



Σχολή Θετικών Επιστημών & Τεχνολογίας
Μεταπτυχιακές Σπουδές στα Μαθηματικά

Διπλωματική Εργασία

Ticket Queues:
Μοντελοποίηση και ανάλυση επίδοσης

ΡΟΒΙΘΗ ΘΕΟΔΩΡΑ

Επιβλέπων Α΄: ΔΗΜΗΤΡΙΟΥ ΙΩΑΝΝΗΣ

Επιβλέπων Β΄: ΣΕΒΡΟΓΛΟΥ ΒΑΣΙΛΕΙΟΣ

Πάτρα, Σεπτέμβριος 2026

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία της φοιτήτριας Ροβίθη Θεοδώρας που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοιχτής πρόσβασης η συγγραφέας/δημιουργός εκχωρεί στο ΕΑΠ, μη αποκλειστική άδεια χρήσης του δικαιώματος αναπαραγωγής, προσαρμογής, δημόσιου δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσής τους διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος και για όλο το χρόνο διάρκειας των δικαιωμάτων πνευματικής ιδιοκτησίας. Η ανοιχτή πρόσβαση στο πλήρες κείμενο για μελέτη και ανάγνωση δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας της συγγραφέας/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, αποθήκευση, πώληση, εμπορική χρήση, μετάδοση, διανομή, έκδοση, εκτέλεση, «μεταφόρτωση» (downloading), «ανάρτηση» (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση της συγγραφέας/δημιουργού. Η Ροβίθη Θεοδώρα διατηρεί το σύνολο των ηθικών και περιουσιακών του δικαιωμάτων.

*Στον σύντροφό μου και στους γονείς μας,
για όλη τη στήριξη, την υπομονή και την πίστη τους.
Με την ευχή η επόμενη μέρα να μας βρει όλους καλά,
έτοιμους για το μεγάλο όνειρο...*

Ευχαριστίες

Θα ήθελα να εκφράσω τις πιο ειλικρινείς μου ευχαριστίες στον επιβλέποντα καθηγητή μου, κ. Ιωάννη Δημητρίου, για την αμέριστη στήριξη και τις πολύτιμες κατευθύνσεις που μου έδωσε. Η καθοδήγησή του από τα πρώτα κιόλας βήματα και η συμβολή του σε όλη τη διάρκεια αυτής της προσπάθειας υπήρξαν καθοριστικές για την ολοκλήρωση της παρούσας εργασίας.

Σας ευχαριστώ θερμά για την εμπιστοσύνη στο πρόσωπό μου.

Περίληψη

Η παρούσα διπλωματική εργασία εστιάζει στη μελέτη και ανάλυση των συστημάτων αναμονής με χρήση εισιτηρίων, τα οποία συναντώνται σε χώρους υψηλής κίνησης, όπως τράπεζες, δημόσιες υπηρεσίες και νοσοκομεία. Σε αντίθεση με τις φυσικές ουρές, όπου το πλήθος των ενεργών πελατών είναι ορατό, στα συστήματα εισιτηρίων φαίνεται μονάχα μία αριθμητική ένδειξη. Αυτό έχει ως αποτέλεσμα να δημιουργείται η εντύπωση ενός υπερβολικά φορτωμένου συστήματος, οδηγώντας τους πελάτες σε υπερεκτίμηση του χρόνου αναμονής. Κεντρικός στόχος της εργασίας είναι η διερεύνηση αυτής της συμπεριφοράς και η ανάπτυξη κατάλληλων μοντέλων για τον περιορισμό των πρόωρων, άσκοπων αποχωρήσεων και για τη βελτιστοποίηση της ποιότητας εξυπηρέτησης.

Το πρώτο κεφάλαιο της εργασίας θέτει το απαραίτητο θεωρητικό υπόβαθρο, παρουσιάζοντας κρίσιμες έννοιες στοχαστικών διαδικασιών (όπως η Μαρκοβιανή Ιδιότητα, η Στάσιμη Κατανομή και η Οριακή Συμπεριφορά) και αναλύοντας τις διαδικασίες Quasi Birth Death (QBD), ως γενίκευση των διαδικασιών Birth Death (BD), με δομή επιπέδου και φάσης. Παράλληλα, εξετάζονται ο απειροστός γεννήτορας, ο τοπικός χαρακτήρας των μεταβάσεων, η συνθήκη ευστάθειας και ο πίνακας ρυθμού, με τελικό στόχο την ολοκληρωμένη παρουσίαση της Πινακογεωμετρικής Μεθόδου. Το κεφάλαιο κλείνει με τρία αντιπροσωπευτικά παραδείγματα που αποτυπώνουν την εφαρμογή της μεθόδου.

Το δεύτερο κεφάλαιο ολοκληρώνει το θεωρητικό μέρος, καλύπτοντας τη Θεωρία Παιγνίων και τη Στρατηγική Συμπεριφορά στις Ουρές Αναμονής. Σε αυτό το τμήμα αναλύονται θεμελιώδεις έννοιες, όπως η ισορροπία Nash, η συνάρτηση βέλτιστης απόκρισης και οι στρατηγικές κατωφλιού, παρέχοντας τα απαραίτητα μαθηματικά εργαλεία για την κατανόηση της συμπεριφοράς των πελατών στα επόμενα στάδια.

Το τρίτο κεφάλαιο αποτελεί τον πυρήνα του κύριου μέρους της μελέτης, εστιάζοντας στη μοντελοποίηση της ουράς εισιτηρίων με πελάτες που αποχωρούν. Κατά τη διαδικασία αυτή, αναδύεται ένα σοβαρό υπολογιστικό πρόβλημα. Για την αποτελεσματική αντιμετώπισή του, εισάγεται το μοντέλο της τροποποιημένης ουράς, σύμφωνα με το οποίο τα εικονικά εισιτήρια αυτών που αποχώρησαν τοποθετούνται στο τέλος. Με τη νέα αυτή προσέγγιση, η πολυπλοκότητα μειώνεται από εκθετική σε τετραγωνική, ενώ παράλληλα διατηρείται η συμπεριφορά του συστήματος αναλλοίωτη. Η ακρίβεια και η αξιοπιστία αυτής της προσεγγιστικής μεθόδου τεκμηριώνονται από τα αποτελέσματα σχετικής βιβλιογραφίας, τα στοιχεία και τα συμπεράσματα της οποίας μελετήθηκαν σε βάθος και αξιοποιήθηκαν επιλεκτικά στην παρούσα εργασία. Επιπλέον, αναδεικνύεται το κρίσιμο παράδοξο όπου οι πελάτες σε ουρές με εισιτήρια αποχωρούν συχνότερα συγκριτικά με τις φυσικές ουρές. Για την αποφυγή αυτού, προτείνεται το βελτιωμένο μοντέλο Ticket-Plus, το οποίο αναγράφει στο εισιτήριο τον αναμενόμενο χρόνο αναμονής, για τη διόρθωση της υπερβολικής εκτίμησης του πελάτη.

Το τέταρτο και τελευταίο κεφάλαιο εισάγει την έννοια του διπλού κατωφλιού, εξετάζοντας πώς οι πελάτες ερμηνεύουν τα δεδομένα και προσαρμόζουν τις αποφάσεις τους. Μέσω της ανάλυσης της δομής του χώρου και βοηθητικών προτάσεων και λημμάτων, οδηγούμαστε στα κεντρικά θεωρήματα της εργασίας. Σύμφωνα με το πρώτο θεώρημα, όσο μεγαλύτερο είναι το συνολικό μήκος της ουράς,

τόσο πιο πιθανό είναι οι προηγούμενοι πελάτες να έχουν εγκαταλείψει, φτάνοντας στο συμπέρασμα ότι αν ο παρατηρούμενος αριθμός πελατών γίνει τεράστιος, με βεβαιότητα παραμένει ενεργός αποκλειστικά ο πελάτης που εξυπηρετείται τη δεδομένη στιγμή, ενώ όλοι οι υπόλοιποι έχουν αποχωρήσει. Στη συνέχεια, το δεύτερο και τελικό θεώρημα αποδεικνύει ότι οι απλές στρατηγικές μονού κατωφλιού δεν μπορούν να σταθούν ως στρατηγικές ισορροπίας Nash. Αντίθετα, η βέλτιστη αντίδραση για τον πελάτη, εάν όλοι οι άλλοι υιοθετήσουν την ίδια στρατηγική μονού κατωφλιού, είναι να ακολουθήσει στρατηγική διπλού κατωφλιού, επιλέγοντας τελικά να εισέλθει και στην επιπλέον περίπτωση που η ένδειξη του εισιτηρίου ξεπερνά μια δεύτερη πιο ανεβασμένη τιμή.

Η μελέτη ολοκληρώνεται αναφέροντας κατευθύνσεις στις οποίες μπορεί να επεκταθεί ή να γενικευθεί το βασικό μοντέλο. Ζητήματα όπως η επαναξιολόγηση της θέσης του πελάτη σε μεταγενέστερο χρόνο, οι πολλαπλοί εξυπηρετητές, οι γενικοί χρόνοι εξυπηρέτησης ή η καθυστερημένη ενημέρωση για τις αποχωρήσεις αποτελούν ενδεχόμενες προοπτικές για τέτοιες επεκτάσεις των συστημάτων αναμονής με εισιτήρια.

Λέξεις-κλειδιά:

- Ουρά εισιτηρίων
- Πινακογεωμετρική μέθοδος
- Ορθολογική συμπεριφορά
- Πελάτες που αποχωρούν
- Υπολογιστική πολυπλοκότητα
- Μοντέλο τροποποιημένης ουράς
- Μοντέλο Ticket-Plus
- Στρατηγικές κατωφλιού

Abstract

This thesis focuses on the study and analysis of ticket queueing systems, which are commonly found in high-traffic environments such as banks, public services and hospitals. In contrast to physical queues, where the actual queue size is directly visible, ticket systems display only a numerical indicator. This creates the impression of an overly congested system, leading customers to overestimate their waiting time. The main objective of this work is to investigate this behavior and develop appropriate models to prevent unnecessary abandonments and to optimize the quality of services.

The first chapter establishes the necessary theoretical background by presenting crucial concepts of stochastic processes (such as the Markovian property, stationary distribution and limiting behavior) and analyzing Quasi-Birth-Death (QBD) processes as a generalization of Birth-Death (BD) processes, featuring a level and phase structure. Furthermore, it examines the infinitesimal generator matrix, the local character of transitions, the stability condition and the rate matrix, with the ultimate goal of providing a comprehensive presentation of the Matrix-Geometric Method (MGM). The chapter concludes with three representative examples illustrating its application.

The second chapter completes the theoretical part by covering Game Theory and Strategic Behavior in Queues. This section analyzes fundamental concepts such as Nash equilibrium, the best-response function and the threshold strategies, providing the necessary mathematical tools for understanding customer behavior in subsequent stages.

The third chapter forms the core of the thesis, focusing on the modeling of ticket queues with balking customers. During this process, a serious computational problem arises. To address it effectively, a modified queue model is introduced, according to which the tickets of balking customers are placed at the end. With this new approach, the complexity is reduced from exponential to quadratic, while the system's behavior remains the same. The accuracy and reliability of this approximation method are supported by the literature, whose data and conclusions were carefully selected and used. Additionally, an important paradox is observed where customers in ticket queues balk more frequently compared to physical queues. To prevent this, the improved Ticket-Plus model is proposed, which prints the expected waiting time on the ticket to correct the customer's overestimation.

The fourth and final chapter introduces the concept of double thresholds, examining how customers interpret data and adjust their decisions. Through the analysis of the state space structure, auxiliary propositions and lemmas, we reach the central theorems of the thesis. According to the first theorem, the larger the total queue length, the more likely it is that preceding customers have abandoned the line, concluding that if the observed number of customers becomes extremely large, with certainty only the customer currently in service remains in the system, while all the others have departed. Then, the second and final theorem proves that simple single-threshold strategies do not constitute a Nash equilibrium. Instead, the best response for a customer, if all the others adopt the same single-threshold strategy, is to follow a double-threshold strategy, choosing to enter even when the ticket indication exceeds a second, higher threshold.

The study concludes by outlining directions in which the basic model can be extended or generalized. Issues such as customer re-evaluation of position at a later time, multiple servers, general service times or delayed notification of departures constitute potential prospects for such extensions of ticket queueing systems.

Keywords:

- Ticket queue
- Matrix-geometric method (MGM)
- Rational behavior
- Balking customers
- Computational complexity
- Modified queue model
- Ticket-Plus model
- Threshold strategies

Διευκρινιστικές σημειώσεις για τον ιδιαίτερο συμβολισμό και τη μορφοποίηση

- Στο πλαίσιο της παρούσας ανάλυσης, τα έντονα γράμματα (bold) χρησιμοποιούνται για τον προσδιορισμό των διανυσμάτων, ενώ οι πίνακες αναπαρίστανται με κεφαλαία γράμματα.
- Στους αριθμητικούς πίνακες και στα διαγράμματα, οι δεκαδικοί αριθμοί διαχωρίζονται με τελεία αντί υποδιαστολής.

