minimizes it.

The controlled process (ξ, δ) is ergotic on $S(K, J)$ state space with distribution π for any strategy δ which results in probabilities $p_j(h, u)$, where h is the number of customers at the node. Derman's *strategy improvement procedure* should be used to determine the optimal strategy δ^*.

At first select some strategy δ and consider the next equations:

$$R_y^{\delta} + \upsilon(y) = r(y, \delta(y)) + \sum_{x \in S(K,J)} Q(x/y, \delta(y))\, \upsilon(x), \quad \text{where } y \in S(K, J) \qquad (8)$$

$$\sum_{n \in S(K,J)} \pi^{\delta}(x)\upsilon(x) = 0 \qquad (9)$$

$$R^{\delta} = \sum_{x \in S(K,J)} \pi^{\delta}(x) r(x) \qquad (10)$$

By solving unknown functionυ for $\{(\upsilon(x),\ R^{\delta}_{y}): x, y \in S(K, J)\}$and getting R^{δ}_{y} asymptotic average expected costs, the policy improvement algorithm would be the following:

Define U^{y}which is set of actions for each $y \in S(K, J)$for which

$\sum_{x \in S(K,J)} Q(x/y,\ u)R^{\delta}_{x} < R^{\delta}_{y},$ or, if no actions satisfy that inequality, the set

that would satisfy both $\sum_{x \in S(K,J)} Q(x/y,\ u)R^{\delta}_{x} = R^{\delta}_{y}$ and

$$r(y, u) + \sum_{x \in S(K,J)} Q(x/y,\ u)\upsilon^{\delta}(x) < r(y, \delta(y)) + \sum_{x \in S(K,J)} Q(x/y,\ \delta(y))\upsilon^{\delta}(x) = R^{\delta}_{y} + \upsilon^{\delta}(y)$$

After starting with some strategy δwe define some local strategy δ'. It takes some action $u \in U^{y}$in at least one state y for which $U^{y}$is not empty, otherwise the action that has been taken is the one dedicated by δ.

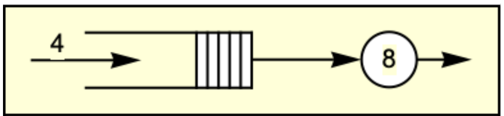
Derman theorem:

The strategy improvement procedure leads to an optimal strategy within a finite number of iterations. The optimal strategy is obtained if U y is empty for all y and then the actual policy fulfills: $\delta' =: \delta^{*}$ *is optimal.*

To sum everything up, we have some starting strategy $\delta$and use Gordon-Newell theory(5) for finding distribution π^{δ}, then find average costs R^{δ} (10). By using equations (8), (9) we get υ^{δ}functions. Then we determine the sets of actions U^{y} and define a new strategy δ'. This goes on until we get the optimal one.

4. Conclusion

With the help of Gordon-Newell Networks, it is possible to obtain crucial data about the given system (Figure 3). These metrics can show the problems with specific nodes and predict future data about the queues.



Basic Properties	
NetworkType	Open
NodeCount	2
SelectedNode	1
Performance Measures	
MeanSystemSize	3
MeanSystemTime	$\frac{1}{2}$
MeanQueueSize	$\frac{9}{4}$
MeanQueueTime	$\frac{3}{8}$

Figure 3: Example of queue properties in Gordon-Newell network

Visualizations can provide a new look at the system and help prevent future bottlenecks in the system. In Figure 4, there are seven customers in the GNN and four nodes (blue - 1, orange - 2, green - 3, red - 4). Based on the image, it is evident that node 1 is the stepping stone in the system and should be further improved to increase the whole system's efficiency.