Περιεχόμενα

Περίληψη

Abstract

Διευκρινιστικές σημειώσεις για τον ιδιαίτερο συμβολισμό και τη μορφοποίηση

Περιεχόμενα

Μέρος I: Θεωρητικό Υπόβαθρο – Βασική Μεθοδολογία

Κεφάλαιο 1^ο: Ανάλυση Διαδικασιών Quasi Birth Death (QBD)

1.1 Εισαγωγή στις Στοχαστικές Διαδικασίες	1
1.1.1 Ορισμοί και Θεμελιώδεις Έννοιες των Στοχαστικών Διαδικασιών	1
1.1.2 Η Μαρκοβιανή Ιδιότητα	2
1.1.3 Χρονική Ομοιογένεια – Στοχαστικός Πίνακας Πιθανοτήτων Μετάβασης	3
1.1.4 Εξισώσεις Chapman – Kolmogorov	5
1.1.5 Στάσιμη Κατανομή και Οριακή Συμπεριφορά ΜΑΣΧ	7
1.1.6 Επικοινωνία, Επαναληπτικότητα και Παροδικότητα	9
1.2 Ορισμός και Αναγνώριση της Διαδικασίας Quasi Birth Death	11
1.2.1 Γενική Έννοια των Διαδικασιών Γεννήσεων – Θανάτων (BD)	11
1.2.2 Αναγκαιότητα των QBDs	13
1.2.3 Ορισμός των QBDs και Δομή του Χώρου Καταστάσεων	14
1.3 Συνθήκη Ευστάθειας και Πινακογεωμετρική Μέθοδος Επίλυσης	16
1.3.1 Συνθήκη Ευστάθειας	17
1.3.2 Πινακογεωμετρική Μορφή και Βασικές Ιδιότητες του Rate Matrix R	18
1.3.3 Μέθοδοι Υπολογισμού του R	20
1.3.4 Πλήρης Μεθοδολογία Επίλυσης της QBD μέσω της MGM	22
1.4 Κλασικά Παραδείγματα	24
1.4.1 M/M/1 με Λειτουργία Set-Up/Turn-Off	24
1.4.2 M/M/1 με Διακοπές Λειτουργίας και Επιδιόρθωση (Breakdown/Repair)	28
1.4.3 M/Cox/1	30

Κεφάλαιο 2^ο: Στοιχεία Θεωρίας Παιγνίων και Στρατηγική Συμπεριφορά στις Ουρές Αναμονής

2.1 Εισαγωγή στην Θεωρία Παιγνίων	34
2.2 Στρατηγικές και Ισορροπία Nash στα Συστήματα Ουρών	36
2.3 Μονοτονία των Συναρτήσεων Βέλτιστης Απόκρισης και Μορφές Ισορροπίας	38
2.4 Στρατηγικές Κατωφλιού	40

Μέρος II: Κύριο Τμήμα – Ερευνητική Εφαρμογή και Αποτελέσματα

Κεφάλαιο 3^ο: Μοντελοποίηση και Ανάλυση Ουράς Εισιτηρίων με Πελάτες που Αποχωρούν

3.1 Περιγραφή και Λειτουργία του Συστήματος Εισιτηρίων – Ανάλυση των Ρυθμών Μετάβασης	44
3.2 Ομαδοποίηση Καταστάσεων και Υπολογισμός της Στάσιμης Κατανομής τους	48
3.3 Προσδιορισμός των Στάσιμων Πιθανοτήτων Εντός των Υπερ-Καταστάσεων	55
3.4 Βασικά Μέτρα Απόδοσης της Ουράς Εισιτηρίων	61
3.5 Αντιμετώπιση της Υπολογιστικής Πολυπλοκότητας: Το Μοντέλο της Τροποποιημένης Ουράς	65
3.6 Μελέτη Απόδοσης και Χαρακτηριστικά της Συμπεριφοράς του Συστήματος	73
3.6.1 Συγκριτική Ανάλυση Ακρίβειας και Δείκτες Απόκλισης	73
3.6.2 Ποιοτική Αξιολόγηση και Ιδιότητες των Βασικών Μέτρων Απόδοσης	77
3.7 Βελτιστοποίηση της Απόδοσης των Υπηρεσιών του Συστήματος Εισιτηρίων	80
3.7.1 Στοχαστική και Αριθμητική Συγκριτική Μελέτη: Ουρές με Εισιτήρια (Απλή και Τροποποιημένη) έναντι της Φυσικής Ουράς M/M/1/K	81
3.7.2 Η Ticket-Plus Queue και η Συγκριτική Ανάλυσή της με την M/M/1/K	88

Κεφάλαιο 4^ο: Ανάλυση Συμπεριφοράς Πελατών και Στρατηγικές Κατωφλιού σε Συστήματα Εξυπηρέτησης με Εισιτήρια

4.1 Εισαγωγή στην Στρατηγική Συμπεριφορά των Πελατών	93
4.2 Μοντέλο και Αναπαράσταση του Χώρου Καταστάσεων	94
4.3 Ανάλυση Φάσεων, Επιτρεπτές Μεταβάσεις και Προσδιορισμός της Στάσιμης Κατανομής	97
4.4 Δεσμευμένη Κατανομή των Πραγματικών Πελατών και Ανάδειξη της Αναγκαιότητας του Διπλού Κατωφλιού	103
4.5 Ανασκόπηση, Συμπεράσματα και Προοπτικές Γενίκευσης	110

Βιβλιογραφία/Παραπομπές	114
-------------------------------	-----

Παραρτήματα Α: Κώδικας Αριθμητικής Εφαρμογής της MGM στο M/Cox/1	115
--	-----

Κατάλογος και Επεξήγηση Όρων	117
------------------------------------	-----

Κεφάλαιο 1^ο: Ανάλυση Διαδικασιών Quasi Birth Death (QBD)

Το παρόν κεφάλαιο θεμελιώνει τις μαθηματικές βάσεις και τα αναλυτικά εργαλεία που καθίστανται απαραίτητα για τη Μοντελοποίηση και Ανάλυση Επίδοσης των Ticket Queues. Η ανάπτυξη της μεθοδολογίας εκκινεί με την παράθεση των θεμελιωδών εννοιών των Στοχαστικών Διαδικασιών, οι οποίες συνιστούν το γενικό πλαίσιο για την περιγραφή της δυναμικής των συστημάτων εξυπηρέτησης. Στη συνέχεια, παρουσιάζονται και αναλύονται οι Quasi Birth Death (QBD) Διαδικασίες, μια ειδική κλάση Μαρκοβιανών Αλυσίδων, η οποία επιτρέπει τη συστηματική μελέτη ουρών αναμονής με δομή επιπέδων. Η κατανόηση των QBD και της Πινακογεωμετρικής Μεθόδου (MGM) είναι κρίσιμη, δεδομένου ότι τα μοντέλα Ticket Queues αναλύονται μέσω αυτής της δομής. Το κεφάλαιο ολοκληρώνεται με κλασικά παραδείγματα εφαρμογής της MGM, τα οποία αποσκοπούν στην εξοικείωση με τη μέθοδο, προκειμένου αυτή να αξιοποιηθεί καταλλήλως στην ανάλυση των μέτρων επίδοσης της παρούσας διπλωματικής εργασίας.

1.1 Εισαγωγή στις Στοχαστικές Διαδικασίες

Το υλικό που παρουσιάζεται στην εν λόγω ενότητα μπορεί να βρεθεί στα ακόλουθα συγγράμματα: *Στοχαστικές Διαδικασίες* (Μ. Λουλάκης), *Introduction to Probability Models* (Sheldon M. Ross), *Performance Modeling and Design of Computer System: Queueing Theory in Action* (M. Harchol-Balter) και *Introduction to Modeling and Analysis of Stochastic Systems* (V. Kulkarni).

1.1.1 Ορισμοί και Θεμελιώδεις Έννοιες των Στοχαστικών Διαδικασιών

Η μοντελοποίηση της δυναμικής συστημάτων, υπό συνθήκες αβεβαιότητας και τυχαίας εξέλιξης, αποτελεί τον πυρήνα της θεωρίας των Στοχαστικών Διαδικασιών. Μία Στοχαστική Διαδικασία $\{X(t), t \in T\}$ ορίζεται ως μία οικογένεια τυχαίων μεταβλητών, όπου $X(t)$ αντιπροσωπεύει την κατάσταση του συστήματος, τη χρονική στιγμή t , ορισμένη σε ένα χώρο καταστάσεων.

Μια Στοχαστική Διαδικασία χαρακτηρίζεται από τα εξής δύο βασικά στοιχεία:

- **Χώρος Καταστάσεων (E):** Το σύνολο αυτό αποτελείται από τις πιθανές τιμές που μπορεί να λάβει η τυχαία μεταβλητή $X(t)$. Στο πλαίσιο της μοντελοποίησης συστημάτων, όπως τα συστήματα εξυπηρέτησης, το σύνολο E συχνά θεωρείται διακριτό και αριθμήσιμο, άρα $E = \{0, 1, 2, 3, \dots\}$.
- **Παραμετρικός Χώρος (T):** Το σύνολο αυτό εκφράζει τις χρονικές στιγμές στις οποίες εξετάζεται η στοχαστική διαδικασία. Στην περίπτωση των Μαρκοβιανών Διαδικασιών, οι οποίες θα εξεταστούν λεπτομερώς στην πορεία της παρούσας υποενότητας, το T μπορεί να είναι είτε διακριτό $T = \{0, 1, 2, 3, \dots\}$, είτε συνεχές $T = [0, +\infty)$.

Η γνώση των κατανομών πεπερασμένων διαστάσεων, δηλαδή των από κοινού κατανομών πιθανότητας των τυχαίων μεταβλητών $X(t_1), X(t_2), \dots, X(t_n)$, που αντιστοιχούν στις χρονικές στιγμές $t_1 < t_2 < \dots < t_n$ στο σύνολο του χρόνου T , είναι καίριας σημασίας για την πλήρη μελέτη των χαρακτηριστικών ιδιοτήτων μιας στοχαστικής διαδικασίας.

Σε πολλές εφαρμογές, το κύριο ερευνητικό αντικείμενο είναι η ασυμπτωτική συμπεριφορά της στοχαστικής διαδικασίας, δηλαδή η συμπεριφορά της καθώς ο χρόνος τείνει στο άπειρο.

Κεντρικό ρόλο στη μελέτη αυτή κατέχει η έννοια της στάσιμης κατανομής, η οποία ορίζεται ως η κατανομή πιθανότητας π , που παραμένει αναλλοίωτη με την πάροδο του χρόνου. Εάν η διαδικασία ξεκινήσει με αυτή την κατανομή ως αρχική συνθήκη, τότε η κατανομή της παραμένει σταθερή για κάθε χρονική στιγμή $t > 0$.

Παράλληλα, η οριακή κατανομή περιγράφει την τάση της διαδικασίας να σταθεροποιείται μακροπρόθεσμα σε ένα συγκεκριμένο διάνυσμα πιθανοτήτων, ανεξάρτητα από την αρχική της κατάσταση. Εφόσον η οριακή κατανομή υπάρχει, η πιθανότητα να βρεθεί το σύστημα σε μία συγκεκριμένη κατάσταση καθίσταται σταθερή σε βάθος χρόνου.

Συνεπώς, εάν υπάρχει ένα διάνυσμα πιθανοτήτων $\pi = (\pi_x)_{x \in E}$, τέτοιο ώστε $\sum_{x \in E} \pi_x = 1$ και ισχύει :

$$\lim_{t \rightarrow \infty} P(X(t) = x) = \pi_x, \quad \forall x \in E,$$

τότε το διάνυσμα π ονομάζεται οριακή κατανομή της διαδικασίας.

Η σχέση μεταξύ των δύο αυτών εννοιών είναι καθοριστική, καθώς όταν η αλυσίδα είναι επαναληπτική¹, η οριακή κατανομή υπάρχει, είναι μοναδική και ταυτίζεται με τη στάσιμη κατανομή. Η σύγκλιση αυτή εξασφαλίζει ότι η διαδικασία επιδεικνύει συμπεριφορά ισορροπίας, όπου οι μακροπρόθεσμες ιδιότητες του συστήματος παραμένουν αναλλοίωτες και ανεξάρτητες από τις αρχικές συνθήκες.

1.1.2 Η Μαρκοβιανή Ιδιότητα

Όταν επιδιώκεται η μοντελοποίηση της εξέλιξης πραγματικών συστημάτων, υιοθετείται συχνά η παραδοχή ότι η μελλοντική συμπεριφορά του συστήματος εξαρτάται αποκλειστικά από την τρέχουσα κατάστασή του και είναι ανεξάρτητη από την ιστορία των καταστάσεων που προηγήθηκαν. Αυτή η θεμελιώδης αρχή, γνωστή ως Μαρκοβιανή Ιδιότητα (Markov Property) ή Ιδιότητα της Αμνησίας, επιτρέπει μια σημαντική αναλυτική απλοποίηση, καθώς η πρόβλεψη της μελλοντικής εξέλιξης καθίσταται εφικτή μόνο βάσει της πιο πρόσφατης διαμόρφωσης του συστήματος. Τα στοχαστικά συστήματα που πληρούν αυτή την προϋπόθεση χαρακτηρίζονται ως Μαρκοβιανά και αποτελούν κεντρική κλάση μοντέλων στην θεωρία πιθανοτήτων και ουρών αναμονής.

¹ Για τον πλήρη ορισμό της έννοιας, βλέπε στην υποπαράγραφο 1.1.6.

Ορισμός 1.1: (ΜΑΔΧ) Μια στοχαστική διαδικασία $\{X_n\}_{n \in N_0}$ λέγεται Μαρκοβιανή Αλυσίδα Διακριτού Χρόνου (Discrete Time Markov Chain (DTMC)), αν για κάθε $n \in N$, η δεσμευμένη κατανομή της X_{n+1} δοθέντων των $(X_0, \dots, X_n)$, ταυτίζεται με την δεσμευμένη κατανομή της X_{n+1} με μόνη δοθείσα την X_n . Επομένως, η $\{X_n\}_{n \in N_0}$, με τιμές σε ένα αριθμήσιμο χώρο καταστάσεων E , είναι Μαρκοβιανή Αλυσίδα, αν για κάθε $n \in N$ και κάθε $i_0, \dots, i_{n-1}, i, j \in E$, ισχύει:

$$P(X_{n+1} = j | X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i).$$

Οι $p_{i,j}^{(n,n+1)} = P(X_{n+1} = j | X_n = i)$ ονομάζονται πιθανότητες μετάβασης ενός βήματος την χρονική στιγμή n .

Ορισμός 1.2: (ΜΑΣΧ) Μια στοχαστική διαδικασία $\{X(t); t \in T\}$ λέγεται Μαρκοβιανή Αλυσίδα Συνεχούς Χρόνου (Continuous Time Markov Chain (CTMC)), αν για κάθε $t \in R^+$, η δεσμευμένη κατανομή της $X(t+s)$ δοθέντων όλων των προηγούμενων $X(u), 0 \leq u < s$, ταυτίζεται με την δεσμευμένη κατανομή της $X(t+s)$ με μόνη δοθείσα την $X(s)$. Επομένως, η $\{X(t); t \in T\}$, με τιμές σε ένα διακριτό χώρο καταστάσεων E , είναι Μαρκοβιανή Αλυσίδα, αν για κάθε $s, t > 0$ και κάθε $i, j \in E$, ισχύει:

$$P(X(t+s) = j | X(u), 0 \leq u < s, X(s) = i) = P(X(t+s) = j | X(s) = i).$$

Οι $p_{i,j}(s, t+s) = P(X(t+s) = j | X(s) = i)$ ονομάζονται πιθανότητες μετάβασης από την i στην j σε χρόνο t , την στιγμή s .

1.1.3 Χρονική Ομοιογένεια – Στοχαστικός Πίνακας Πιθανοτήτων Μετάβασης

Στην πλειονότητα των εφαρμογών των μαρκοβιανών αλυσίδων που θα μας απασχολήσουν, η λειτουργία που περιγράφει την εξέλιξη της αλυσίδας δεν έχει χρονική εξάρτηση. Ως εκ τούτου, η πιθανότητα μετάβασης μεταξύ δύο καταστάσεων παραμένει σταθερή σε κάθε χρονικό βήμα, χαρακτηρίζοντας την αλυσίδα ως Χρονικά Ομοιογενή (Time Homogeneous). Στο πλαίσιο της παρούσας εργασίας, η χρονική ομοιογένεια των Μαρκοβιανών Αλυσίδων θα θεωρείται αυτονόητη παραδοχή και δεν θα απαιτείται εκ νέου διαπίστωση στις ακόλουθες ενότητες.

Ορισμός 1.3: (Χρονική Ομοιογένεια) Η ΜΑΔΧ $\{X_n\}_{n \in N_0}$, με τιμές σε ένα αριθμήσιμο χώρο καταστάσεων E , λέγεται χρονικά ομοιογενής, αν για κάθε $n \geq 0$ και κάθε $i, j \in E$, ισχύει:

$$p_{i,j}^{(n,n+1)} = P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i) = p(i, j) = p_{i,j}.$$

Ορισμός 1.4: (Χρονική Ομοιογένεια) Η ΜΑΣΧ $\{X(t); t \in T\}$, με τιμές σε ένα διακριτό χώρο καταστάσεων E , λέγεται χρονικά ομοιογενής, αν για κάθε $s, t > 0$ και κάθε $i, j \in E$, ισχύει:

$$p_{i,j}(s, t+s) = P(X(t+s) = j | X(s) = i) = P(X(t) = j | X(0) = i) = p_{i,j}(t).$$

Δεδομένου ότι το δομικό στοιχείο μελέτης της παρούσας διπλωματικής απαιτεί την αποκλειστική αξιοποίηση Μαρκοβιανών Αλυσίδων Συνεχούς Χρόνου (ΜΑΣΧ), η ανάλυση που ακολουθεί θα επικεντρωθεί σε αυτές. Όλα τα απαραίτητα θεωρητικά στοιχεία θα εξειδικευτούν, διαμορφώνοντας ένα ολοκληρωμένο πλαίσιο βασισμένο σε φαινόμενα που εξελίσσονται σε συνεχή χρόνο.

Οι πιθανότητες μετάβασης ορίζουν τον εξής πίνακα μετάβασης:

$$P(t) = (p_{i,j}(t))_{i,j \in E} = \begin{bmatrix} p_{0,0}(t) & p_{0,1}(t) & \cdots & p_{0,j}(t) & \cdots \\ p_{1,0}(t) & p_{1,1}(t) & \cdots & p_{1,j}(t) & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ p_{i,0}(t) & p_{i,1}(t) & & p_{i,j}(t) & \\ \vdots & \vdots & & \vdots & \ddots \end{bmatrix},$$

όταν $|E| = |N|$, δηλαδή ισοπληθικό με το σύνολο των φυσικών αριθμών,

$$\text{ενώ έχουμε: } P(t) = (p_{i,j}(t))_{i,j \in E} = \begin{bmatrix} p_{0,0}(t) & p_{0,1}(t) & \cdots & p_{0,j}(t) & \cdots & p_{0,N}(t) \\ p_{1,0}(t) & p_{1,1}(t) & \cdots & p_{1,j}(t) & \cdots & p_{1,N}(t) \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ p_{i,0}(t) & p_{i,1}(t) & & p_{i,j}(t) & \cdots & p_{i,N}(t) \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ p_{N,0}(t) & p_{N,1}(t) & \cdots & p_{N,j}(t) & \cdots & p_{N,N}(t) \end{bmatrix},$$

όταν $|E| = N < \infty$.

Ο πίνακας πιθανοτήτων μετάβασης $P(t)$ εξασφαλίζει τον στοχαστικό του χαρακτήρα υπό τις προϋποθέσεις:

- $p_{i,j}(t) \geq 0, \forall i, j \in E$,
- $\sum_{j=0}^{\infty} p_{i,j}(t) = 1, \forall i \in E, t \geq 0$.

Αυτό σημαίνει ότι αν ξεκινήσουμε από οποιαδήποτε κατάσταση, η πιθανότητα μετάβασης σε μία από όλες τις δυνατές καταστάσεις είναι 1, δηλαδή, σε πεπερασμένο χρόνο $[0, t]$, η μετάβαση από την κατάσταση i σε κάποια κατάσταση j είναι βέβαιη.

Η προϋπόθεση αυτή αποτελεί συνθήκη κανονικοποίησης (normalization condition), επιβάλλοντας τα στοιχεία κάθε γραμμής του πίνακα μεταβάσεων να αθροίζουν στη μονάδα.

Η διαδικασία $\{X(t); t \in T\}$ είναι πλήρως ορισμένη αν είναι γνωστά: ο πίνακας $P(t)$ και η αρχική κατανομή $p_i(0) = P(X(0) = i), i \in E$ της αλυσίδας, καθορίζοντας πλήρως τη στοχαστική της συμπεριφορά.

1.1.4 Εξισώσεις Chapman – Kolmogorov

Για την πορεία της ανάλυσης, η μεθοδολογία απαιτεί την εισαγωγή των ρυθμών μετάβασης, οι οποίοι περιγράφουν την στιγμιαία ένταση των μεταβάσεων μεταξύ καταστάσεων, θεμελιώνοντας έτσι την μαθηματική βάση για τις Εξισώσεις Chapman – Kolmogorov.

Θεώρημα 1.1: Για κάθε Μαρκοβιανή Αλυσίδα Συνεχούς Χρόνου αποδεικνύεται ότι οι πιθανότητες μετάβασης $p_{i,j}(t)$, $p_{i,i}(t)$, $i, j \in E$ ικανοποιούν τα παρακάτω όρια, καθώς $t \rightarrow 0$,

$$\lim_{t \rightarrow 0} \frac{p_{i,j}(t)}{t} = q_{i,j}, i \neq j, \quad \lim_{t \rightarrow 0} \frac{1 - p_{i,i}(t)}{t} = v_i.$$

Όμως, στα παραπάνω όρια θεωρώντας $p_{i,j}(0) = 0$, $p_{i,i}(0) = 1$, προκύπτει:

$$\left. \frac{d}{dt} p_{i,j}(t) \right|_{t=0} = q_{i,j}, i \neq j \text{ και } -\left. \frac{d}{dt} p_{i,i}(t) \right|_{t=0} = v_i.$$

Με χρήση αναπτύγματος Taylor, γύρω από το 0, έχουμε:

- Για $i \neq j$: $p_{i,j}(t) = p_{i,j}(0) + p'_{i,j}(0) \cdot t + o(t)$,
- Για $i = j$: $p_{i,i}(t) = p_{i,i}(0) + p'_{i,i}(0) \cdot t + o(t)$.

Αντικαθιστώντας όπου t το 0 και χρησιμοποιώντας ότι $p_{i,j}(0) = 0$, $p'_{i,j}(t) = q_{i,j}$ και $p_{i,i}(0) = 1$, $p'_{i,i}(t) = -v_{i,i}$, αντίστοιχα σε κάθε σχέση, λαμβάνουμε:

- Για $i \neq j$: $p_{i,j}(t) \approx q_{i,j} \cdot t$,
- Για $i = j$: $p_{i,i}(t) \approx 1 - v_{i,i} \cdot t$.

Συνεπώς, θεωρώντας

- $q_{i,j}$ τον ρυθμό μετάβασης από την κατάσταση i στην κατάσταση j , με $i \neq j$ και
- $v_{i,i}$ τον ρυθμό εξόδου από την κατάσταση i ,

τότε από το Θεώρημα Ολικής Πιθανότητας $p_{i,i}(t) + \sum_{i \neq j} p_{i,j}(t) = 1$, προκύπτει η θεμελιώδης σχέση που συνδέει τον ρυθμό εξόδου με τους επιμέρους ρυθμούς μετάβασης, ως εξής:

$$v_i = \sum_{i \neq j} q_{i,j}.$$

Αυτοί οι ρυθμοί συγκροτούν τον πίνακα, του οποίου τα στοιχεία ορίζονται ως:

$$Q_{i,j} = \begin{cases} q_{i,j}, & \text{αν } i \neq j \\ -v_{i,i}, & \text{αν } i = j \end{cases}.$$

Ο πίνακας $Q_{i,j} = Q$, που καθορίζει πλήρως την συμπεριφορά της αλυσίδας, καλείται *Απειροστός Γεννήτορας (Infinitesimal Generator)* της ΜΑΣΧ ή, εναλλακτικά, *Πίνακας Ρυθμών Μετάβασης* και διαθέτει την μορφή:

$$Q_{i,j} = \begin{bmatrix} -\nu_0 & q_{0,1} & q_{0,2} & \cdots \\ q_{1,0} & -\nu_1 & q_{1,2} & \cdots \\ q_{2,0} & q_{2,1} & -\nu_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Υπό την απαραίτητη προϋπόθεση ικανοποίησης των εξής σχέσεων :

- $q_{i,j} \geq 0, \forall i, j \in E,$
- $-\nu_i + \sum_{i \neq j} q_{i,j} = 0,$

επιβάλλεται ότι τα στοιχεία κάθε γραμμής του απειροστού γεννήτορα αθροίζουν στο μηδέν.