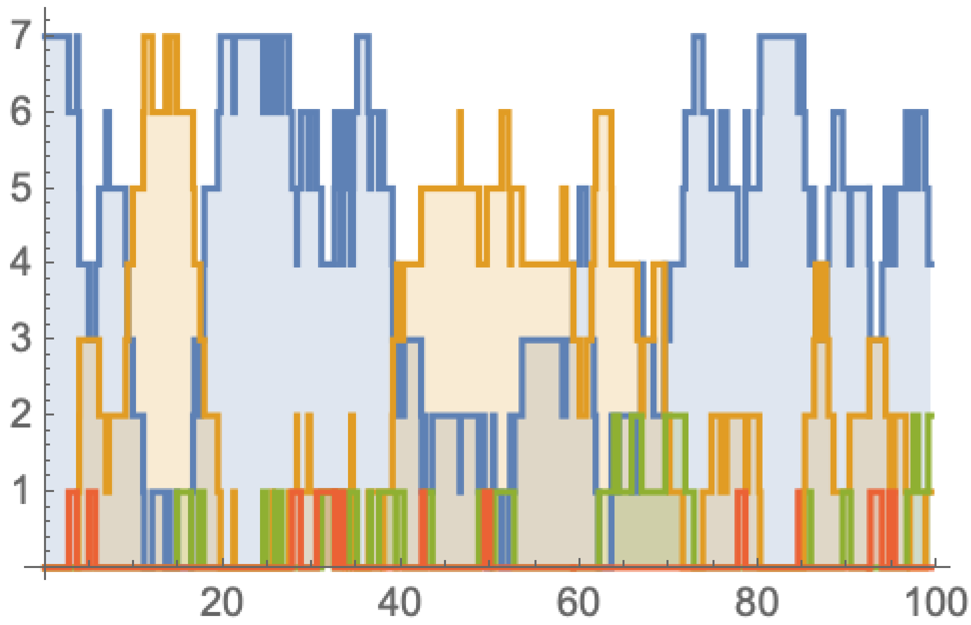


Figure 4: Gordon-Newell network’s queues capacities over time

In Figure 5, the GNN with nine workers and two nodes is displaying the system's load over some period of time.

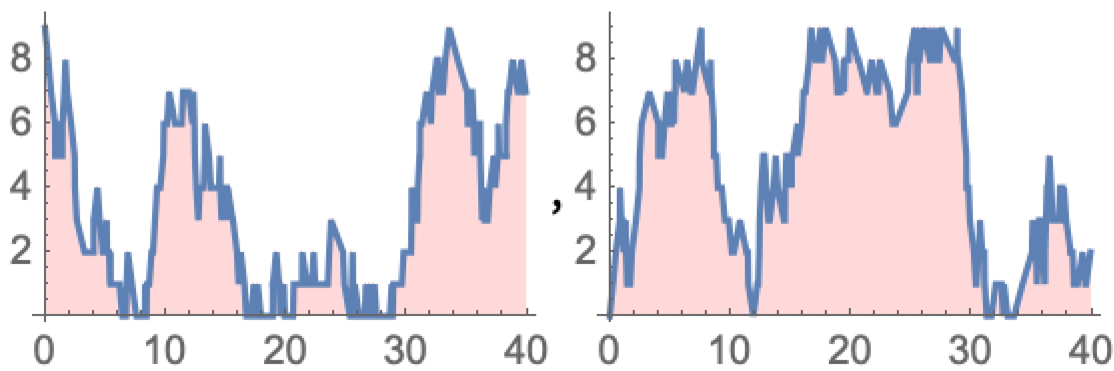


Figure 5: The load of 2 nodes GNN

Using Gordon-Newell theory for closed interconnected queues and applying it to the policy improvement algorithm benefits GNN. By obtaining optimal local strategy with local decision-makers, the efficiency of Gordon-Newell network increases and can solve the problems that the system had.

The algorithm and formulas presented in this paper are constantly being worked on. There are numerous publications that extend the Gordon-Newell theorem for more specific cases but were not touched upon in this paper. This means that there is still more work that can be done in the field of queueing theory that can provide a considerable benefit to the world.

Bibliography

1. Jeffrey P. Buzen Harvard University and Honeywell Information Systems “Computational Algorithms for Closed Queueing Networks with Exponential Servers” — 1973
2. 170. Jackson J. R. Jobshop-like queueing systems // Management Science. — 1963. — Vol. 10.
3. Gordon W. J., Newell G. F. Closed queueing networks with exponential servers // Operations Research. — 1967. — Vol. 15.
4. Kendall, D. G. Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain // *The Annals of Mathematical Statistics*. — 1953 — Vol. 24
5. Daduna, H. Passage Times for Overtake-Free Paths in Gordon-Newell Networks // *Advances in Applied Probability*. — 1982. — Vol. 14.
6. Gong, Q.; Lai, K. K.; Wang, S. Supply chain networks: Closed Jackson network models and properties // I*nternational Journal of Production Economics*. — 2008. — Vol. 2.
7. Erlang, Agner K. Solution of some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges — 1948. — Vol. 2.
8. Chornei, R. K.; Daduna, H.; Knopov, P. S. Controlled Markov fields with finite state space on graphs // *Stochastic Models* — 2005. — Vol. 21. — p. 847–874.
9. Chornei, R. K. Local control in networks // *Qualification scientific work in the form of a manuscript* — 2021. — p. 191-220