Ο Απειροστός Γεννήτορας Q μας επιτρέπει την άμεση παραγωγή των Προδρομικών Εξισώσεων Kolmogorov (Kolmogorov's Forward Equations), οι οποίες προκύπτουν από τις Εξισώσεις Chapman – Kolmogorov.

Θεώρημα 1.2: Για κάθε $s, t \geq 0$, ισχύει:

$$p_{i,j}(t+s) = \sum_{k \in E} p_{i,k}(t) p_{k,j}(s).$$

Αυτή η σχέση απορρέει και πάλι από το Θεώρημα Ολικής Πιθανότητας, υπό την προϋπόθεση ότι η διαδικασία είναι μαρκοβιανή. Συγκεκριμένα, αντιπροσωπεύει τον συλλογισμό ότι για να διανύσουμε την απόσταση από την κατάσταση i στην κατάσταση j , σε συνολικό χρόνο $t+s$, είναι αναγκαίο να μεταβούμε πρώτα από την i σε κάθε πιθανή ενδιάμεση κατάσταση k , του συνόλου καταστάσεων E , σε χρόνο t , και στη συνέχεια από εκεί να διανύσουμε την υπόλοιπη απόσταση μέχρι και την κατάσταση j , στον υπολειπόμενο χρόνο s .

Με αφετηρία τον παραπάνω τύπο του Θεωρήματος 1.2, για το χρονικό διάστημα $t+h$, έχουμε:

$$p_{i,j}(t+h) = \sum_{k \in E} p_{i,k}(t) p_{k,j}(h) = p_{i,j}(t) p_{j,j}(h) + \sum_{k \neq j} p_{i,k}(t) p_{k,j}(h). \quad (I)$$

Έπειτα, αξιοποιώντας την σειρά Taylor πρώτης τάξης στις πιθανότητες μετάβασης, για πολύ μικρό $h \rightarrow 0$, προκύπτουν:

- Για $k \neq j$: $p_{k,j}(t) = q_{k,j} \cdot h + o(h)$, μετάβαση στο j ,
- Για $k = j$: $p_{j,j}(t) = 1 - \nu_j \cdot h + o(h)$, παραμονή στο j .

Αντικαθιστώντας τις $p_{j,j}(t), p_{k,j}(t)$ στην (I), έχουμε:

$$p_{i,j}(t+h) = p_{i,j}(t) [1 - \nu_j \cdot h + o(h)] + \sum_{k \neq j} p_{i,k}(t) [q_{k,j} \cdot h + o(h)].$$

Με διαδοχικές πράξεις απλοποίησης, αναδιάταξη των όρων και φέρνοντας τον όρο $p_{i,j}(t)$ στο αριστερό μέλος, έχουμε:

$$p_{i,j}(t+h) - p_{i,j}(t) = \sum_{k \neq j} p_{i,k}(t)q_{k,j}h - p_{i,j}(t)v_jh + O(h).$$

Εν συνεχεία, διαιρούμε και τα δύο μέλη με h και λαμβάνουμε το όριο καθώς $h \rightarrow 0$ και έτσι καταλήγουμε σε:

$$\lim_{h \rightarrow 0} \frac{p_{i,j}(t+h) - p_{i,j}(t)}{h} = \lim_{h \rightarrow 0} \left[\sum_{k \neq j} p_{i,k}(t)q_{k,j} - p_{i,j}(t)v_j \right].$$

Παρατηρούμε ότι $\frac{O(h)}{h} \rightarrow 0$ όταν $h \rightarrow 0$, το οποίο οφείλεται στο γεγονός ότι $O(h)$ είναι αμελητέο και συγκλίνει γρηγορότερα στο 0 από ότι το h .

Το αριστερό μέλος της τελευταίας ισότητας είναι εξ' ορισμού η παράγωγος της $p_{i,j}(t)$ ως προς t , ενώ το δεξί μέλος επιδέχεται απλοποίηση:

$$\frac{dp_{i,j}(t)}{dt} = \sum_{k \neq j} p_{i,k}(t)q_{k,j} - p_{i,j}(t)v_j.$$

Πλέον, έχουμε καταλήξει στις Προδρομικές Εξισώσεις Kolmogorov (Kolmogorov's Forward Equations), που ανακεφαλαιώνονται στο ακόλουθο θεώρημα.

Θεώρημα 1.3: Για κάθε $i, j \in E$, οι πιθανότητες μετάβασης $p_{i,j}(t)$ μιας ΜΑΣΧ ικανοποιούν το σύστημα διαφορικών εξισώσεων διαφορών:

$$\frac{d}{dt} p_{i,j}(t) = \sum_{k \neq j} p_{i,k}(t)q_{k,j} - p_{i,j}(t)v_j. \quad (\text{Kolmogorov's Forward Equations})$$

Σε πίνακική μορφή οι Προδρομικές Εξισώσεις Kolmogorov γράφονται $\frac{d}{dt} P(t) = P(t)Q$, όπου $P(t)$ ο πίνακας μετάβασης και Q ο απειροστός γεννήτορας της αλυσίδας.

1.1.5 Στάσιμη Κατανομή και Οριακή Συμπεριφορά ΜΑΣΧ

Ενώ η στάσιμη κατανομή ορίζει τη συμπεριφορά ισορροπίας της διαδικασίας, ο προσδιορισμός αυτής στις ΜΑΣΧ απαιτεί την αξιοποίηση του Απειροστού Γεννήτορα Q και των Προδρομικών Διαφορικών Εξισώσεων Kolmogorov για την καταγραφή των εξισώσεων ισορροπίας.

Θεώρημα 1.4: Για μία αδιαχώριστη² και θετικά επαναληπτική³ ΜΑΣΧ, με ρυθμούς μεταβάσεων $v_i, q_{i,j}$, ισχύουν οι εξής προτάσεις για την οριακή συμπεριφορά της αλυσίδας:

- Τα όρια $\pi_j = \lim_{t \rightarrow \infty} p_{i,j}(t)$ υπάρχουν για όλα τα $(i, j) \in E$ και είναι ανεξάρτητα από την αρχική κατάσταση.
- Η οριακή κατανομή $\pi = (\pi_j)_{j \in E}$ ταυτίζεται με τη στάσιμη κατανομή και είναι η μοναδική μη αρνητική λύση του συστήματος των εξισώσεων ισορροπίας.

Στην ισορροπία, η κατανομή π πρέπει να παραμένει αναλλοίωτη, πράγμα που συνεπάγεται ότι ο ρυθμός μεταβολής της κατανομής $\frac{d\pi(t)}{dt}$ είναι μηδέν, καθώς $t \rightarrow \infty$.

Συνεπώς, στέλνοντας τον χρόνο στο άπειρο στις Kolmogorov's Forward Equations το αριστερό μέλος εξαφανίζεται, αφού $\lim_{t \rightarrow \infty} \frac{dp_{i,j}(t)}{dt} = 0$, ενώ συγχρόνως το δεξί μέλος οδηγεί στην θεμελιώδη σχέση:

$$\lim_{t \rightarrow \infty} \frac{dp_{i,j}(t)}{dt} = 0 \Rightarrow \sum_{k \neq j} q_{k,j} \pi_k - \pi_j v_j = 0 \Rightarrow \pi_j v_j = \sum_{k \neq j} q_{k,j} \pi_k,$$

όπου συνοδεύεται από την απαραίτητη συνθήκη κανονικοποίησης:

$$\sum_{i \in E} \pi_i = 1,$$

διαμορφώνουν το σύστημα του οποίου μοναδική λύση είναι η οριακή κατανομή.

Σαφώς, αφού το άθροισμα των στοιχείων κάθε γραμμής του πίνακα Q ισούται με το 0, τότε $-v_i + \sum_{k \neq j} q_{k,j} = 0 \Leftrightarrow v_i = \sum_{k \neq j} q_{k,j}$ και, έτσι καταλήγουμε στην σχέση $\pi_j \sum_{k \neq j} q_{k,j} = \sum_{k \neq j} q_{k,j} \pi_k$, που αποτελεί το σύστημα των εξισώσεων ισορροπίας.

Παρατίθενται οι εννοιολογικές ερμηνείες των όρων του συστήματος των εξισώσεων ολικής ισορροπίας, $\pi Q = 0$:

- Ο όρος $\sum_{k \neq j} \pi_k(t) q_{k,j}$ αποτελεί το συνολικό ρυθμό εισόδου στην κατάσταση j , δηλαδή το ρυθμό μετάβασης από κάθε κατάσταση k προς την j .
- Ο όρος $\pi_j \sum_{k \neq j} q_{k,j}$ αποτελεί το συνολικό ρυθμό εξόδου από την κατάσταση j .

Στην ειδική, ιδιάζουσα περίπτωση, που το σύστημα των εξισώσεων ισορροπίας και της εξίσωσης κανονικοποίησης δεν έχει λύση, τότε $\pi_j = 0, \forall j \in E$.

² Για τον πλήρη ορισμό της έννοιας, βλέπε στην υποπαράγραφο 1.1.6.

³ Βλέπε προηγούμενη υποσημείωση.

Η λύση του συστήματος των εξισώσεων $\pi Q = 0, \sum_{i \in E} \pi_i = 1$, παρέχει το π , το οποίο αποτελεί το ποσοστό του χρόνου που η αλυσίδα βρίσκεται στην κατάσταση j , μακροπρόθεσμα, καθιστώντας το ένα σημαντικό εργαλείο για τη μελέτη μέτρων απόδοσης ενός Μαρκοβιανού Συστήματος. Η $\pi = (\pi_j)_{j \in E}$ αναφέρεται ως οριακή κατανομή ή στάσιμη κατανομή ή κατανομή ισορροπίας της $\{X(t)\}$.

1.1.6 Επικοινωνία, Επαναληπτικότητα και Παροδικότητα

Για τη μελέτη της ευστάθειας και της ύπαρξης στάσιμης κατανομής σε μία ΜΑΔΧ, είναι απαραίτητος ο χαρακτηρισμός των καταστάσεων με βάση την ιδιότητα της επιστροφής.

Παρά το γεγονός ότι η QBD Διαδικασία, η οποία αποτελεί το αντικείμενο της μελέτης μας, είναι μια Μαρκοβιανή Αλυσίδα Συνεχούς Χρόνου (ΜΑΣΧ), οι ιδιότητες της Επαναληπτικότητας και της Παροδικότητας εφαρμόζονται και σε αυτήν. Αυτό συμβαίνει διότι η κατάταξη των καταστάσεων και η ύπαρξη στάσιμης κατανομής σε κάθε ΜΑΣΧ καθορίζονται από τις θεμελιώδεις ιδιότητες του Γεννήτορα Πίνακα Q και την Θεωρία Ισορροπίας του συστήματος, η οποία είναι άρρηκτα συνδεδεμένη με αυτή των ΜΑΔΧ. Ως εκ τούτου, δικαιολογείται έτσι η παράθεση των ακόλουθων ορισμών.

Επικοινωνία και Δομή Χώρου Καταστάσεων

Για τον χαρακτηρισμό των καταστάσεων, ορίζεται η σχέση της μεταξύ τους επικοινωνίας:

- Προσιτή: Η κατάσταση $j \in E$ λέγεται προσιτή από την $i \in E$ ($i \rightarrow j$) εάν υπάρχει $n \geq 0$ τέτοιο ώστε $p_{i,j}^{(n)} = P(X_n = j | X_0 = i) > 0$.
- Επικοινωνία: Οι καταστάσεις $i, j \in E$ επικοινωνούν ($i \leftrightarrow j$) εάν είναι αμφίδρομα προσιτές.
- Κλειστότητα: Μία κλάση επικοινωνίας $C \subseteq E$ λέγεται κλειστή αν $p_{i,j}^{(n)} = 0, \forall i \in C, j \notin C, n \geq 1$, δηλαδή όταν $i \not\rightarrow j$.
- Αδιαχωριστότητα: Μια ΜΑΔΧ είναι αδιαχώριστη εάν όλες οι καταστάσεις της επικοινωνούν μεταξύ τους, δηλαδή αν ο χώρος καταστάσεων E αποτελεί μία μόνο κλάση επικοινωνίας.

Παροδικότητα και Επαναληπτικότητα

Ο χαρακτηρισμός μιας κατάστασης i εξαρτάται από την πιθανότητα επιστροφής $f_{i,i}$ και το μέσο αριθμό επιστροφών $E(R_i)$.

Θεωρούμε $T_i^r = \min \{n > 0 : X_n = i\}$ τον χρόνο πρώτης επανόδου στην κατάσταση $i \in E$ και $f_{i,i}^* = \sum_{n=1}^{\infty} f_{i,i}^{(n)}$, με $f_{i,i}^{(n)} = P(T_i^r = n | X_0 = i)$ πιθανότητα μετάβασης πρώτης επανόδου με n βήματα.

Ορισμός 1.5: Η κατάσταση $i \in E$ λέγεται *Επαναληπτική* (Recurrent) εάν η επιστροφή στην i είναι βέβαιη, δηλαδή εάν ισχύει $f_{i,i}^* = P(T_i^r < \infty | X_0 = i) = 1$.

Πρόταση 1.1: Για κάθε κατάσταση $i \in E$, ισχύουν οι παρακάτω ισοδύναμες προτάσεις:

- i. Η κατάσταση i είναι επαναληπτική.
- ii. Το πλήθος των επιστροφών στην i απειρίζεται:

$$P(R_i = \infty | X_0 = i) = 1 \text{ και } P(R_i < \infty | X_0 = i) = 0,$$

δηλαδή η πιθανότητα το πλήθος των επιστροφών να είναι άπειρο είναι βέβαιη και ισοδύναμα η πιθανότητα το πλήθος των επιστροφών να είναι πεπερασμένο είναι μηδενική.

- iii. Ο μέσος αριθμός επιστροφών στην i απειρίζεται:

$$E(R_i | X_0 = i) = \sum_{n=1}^{\infty} p_{i,i}^{(n)} = \infty.$$

Ιδιότητα Κλάσης: Εάν η κατάσταση $i \in E$ είναι επαναληπτική, τότε και κάθε κατάσταση j , που επικοινωνεί με την i , είναι εξίσου επαναληπτική.

Ορισμός 1.6: Η κατάσταση $i \in E$ λέγεται *Παροδική* (Transient) αν δεν είναι επαναληπτική, που σημαίνει ότι $f_{i,i}^* = P(T_i^r < \infty | X_0 = i) < 1$, δηλαδή αν η επιστροφή στην i δεν είναι βέβαιη.

Πρόταση 1.2: Για κάθε κατάσταση $i \in E$, ισχύουν οι παρακάτω ισοδύναμες προτάσεις:

- i. Η κατάσταση i είναι παροδική.
- ii. Το πλήθος των επιστροφών στην i είναι πεπερασμένο:

$$P(R_i = \infty | X_0 = i) = 0 \text{ και } P(R_i < \infty | X_0 = i) = 1,$$

δηλαδή η πιθανότητα το πλήθος των επιστροφών να είναι άπειρο είναι μηδενική και ισοδύναμα η πιθανότητα το πλήθος των επιστροφών να είναι πεπερασμένο είναι βέβαιη.

- iii. Ο μέσος αριθμός επιστροφών στην i είναι πεπερασμένος:

$$E(R_i | X_0 = i) = \sum_{n=1}^{\infty} p_{i,i}^{(n)} < \infty.$$

Ιδιότητα Κλάσης: Εάν η κατάσταση $i \in E$ είναι παροδική, τότε και κάθε κατάσταση j , που επικοινωνεί με την i , είναι εξίσου παροδική.

Κατηγορίες Επαναληπτικότητας

Οι επαναληπτικές καταστάσεις διακρίνονται περαιτέρω με βάση τον μέσο χρόνο επιστροφής $\mu(i) = E(T_i | X_0 = i)$.

- **Θετικά Επαναληπτική** (Positive Recurrent): Μία επαναληπτική κατάσταση i είναι θετικά επαναληπτική αν ο μέσος χρόνος επιστροφής είναι πεπερασμένος, $\mu(i) < \infty$.

☑ **Συνθήκη Ευστάθειας:** Η Μαρκοβιανή Αλυσίδα είναι ευσταθής αν και μόνον αν όλες οι καταστάσεις της είναι θετικά επαναληπτικές. Αυτή η ιδιότητα αποτελεί την αναγκαία και ικανή συνθήκη για την ύπαρξη μη μηδενικής στάσιμης κατανομής π στο σύστημα.

- **Μηδενικά Επαναληπτική** (Null Recurrent): Μία επαναληπτική κατάσταση i είναι μηδενικά επαναληπτική αν ο μέσος χρόνος επιστροφής είναι άπειρος, $\mu(i) = \infty$.

Θεώρημα 1.5: Έστω μία ΜΑΔΧ $\{X_n; n \geq 0\}$ με πεπερασμένο χώρο καταστάσεων E . Τότε η ΜΑΔΧ έχει τουλάχιστον μία επαναληπτική κατάσταση.

Πρόταση 1.3: Κάθε επαναληπτική κατάσταση σε πεπερασμένο χώρο καταστάσεων είναι θετικά επαναληπτική.

Η παραπάνω πρόταση αποδεικνύεται βασιζόμενη στο ότι ο μέσος χρόνος επιστροφής στην εν λόγω περίπτωση είναι πάντα πεπερασμένος, $\mu(i) < \infty$.

Συνεπώς, η ιδιότητα της κλάσης επικοινωνίας – σύμφωνα με την οποία η επαναληπτικότητα διατηρείται σε όλες τις επικοινωνούσες καταστάσεις – σε συνδυασμό με το Θεώρημα 1.5, που εγγυάται την ύπαρξη επαναληπτικής κατάστασης σε πεπερασμένο χώρο, οδηγούν στο συμπέρασμα ότι κάθε πεπερασμένη κλειστή κλάση επικοινωνίας είναι επαναληπτική.

Κατόπιν, με βάση την Πρόταση 1.3, που δηλώνει ότι η επαναληπτικότητα συνεπάγεται θετική επαναληπτικότητα σε πεπερασμένο χώρο, προκύπτει το ακόλουθο θεώρημα:

Θεώρημα 1.6: Όλες οι καταστάσεις μιας πεπερασμένης κλειστής κλάσης επικοινωνίας είναι θετικά επαναληπτικές.

1.2 Ορισμός και Αναγνώριση της Διαδικασίας Quasi Birth Death

Η ενότητα αυτή χρησιμεύει ως γέφυρα, μεταβαίνοντας από τις απλές μονοδιάστατες Μαρκοβιανές Αλυσίδες (όπως οι Διαδικασίες Birth – Death (BD)) στις πολυδιάστατες, δομημένες διαδικασίες, που θα αποτελέσουν το αντικείμενο της μελέτης μας, τις Quasi Birth – Death Διαδικασίες (QBD), οι οποίες απαιτούν την χρήση πινάκων για την πλήρη περιγραφή των ρυθμών μετάβασης.

Το υλικό που παρουσιάζεται στην εν λόγω ενότητα προέρχεται από τα ακόλουθα σύγγραμματα: *Introduction to Matrix Analytic Methods in Stochastic Modeling* (V. Ramaswani and G. Latouche) και *Analysis of structured Markov processes* (I. Adan, J. van Leeuwen, J. Selen).

1.2.1 Γενική Έννοια των Διαδικασιών Γεννήσεων – Θανάτων (BD)

Οι Διαδικασίες Γεννήσεων – Θανάτων (BD) αποτελούν την απλούστερη και πιο θεμελιώδη κατηγορία Μαρκοβιανών Αλυσίδων σε Συνεχή Χρόνο και χρησιμοποιούνται για τη μοντελοποίηση του μεγέθους ενός πληθυσμού ή του αριθμού των πελατών σε ένα σύστημα.

Ο χώρος καταστάσεων μιας BD Διαδικασίας είναι το σύνολο $E = \{0, 1, 2, \dots\}$ και κάθε κατάσταση i αντιπροσωπεύει το μέγεθος της οντότητας που μοντελοποιείται (π.χ. πλήθος πελατών στην ουρά).

Είναι σαφές ότι σε αυτές της διαδικασίες ο Μαρκοβιανός Χαρακτήρας διατηρείται, εξασφαλίζοντας ότι η μελλοντική εξέλιξη του συστήματος εξαρτάται μόνο από την τρέχουσα κατάσταση και όχι από το παρελθόν.

Το καθοριστικό χαρακτηριστικό των BD Διαδικασιών είναι ο τοπικός περιορισμός στην συνδεσιμότητα των καταστάσεων, σύμφωνα με τον οποίο μεταβάσεις επιτρέπονται αυστηρά μεταξύ γειτονικών καταστάσεων, όπως φαίνεται και στο παρακάτω διάγραμμα μεταβάσεων.



Σχήμα 1.1: Διάγραμμα μεταβάσεων μιας Birth Death διαδικασίας

Έτσι, περιγραφικά έχουμε:

- Γέννηση (Birth): μετάβαση από n σε $n+1$ με ρυθμό λ_n .
- Θάνατος (Death): μετάβαση από n σε $n-1$ με ρυθμό μ_n .
- Δεν επιτρέπονται μεταβάσεις που παρακάμπτουν ενδιάμεσες καταστάσεις, δηλαδή δεν υπάρχουν άλματα από n σε $n \pm k, k \neq 1$. Αυτός ο τοπικός περιορισμός είναι η ουσία των BD.

Σχόλιο: Τα λ_n και μ_n μπορούν να έχουν τόσο μεταβλητή τιμή, προσαρμοζόμενη στον αριθμό n της αφετηριακής κατάστασης, όσο και σταθερή τιμή, ίδια για κάθε μετάβαση.

Ακολουθεί πίνακας συνδεδεμένων καταστάσεων και αντίστοιχων χρόνων των μεταξύ τους μεταβάσεων.

Κατάσταση	Επόμενη Κατάσταση	Χρόνος Μετάβασης σε Επόμενη Κατάσταση
$n \geq 0$	$n+1$	$T_{n,n+1} \sim \exp(\lambda_n)$
$n \geq 1$	$n-1$	$T_{n,n-1} \sim \exp(\mu_n)$

Πίνακας 1.1: Συνδεσιμότητα καταστάσεων και χρόνοι μετάβασης

Η δομή του χαρακτηριστικού τριδιαγώνιου πίνακα ρυθμών μετάβασης Q παρουσιάζεται ευθύς αμέσως, όπου για $n \geq 1$, τα μη μηδενικά στοιχεία του είναι: $q_{n,n+1} = \lambda_n, q_{n,n-1} = \mu_n$ και τα στοιχεία της κύριας διαγώνιου είναι $q_{n,n} = -(\lambda_n + \mu_n)$ (ενώ για την οριακή κατάσταση $n=0$ ισχύει $q_{0,0} = -\lambda_0$), ώστε να εξασφαλίζεται ότι κάθε γραμμή του πίνακα αθροίζει στο μηδέν:

$$Q = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \dots \\ 0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Η απλή αυτή τριδιαγώνια μορφή του πίνακα ρυθμών μετάβασης Q , αποτελεί την βάση, όπου στην περίπτωση των QBDs τα αριθμητικά στοιχεία του Q θα αντικατασταθούν με block πίνακες ρυθμών, δίνοντας μια γενίκευση του φαινομένου.

1.2.2 Αναγκαιότητα των QBDs

Η μετάβαση από τις κλασικές Διαδικασίες Γεννήσεων – Θανάτων (BD) στις Quasi Birth – Death (QBD) Διαδικασίες πηγάζει από την ανεπάρκεια των πρώτων να μοντελοποιήσουν συστήματα με σύνθετη εσωτερική δομή. Ενώ μία BD Διαδικασία μπορεί να περιγράψει αποτελεσματικά συστήματα όπου οι ρυθμοί μετάβασης εξαρτώνται αποκλειστικά από την κατάσταση της διαδικασίας, αποτυγχάνει όταν οι παράμετροι του συστήματος εξαρτώνται από έναν εξωτερικό παράγοντα.

Στα πραγματικά συστήματα, η συμπεριφορά γέννησης/θανάτου (π.χ. άφιξη πελάτη/ ολοκλήρωση εξυπηρέτησης) συχνά επηρεάζεται από την τρέχουσα λειτουργία του συστήματος. Για παράδειγμα, θεωρούμε ένα σύστημα εξυπηρέτησης (π.χ. μια δημόσια υπηρεσία) όπου η διαδικασία άφιξης των πελατών επηρεάζεται έντονα από την χρονική περίοδο της ημέρας. Η παραδοσιακή BD διαδικασία προϋποθέτει ότι ο ρυθμός άφιξης λ_n εξαρτάται μόνο από το πλήθος των πελατών στο σύστημα, δηλαδή διατηρείται σταθερός.

Στην πραγματικότητα, παρατηρούνται διακριτές περιόδους: περίοδος αιχμής, περίοδος ύφεσης.

Χρονική Περίοδος	Ρυθμός άφιξης	Φόρτος Συστήματος
Πρωινή ώρα (φάση 1)	Υψηλός	Μεγάλος
Μεσημεριανή ώρα (φάση 2)	Χαμηλός	Περιορισμένος

Πίνακας 1.2: Χαρακτηριστικά λειτουργίας συστήματος

Η BD Διαδικασία υποθέτει ότι όλοι οι ρυθμοί είναι αριθμητικά μεγέθη και για να ξεπεραστεί αυτό, στις QBDs εισάγουμε μια δεύτερη, πεπερασμένη διάσταση, μετατρέποντας τους ρυθμούς μετάβασης σε πίνακες ρυθμών.

Για να μοντελοποιηθεί αυτό απαιτείται η εισαγωγή μιας εξωτερικής Μαρκοβιανής Φάσης που περιγράφει την κατάσταση του περιβάλλοντος. Το Επίπεδο συνεχίζει να μετρά τον αριθμό των πελατών, ενώ η Φάση μετρά την τρέχουσα χρονική περίοδο (π.χ. 1 για αιχμή/πρωινή ώρα και 2 για ύφεση/μεσημεριανή ώρα).

Η ανάγκη, λοιπόν, για τη διάσταση της Φάσης γίνεται ιδιαίτερα εμφανής όταν οι ρυθμοί λειτουργίας του συστήματος δεν είναι σταθεροί, αλλά εξαρτώνται από μία άλλη κατάσταση, η οποία εξελίσσεται ανεξάρτητα από το πλήθος των πελατών

Η χρήση των QBDs, με την υιοθέτηση του διπλού χώρου καταστάσεων, επιτρέπει την ανάλυση της δυναμικής του συστήματος, λαμβάνοντας υπόψη την αλληλεπίδραση μεταξύ του πλήθους των πελατών και των μεταβαλλόμενων, χρονικά εξαρτώμενων, συνθηκών άφιξης του συστήματος.

1.2.3 Ορισμός των QBDs και Δομή του Χώρου Καταστάσεων

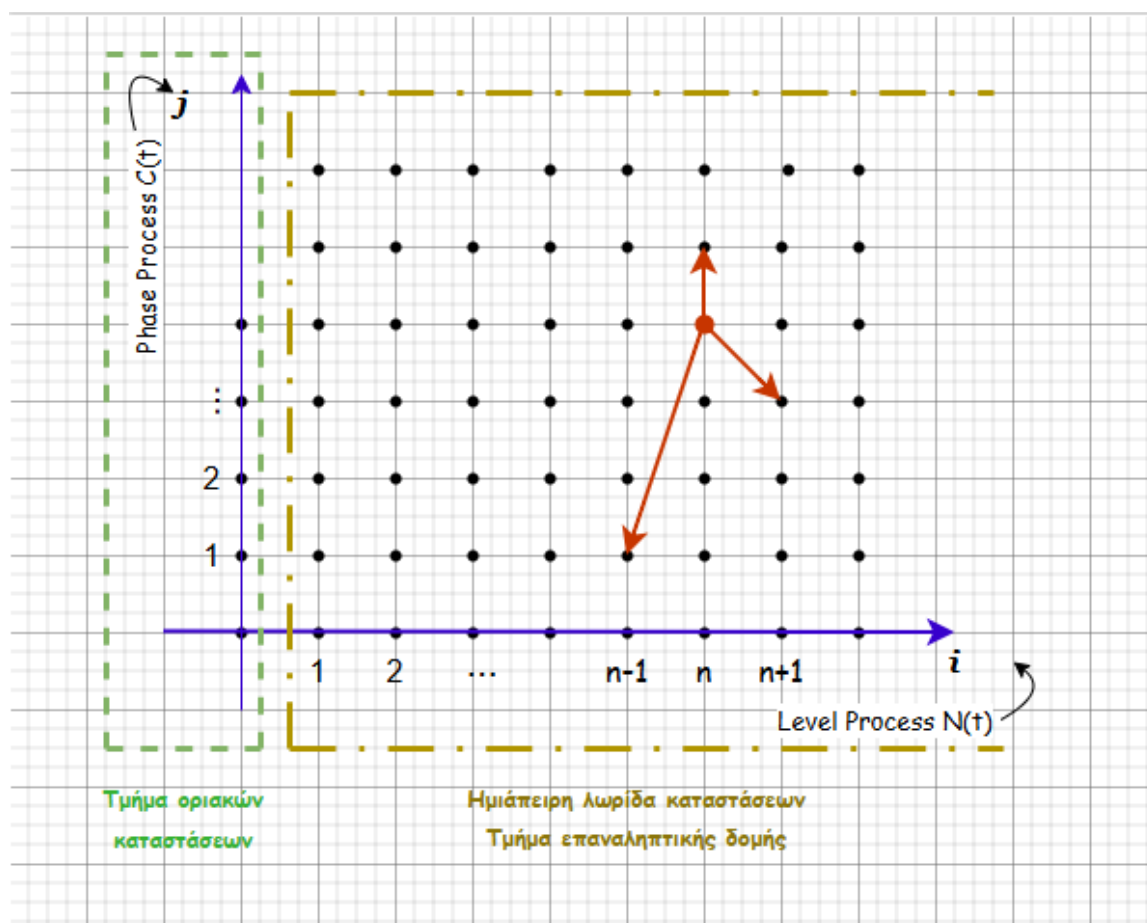
Οι QBDs αποτελούν απλή γενίκευση των Διαδικασιών Γεννήσεων – Θανάτων σε δύο διαστάσεις. Συνεπώς, μια QBD πρόκειται για μια αδιαχώριστη διδιάστατη Μαρκοβιανή Διαδικασία

$$X(t) = \{(N(t), C(t)); t \geq 0\},$$

με χώρο καταστάσεων της μορφής $\{0, 1, 2, \dots\} \times \{0, 1, 2, \dots, N\}$, όπου έχουμε μια μεταβλητή με διπλή είσοδο. Κάθε μία τιμή της $X(t)$ χαρακτηρίζεται ως εξής:

- Επίπεδο (Level), $N(t) = i \in \{0, 1, 2, \dots\}$: η κύρια μεταβλητή (π.χ. μέγεθος ουράς).
- Φάση (Phase), $C(t) = j \in \{0, 1, 2, \dots, N\}$: η δευτερεύουσα μεταβλητή (π.χ. παράμετρος επιρροής).

Για την βαθύτερη κατανόηση της μορφής τέτοιων Διαδικασιών κρίνεται αναγκαία η οπτικοποίηση τόσο του γενικότερου πλαισίου που τις διέπει, όσο και της μορφής των επιτρεπτών μεταβάσεων από μια τυχαία θέση του συστήματος. Με στόχο, λοιπόν, να έχει ο αναγνώστης μια πιο ξεκάθαρη εικόνα του προς μελέτη αντικειμένου, παρουσιάζεται το ακόλουθο γράφημα.



Σχήμα 1.2: Διδιάστατος χώρος καταστάσεων και επιτρεπτές μεταβάσεις μιας διαδικασίας QBD

Οι QBDs διατηρούν τον περιορισμό σε "τοπική εμβέλεια", ως προς το επίπεδο, καθώς επιτρέπονται μεταβάσεις μονάχα μεταξύ γειτονικών επιπέδων ή εντός του ίδιου επιπέδου, όπως φαίνεται και στο παραπάνω σχήμα. Ωστόσο, η φάση μπορεί να αλλάξει (χωρίς περιορισμούς) συγχρόνως με το επίπεδο.

Θεωρούμε μια τυχαία QBD με καταστάσεις (i, j) , όπου $i = 0, 1, 2, \dots$ και $j = 0, 1, \dots, k, \dots, N$, με $N < \infty$. Ο χώρος καταστάσεων διαμερίζεται σε επίπεδα (levels). Στο γενικότερο πλαίσιο, το Επίπεδο Ορίου (Boundary Portion – Level 0) μπορεί να έχει διαφορετική διάσταση (πλήθος φάσεων b) από τα υπόλοιπα επίπεδα και οποιαδήποτε τιμή είναι αποδεκτή, είτε μικρότερη ($b < N$), είτε μεγαλύτερη ($b > N$), είτε ίση ($b = N$). Ωστόσο, η εφαρμογή της Πινακογεωμετρικής Μεθόδου (MGM), πάνω στην οποία στηρίζουμε τη μελέτη των Διαδικασιών QBD προϋποθέτει την ύπαρξη ενός τμήματος επαναληπτικής δομής. Αυτό επιτυγχάνεται υιοθετώντας την παραδοχή ότι όλα τα επίπεδα $i \geq 1$ έχουν την ίδια, κοινή και πεπερασμένη διάσταση N . Αυτή η παραδοχή είναι απαραίτητη διότι εγγυάται την ομοιομορφία στη δομή του πίνακα ρυθμών μετάβασης, με την παρουσία τριών επαναλαμβανόμενων πινάκων. Η μελέτη μοντέλων, όπου ακόμα και τα επίπεδα έχουν διαφορετική διάσταση, λαμβάνοντας διαφορετική τιμή N_i για κάθε level i , έπειτα από το boundary τμήμα, οδηγεί σε αναλυτικά μη διαχειρίσιμες λύσεις, καθιστώντας την προσέγγιση αυτή μη αποδοτική.

Διαμερίζουμε σε $L_0 = \{(0,0), (0,1), \dots, (0,b)\}$, $L_i = \{(i,0), (i,1), \dots, (i,N)\}$, $i \geq 1$, με τα b, N να είναι πεπερασμένοι ακέραιοι, τέτοιοι ώστε: $E = L_0 \cup L_1 \cup L_2 \cup \dots = \bigcup_{i=0}^{\infty} L_i$.

Με "λεξικογραφική προσέγγιση" παρατίθεται η διάταξη των καταστάσεων:

$$\begin{array}{lll} (0,0), (0,1), (0,2), \dots, (0,b), & \text{level 0.} \\ (1,0), (1,1), (1,2), \dots, (1,N), & \text{level 1.} \\ (2,0), (2,1), (2,2), \dots, (2,N), & \text{level 2.} \\ \vdots & \vdots & \vdots \end{array}$$

Μετά την ομαδοποίηση των καταστάσεων ανά επίπεδο, ο άπειρος πίνακας ρυθμών μετάβασης Q της QBD διαθέτει τη χαρακτηριστική block-τριδιαγώνια μορφή, όπου τα στοιχεία του είναι πίνακες αντί για απλούς αριθμούς.

Η δομή του Q είναι η ακόλουθη:

$$Q = \begin{bmatrix} \Lambda_0^{(0)} & \Lambda_1^{(0)} & 0 & 0 & 0 & \dots \\ \Lambda_0^{(1)} & \Lambda_0 & \Lambda_1 & 0 & 0 & \dots \\ 0 & \Lambda_{-1} & \Lambda_0 & \Lambda_1 & 0 & \dots \\ 0 & 0 & \Lambda_{-1} & \Lambda_0 & \Lambda_1 & 0 \\ 0 & 0 & 0 & \Lambda_{-1} & \Lambda_0 & \ddots \\ \vdots & \vdots & \vdots & 0 & \ddots & \ddots \end{bmatrix}.$$

Αναλυτική περιγραφή των block-πινάκων του Q:

$\Lambda_0^{(0)} = \Lambda_{0,0}$: πίνακας ρυθμών μετάβασης μεταξύ των καταστάσεων εντός του level 0, δεν οδηγούμαστε σε αλλαγή level, η διάσταση του είναι $(b+1) \times (b+1)$.

$\Lambda_1^{(0)} = \Lambda_{0,1}$: πίνακας ρυθμών μετάβασης από το level 0 στο level 1, η διάσταση του είναι $(b+1) \times (N+1)$.

$\Lambda_0^{(1)} = \Lambda_{1,0}$: πίνακας ρυθμών μετάβασης από το level 1 στο level 0, η διάσταση του είναι $(N+1) \times (b+1)$.

Οι τρεις ακόλουθοι πίνακες έχουν διάσταση $(N+1) \times (N+1)$ και αποτελούν το μοτίβο των επαναλαμβανόμενων πινάκων:

Λ_{-1} : πίνακας ρυθμών μετάβασης από το level $i > 0$ στο level $i-1$.

Λ_0 : πίνακας ρυθμών μετάβασης μεταξύ των καταστάσεων εντός του level $i > 0$, δεν οδηγούμαστε σε αλλαγή level.

Λ_1 : πίνακας ρυθμών μετάβασης από το level $i > 1$ στο level $i+1$.

Παρατηρήσεις:

1. Τα στοιχεία κάθε γραμμής του πίνακα Q αθροίζουν στο μηδέν.
2. Τα στοιχεία της κύριας διαγωνίου του Q είναι αρνητικά, ενώ τα εκτός διαγωνίου στοιχεία είναι μη αρνητικά.

1.3 Συνθήκη Ευστάθειας και Πινακογεωμετρική Μέθοδος Επίλυσης

Η παρούσα υποενότητα ασχολείται με την Πινακογεωμετρική Μέθοδο (Matrix Geometric Method - MGM), η οποία είναι η κύρια αναλυτική τεχνική για την επίλυση των Διαδικασιών QBD. Η κατάσταση του συστήματος περιγράφεται από το ζεύγος $(N(t), C(t))$, όπου η μεταβλητή $N(t)$ αντιπροσωπεύει το Επίπεδο (Level) και η $C(t)$ την Φάση (Phase). Η μέθοδος αυτή επιτρέπει τον υπολογισμό της στάσιμης κατανομής της QBD, δίνοντας τη δυνατότητα προσδιορισμού των ροών της. Η κεντρική ιδέα πάνω στην οποία λειτουργεί η MGM είναι η εκμετάλλευση της επαναληπτικής δομής του πίνακα ρυθμών μετάβασης Q, με στόχο την εύρεση μιας γεωμετρικής σχέσης μεταξύ των διανυσμάτων πιθανότητας διαδοχικών επιπέδων.

Το υλικό που παρουσιάζεται στην εν λόγω ενότητα προέρχεται από τα ακόλουθα συγγράμματα: *Introduction to Matrix Analytic Methods in Stochastic Modeling* (V. Ramaswani and G. Latouche), *Matrix Geometric Solutions in Stochastic Models: An Algorithmic Approach* (M. F. Neuts) και *Analysis of structured Markov processes* (I. Adan, J. van Leeuwen, J. Selen).

1.3.1 Συνθήκη Ευστάθειας

Η Συνθήκη Ευστάθειας εξασφαλίζει την ύπαρξη της οριακής κατανομής της QBD Διαδικασίας. Για να υπάρχει μοναδική στάσιμη κατανομή, η οποία θα συμπίπτει με την οριακή, θα πρέπει η QBD να είναι ευσταθής. Αυτή η απαίτηση εξαρτάται από το ποσοστό του χρόνου που παραμένει η QBD σε κάθε φάση και από τους ρυθμούς μετάβασης μεταξύ των επιπέδων.

Έτσι, θεωρούμε:

- $\Lambda := \Lambda_{-1} + \Lambda_0 + \Lambda_1$ τον γεννήτορα πίνακα, που δείχνει πώς κινείται η αλυσίδα εντός κάθε level.
- $\mathbf{x} = (x(0), x(1), x(2), \dots, x(r))$ το διάνυσμα, που αποτελεί την μοναδική θετική λύση του συστήματος των εξισώσεων:

$$\mathbf{x}\Lambda = \mathbf{0} \text{ και } \mathbf{x}\mathbf{e} = 1,$$

όπου $\mathbf{0}$ είναι το μηδενικό διάνυσμα και $\mathbf{e}$ το διάνυσμα μονάδων.

Το $x(i)$ είναι το ποσοστό του χρόνου που η QBD παραμένει στην φάση i , χωρίς την επιρροή των αρχικών καταστάσεων (level 0).

- $\mathbf{x}\Lambda_1\mathbf{e} \equiv$ μέση μετατόπιση προς τα δεξιά (mean drift to the right).
- $\mathbf{x}\Lambda_{-1}\mathbf{e} \equiv$ μέση μετατόπιση προς τα αριστερά (mean drift to the left).
- $\mathbf{x}\Lambda_1\mathbf{e} - \mathbf{x}\Lambda_{-1}\mathbf{e} \equiv$ καθαρή μετατόπιση της QBD (net drift).

Για να είναι η QBD ευσταθής, απαιτείται η μέση μετατόπιση προς τα δεξιά να είναι μικρότερη από την μέση μετατόπιση προς τα αριστερά. Σε διαφορετική περίπτωση, το σύστημα θα ολισθαίνει στο άπειρο (drift off), γεγονός που θα δημιουργούσε πρόβλημα στην ευστάθεια και στην ισορροπία του συστήματος. Αυτή η συνθήκη είναι γνωστή ως Συνθήκη Μέσης Ολίσθησης του Neuts (Neuts' Mean Drift Condition) [[9]] ή ως Συνθήκη Ευστάθειας (Stability Condition) και παρουσιάζεται στο παρακάτω βασικό θεώρημα.

Θεώρημα 1.7: Η QBD $\{X(t)\}$ είναι θετικά επαναληπτική (άρα ευσταθής) αν και μόνο αν

$$\mathbf{x}\Lambda_1\mathbf{e} < \mathbf{x}\Lambda_{-1}\mathbf{e},$$

όπου $\mathbf{x} = (x(0), x(1), \dots, x(r))$ είναι κατανομή ισορροπίας της Μαρκοβιανής Διαδικασίας, με γεννήτορα τον πίνακα ρυθμών μετάβασης $\Lambda := \Lambda_{-1} + \Lambda_0 + \Lambda_1$, έτσι ώστε $\mathbf{x}\Lambda = \mathbf{0}$ και $\mathbf{x}\mathbf{e} = 1$.

Παρατήρηση:

Η παραπάνω συνθήκη αποτελεί τη γενίκευση της κλασικής συνθήκης ευστάθειας για τις απλές ομογενείς Birth – Death Διαδικασίες. Στις απλές BD Διαδικασίες, όπου ο ρυθμός γεννήσεων είναι λ και ο ρυθμός θανάτων είναι μ , οι πίνακες Λ_1 και Λ_{-1} είναι

μονοδιάστατοι (1×1), δηλαδή οι $\Lambda_1 = [\lambda]$ και $\Lambda_{-1} = [\mu]$, ενώ το διάνυσμα x είναι απλά ο αριθμός 1. Επομένως, η συνθήκη ευστάθειας ανάγεται στην απλούστερη μορφή:

$$x\Lambda_1 e < x\Lambda_{-1} e \Rightarrow \lambda < \mu.$$

Αυτή η συνθήκη σημαίνει ότι η διαδικασία είναι ευσταθής όταν ο ρυθμός γεννήσεων λ είναι μικρότερος από τον ρυθμό θανάτων μ , σε μία BD Διαδικασία.

1.3.2 Πινακογεωμετρική Μορφή και Βασικές Ιδιότητες του Rate Matrix R

Αφού στόχος της Matrix Geometric Method (MGM) είναι ο υπολογισμός της στάσιμης κατανομής $\pi(i, j) = \lim_{t \rightarrow \infty} P((N(t), C(t) = (i, j)), (i, j) \in E$, της θετικά επαναληπτικής QBD Διαδικασίας, η μέθοδος βασίζεται στην εκμετάλλευση της επαναλαμβανόμενης δομής των πινάκων μετάβασης. Η στάσιμη κατανομή υπάρχει και είναι μοναδική, ικανοποιώντας το βασικό ζεύγος εξισώσεων ισορροπίας $\pi Q = 0$ και $\pi e = 1$.

Θεωρούμε $\pi = (\pi_0, \pi_1, \pi_2, \dots)$, όπου:

$$\pi_0 = (\pi(0, 0), \pi(0, 1), \dots, \pi(0, b)), \text{ με } b+1 \text{ σε πλήθος καταστάσεις,}$$

$$\pi_i = (\pi(i, 0), \pi(i, 1), \dots, \pi(i, r)), \text{ με } r+1 \text{ σε πλήθος καταστάσεις.}$$

Οι εξισώσεις $\pi Q = 0$ ισοδυναμούν με:

$$\pi_0 \Lambda_{0,0} + \pi_1 \Lambda_{1,0} = 0, \quad (1.1)$$

$$\pi_0 \Lambda_{0,1} + \pi_1 \Lambda_0 + \pi_2 \Lambda_{-1} = 0, \quad (1.2)$$

$$\pi_{i-1} \Lambda_1 + \pi_i \Lambda_0 + \pi_{i+1} \Lambda_{-1} = 0, i \geq 2. \quad (1.3)$$

Η γεωμετρική μορφή της στάσιμης κατανομής στα επίπεδα $i \geq 1$ καθορίζεται από το ακόλουθο βασικό θεώρημα, που εμφανίζεται στο [[9]] και που αποτελεί τον πυρήνα της MGM.

Θεώρημα 1.8: Δεδομένης της QBD Διαδικασίας $\{X(t)\}$, που είναι αδιαχώριστη και θετικά επαναληπτική, το διάνυσμα πιθανότητας ισορροπίας π , που ικανοποιεί τις $\pi Q = 0$ και $\pi e = 1$, δίνεται από:

$$\pi_{i+1} = \pi_i R, \quad (1.4)$$

Ως εκ τούτου:

$$\pi_{i+1} = \pi_i R = \pi_{i-1} R^2 = \pi_{i-2} R^3 = \dots = \pi_{i-k} R^{k+1} = \dots = \pi_{i-(i-1)} R^{i-1+1}.$$

Άρα, καταλήγουμε σε:

$$\pi_{i+1} = \pi_1 R^i, i \geq 1, \quad (1.5)$$

όπου ο πίνακας ρυθμού (rate matrix) R είναι η ελάχιστη μη αρνητική ρίζα της πινακικής δευτεροβάθμιας χαρακτηριστικής εξίσωσης:

$$R^2 \Lambda_{-1} + R \Lambda_0 + \Lambda_1 = \mathbf{0}. \quad (1.6)$$

Η σχέση (1.6) προέκυψε από την (1.3), μέσω της (1.5), έπειτα από απλοποίηση με το R^i .

Τα διανύσματα πιθανοτήτων ισορροπίας π_0, π_1 προκύπτουν από τις (1.1), (1.2), (1.5) και τη συνθήκη κανονικοποίησης:

$$\pi e = 1 \Leftrightarrow \pi_0 e + \pi_1 (I - R)^{-1} e = 1, \quad (1.7)$$

της οποίας η απόδειξη παρουσιάζεται ευθύς αμέσως.

Απόδειξη της Συνθήκης Κανονικοποίησης (1.7):

Θέτοντας $i = j - 1$ στην (1.5) και αντικαθιστώντας στην $\sum_{j=0}^{\infty} \pi_j e = 1 \Leftrightarrow \pi_0 e + \sum_{j=1}^{\infty} \pi_j e = 1$, έχουμε:

$$\pi_0 e + \sum_{j=1}^{\infty} \pi_1 R^{j-1} e = 1.$$

Με αλλαγή $j - 1 = k$, το παραπάνω άθροισμα ισοδυναμεί με:

$$\pi_0 e + \pi_1 \left(\sum_{k=0}^{\infty} R^k \right) e = 1.$$

Επιπλέον, αποδεικνύεται ότι από πινακική γεωμετρική σειρά έχουμε: $\sum_{k=0}^{\infty} R^k = (I - R)^{-1}$,

όπως φαίνεται αναλυτικά:

$$\begin{aligned} \sum_{k=0}^{\infty} R^k &= R^0 + R^1 + R^2 + R^3 + \dots \Rightarrow \sum_{k=0}^{\infty} R^k = I + R + R^2 + R^3 + \dots \Rightarrow \sum_{k=0}^{\infty} R^k = I + R(1 + R + R^2 + \dots) \\ &\Rightarrow \sum_{k=0}^{\infty} R^k = I + R \sum_{k=0}^{\infty} R^k, \end{aligned}$$

όπου $R^0 = I$ μοναδιαίος πίνακας.

Υπό την προϋπόθεση σύγκλισης της σειράς, απαιτείται:

$$\sum_{k=0}^{\infty} R^k = (I - R)^{-1}.$$

Συνεπώς αντικαθιστώντας το άθροισμα, καταλήγουμε σε:

$$\pi_0 e + \pi_1 (I - R)^{-1} e = 1. \quad \square$$

Παρατηρήσεις:

1. Η μέθοδος χαρακτηρίζεται ως Πινακογεωμετρική διότι αφενός η ανάλυση διεξάγεται με χρήση πινάκων (αντί για βαθμωτές τιμές) για την περιγραφή των ρυθμών μετάβασης και αφετέρου γιατί η στάσιμη κατανομή των επιπέδων μειώνεται γεωμετρικά με λόγο τον πίνακα R , καθώς το επίπεδο i αυξάνεται.
2. Για να είναι εφαρμόσιμη η συνθήκη κανονικοποίησης απαιτείται ο R να είναι τέτοιος ώστε να επιτρέπεται ο υπολογισμός του όρου $(I - R)^{-1}$, εξασφαλίζοντας ότι η ουρά του συστήματος παραμένει υπό έλεγχο (δεν μεγαλώνει επ' άπειρον) και το άθροισμα όλων των πιθανοτήτων είναι ίσο με 1.

1.3.3 Μέθοδοι Υπολογισμού του R

Αριθμητικός Υπολογισμός – Γενική Περίπτωση

Ο υπολογισμός του R γίνεται συνήθως αριθμητικά, μέσω μίας επαναληπτικής μεθόδου. Η μη γραμμική εξίσωση (1.6) αναδιατάσσεται και παίρνει την ακόλουθη μορφή:

$$R = -(\Lambda_1 + R^2 \Lambda_{-1}) \Lambda_0^{-1}. \quad (1.8)$$

Η παραπάνω μορφή απαιτεί ο Λ_0 να είναι αντιστρέψιμος. Αυτό το γεγονός ικανοποιείται, καθώς ο Λ_0 περιγράφει τις μεταβάσεις εντός του ίδιου επιπέδου, δηλαδή αποτελεί πίνακα ρυθμών μετάβασης μιας επιμέρους Μαρκοβιανής Αλυσίδας, στην οποία δεν εμφανίζονται μεταβάσεις προς μεγαλύτερο ή μικρότερο επίπεδο, αλλά μονάχα μεταβάσεις που αφορούν αλλαγή φάσης. Εφόσον η QBD επιτρέπει την μετακίνηση εκτός των επιπέδων μέσω των Λ_1 και Λ_{-1} , τότε ο Λ_0 διαθέτει τον χαρακτηρισμό της παροδικότητας.

Η (1.8) αποτελεί Εξίσωση Σταθερού Σημείου (fixed point equation), η οποία επιλύεται επαναληπτικά, δηλαδή ξεκινώντας από μία αρχική εκτίμηση και βελτιώνοντας την λύση σε κάθε βήμα.

Για να απλοποιηθεί η μορφή της $R_{n+1} = -(\Lambda_1 + R_n^2 \Lambda_{-1}) \Lambda_0^{-1}$, ορίζονται οι πίνακες:

$$V := \Lambda_1 \Lambda_0^{-1},$$

$$W := \Lambda_{-1} \Lambda_0^{-1},$$

και έτσι η (1.8) παίρνει την εξής αναδρομική μορφή:

$$R_{n+1} = -V - R_n^2 W. \quad (1.9)$$

Θέτουμε αρχική εκτίμηση $R_0 = 0$ και αφήνουμε το R_n να προσεγγίσει το R , για πολύ μεγάλες τιμές του n , δηλαδή $R_n \xrightarrow{n \rightarrow \infty} R$.

Δηλαδή, υπολογίζουμε διαδοχικά τα $R_1, R_2, \dots$ χρησιμοποιώντας την (9) και η επαναληπτική διαδικασία συνεχίζεται μέχρι η διαφορά μεταξύ δύο διαδοχικών επαναλήψεων R_{n+1}, R_n να είναι ικανοποιητικά μικρή, μικρότερη από ένα φράγμα ε .

Μάλιστα, η επαναληπτική διαδικασία τερματίζεται όταν ισχύει ότι $\|R_{n+1} - R_n\| < \varepsilon$, όπου $\|A\| = \max_{i,j} [A_{i,j}]$ είναι η μέγιστη απόλυτη τιμή των στοιχείων του πίνακα που προκύπτει από τη διαφορά των διαδοχικών πινάκων.

Παρατηρήσεις:

1. Προϋπόθεση για να εφαρμοστεί η MGM είναι η ύπαρξη των πινάκων V, W , που πληρείται υπό συνθήκη αντιστρεψιμότητας του Λ_0 .
2. Η ακολουθία των $R_k, k = 0, 1, 2, \dots$ είναι μη φθίνουσα και συγκλίνει στο R , πιθανότατα με αργή σύγκλιση, βάσει του ορίου ανοχής ε .

Έπειτα από τον υπολογισμό του πίνακα R , το επόμενο και τελικό βήμα της μεθόδου περιλαμβάνει τον προσδιορισμό των στάσιμων πιθανοτήτων π_0, π_1 , που προκύπτουν από τις εξισώσεις (1), (2), του boundary portion, συνδυαστικά με τη συνθήκη κανονικοποίησης. Δηλαδή, αντικαθιστώντας $\pi_2 = \pi_1 R$ στην (1.2), επιλύουμε το σύστημα των εξισώσεων.

$$\begin{aligned}\pi_0 \Lambda_{0,0} + \pi_1 \Lambda_{1,0} &= \mathbf{0}, \\ \pi_0 \Lambda_{0,1} + \pi_1 [\Lambda_0 + R \Lambda_{-1}] &= \mathbf{0}, \\ \pi_0 \mathbf{e} + \pi_1 (I - R)^{-1} \mathbf{e} &= 1.\end{aligned}$$

Αναλυτικός Υπολογισμός – Ειδική Περίπτωση

Πέρα από τη γενική αριθμητική επίλυση, σε ειδικές κατηγορίες διαδικασιών QBD, ο πίνακας ρυθμού R μπορεί να υπολογιστεί αναλυτικά.

Τα όσα ακολουθούν βασίζονται στο έργο των Adan, van Leeuwen, Selen [[1]], οι οποίοι εξειδικεύουν τα γενικά αποτελέσματα του Neuts [[9]], εξετάζοντας τη δομή του πίνακα ρυθμών μετάβασης προς το χαμηλότερο επίπεδο, ώστε να προκύψει μια πιο απλή, αναλυτική διαδικασία για τον υπολογισμό του R .

Λήμμα 1.1: Θεωρούμε μια Διαδικασία QBD με πίνακα ρυθμών μετάβασης Q ⁴. Αν ο πίνακας ρυθμών προς το χαμηλότερο επίπεδο (θανάτων) Λ_{-1} έχει την μορφή:

$$\Lambda_{-1} = \mathbf{a} \times \mathbf{b},$$

όπου $\mathbf{a}$ είναι ένα διάνυσμα στήλης και $\mathbf{b}$ ένα διάνυσμα γραμμής και τα δύο διαστάσεων $r+1$, με το $\mathbf{b}$ να έχει άθροισμα στοιχείων 1 και $\mathbf{a}, \mathbf{b} > 0$,

⁴ Όπως αυτός ορίστηκε στην υποπαράγραφο 1.2.3.

δηλαδή $\mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_r \end{pmatrix}$, $\mathbf{b} = (b_0 \ b_1 \ \dots \ b_r)$, $\mathbf{b}\mathbf{e}_r = 1$, $\mathbf{a}, \mathbf{b} > 0$, τότε ο πίνακας R δίνεται από τον τύπο:

$$R = -\Lambda_1 (\Lambda_0 + \Lambda_1 \mathbf{e}\mathbf{b})^{-1},$$

όπου ο πίνακας $(\Lambda_0 + \Lambda_1 \mathbf{e}\mathbf{b})^{-1}$ είναι αντιστρέψιμος.

1.3.4 Πλήρης Μεθοδολογία Επίλυσης της QBD μέσω της MGM

Η πλήρης διαδικασία επίλυσης για την εύρεση της στάσιμης κατανομής π μίας QBD διαδικασίας μέσω της MGM ακολουθεί τα παρακάτω διαδοχικά βήματα:

Βήμα 1: Εξασφάλιση της Block – Τριδιαγώνιας Μορφής

Πριν την έναρξη του αλγορίθμου, εξασφαλίζεται η σωστή λεξικογραφική διάταξη των καταστάσεων (επίπεδα, φάσεις) στον πίνακα ρυθμού μετάβασης Q . Αυτό έχει ως αποτέλεσμα ο Q να έχει την χαρακτηριστική Block – Τριδιαγώνια Μορφή, με τους πίνακες Λ_{-1}, Λ_0 και Λ_1 να περιγράφουν τις μεταβάσεις προς το προηγούμενο, στο ίδιο και στο επόμενο επίπεδο, αντίστοιχα.

Βήμα 2: Εύρεση της Συνθήκης Ευστάθειας

Αφού έχουμε ήδη υπολογίσει το διάνυσμα $\mathbf{x}$, μέσω των σχέσεων $\mathbf{x}\Lambda = \mathbf{0}$ και $\mathbf{x}\mathbf{e} = 1$, προχωράμε στον έλεγχο ευστάθειας του συστήματος. Προκειμένου η QBD να είναι θετικά επαναληπτική, απαιτείται ο μέσος ρυθμός ανόδου να είναι μικρότερος από τον μέσο ρυθμό καθόδου, δηλαδή:

$$\mathbf{x}\Lambda_1 \mathbf{e} < \mathbf{x}\Lambda_{-1} \mathbf{e}.$$

Βήμα 3: Υπολογισμός του Πίνακα Ρυθμού R

Υπολογίζεται ο πίνακας ρυθμού (Rate Matrix) R , μέσω της επίλυσης της εξίσωσης:

$$R^2 \Lambda_{-1} + R \Lambda_0 + \Lambda_1 = \mathbf{0}.$$

Στη γενική περίπτωση, ο R δεν μπορεί να υπολογιστεί αναλυτικά σε κλειστή μορφή και απαιτούνται επαναληπτικοί αριθμητικοί αλγόριθμοι.

Ωστόσο, παραμένει χρήσιμη η αξιοποίηση του λήμματος 1.1 για την εξειδικευμένη μορφή του πίνακα Λ_{-1} , όταν αυτός γράφεται ως εξωτερικό γινόμενο δύο διανυσμάτων (με διάνυσμα γραμμής να έχει άθροισμα στοιχείων 1), καθώς απλοποιεί σημαντικά τους υπολογισμούς, επιτρέποντας την απευθείας εύρεση του πίνακα R από τον τύπο:

$$R = -\Lambda_1 (\Lambda_0 + \Lambda_1 \mathbf{e}\mathbf{b})^{-1}.$$

Βήμα 4: Λύση των Εξισώσεων Ισορροπίας Ορίου

Επιλύεται το σύστημα γραμμικών εξισώσεων που προκύπτει από τις εξισώσεις ισορροπίας του πίνακα Q , για τα Level 0 και Level 1 (boundary portion):

$$\begin{aligned}\pi_0 \Lambda_{0,0} + \pi_1 \Lambda_{1,0} &= \mathbf{0}, \\ \pi_0 \Lambda_{0,1} + \pi_1 [\Lambda_0 + R \Lambda_{-1}] &= \mathbf{0},\end{aligned}$$

λαμβάνοντας υπόψιν τον υπολογισμένο πίνακα R .

Αυτό το σύστημα δίνει τα διανύσματα π_0 και π_1 συναρτήσεως μίας άγνωστης σταθεράς κανονικοποίησης. Επί της ουσίας, το σύστημα αποτελείται από γραμμικά εξαρτημένες εξισώσεις και η μοναδικότητα της λύσης του θα προκύψει στην πορεία από την εφαρμογή της συνθήκης κανονικοποίησης.

Βήμα 5: Εφαρμογή της Συνθήκης Κανονικοποίησης

Για να βρεθεί η μοναδική λύση του συστήματος, εφαρμόζεται η συνθήκη κανονικοποίησης, η οποία απαιτεί το άθροισμα όλων των στάσιμων πιθανοτήτων να ισούται με την μονάδα:

$$\sum_{i=0}^{\infty} \pi_i \mathbf{e} = 1.$$

Λόγω της μορφής της λύσης και υπό την προϋπόθεση ότι το σύστημα είναι ευσταθές, για να ορίζεται ο όρος $(I - R)^{-1}$, η παραπάνω σχέση παίρνει την ακόλουθη μορφή:

$$\pi_0 \mathbf{e} + \pi_1 (I - R)^{-1} \mathbf{e} = 1.$$

Μέσω αυτής προσδιορίζονται πλήρως τα π_0 και π_1 . Αυτό δίνει τη δυνατότητα, για κάθε επίπεδο i , να υπολογιστούν στην συνέχεια όλες οι οριακές πιθανότητες του συστήματος.

Βήμα 6: Χρήση του Αναδρομικού Σχήματος

Στο τελευταίο στάδιο, χρησιμοποιούμε το γεωμετρικό αναδρομικό σχήμα για τον υπολογισμό των στάσιμων κατανομών για όλα τα ανώτερα επίπεδα $i > 1$. Το διάνυσμα πιθανότητας ισορροπίας κάθε επιπέδου προκύπτει μέσω του πίνακα ρυθμού R , σύμφωνα με τον τύπο:

$$\pi_n = \pi_{n-1} R, \text{ για } n \geq 2.$$

Εναλλακτικά, γίνεται χρήση του ισοδύναμου αναδρομικού σχήματος:

$$\pi_n = \pi_1 R^{n-1}, \text{ για } n \geq 2.$$

1.4 Κλασικά Παραδείγματα

Στόχος της ενότητας αυτής είναι η εξοικείωση με την MGM, ώστε να μπορέσει η τεχνική αυτή να αποτελέσει ένα χρήσιμο εργαλείο για την μελέτη των κεφαλαίων που θα ακολουθήσουν. Η ενότητα ξεκινάει με δύο σημαντικά αντιπροσωπευτικά μοντέλα τύπου M/M/1 (το Set-Up/Turn-Off και το Breakdown/Repair). Η επιλογή αυτών των δύο παραδειγμάτων έγινε ώστε να εστιάσουμε στη διαφορά του επιπέδου 0 (*boundary portion*) μεταξύ των δύο μοντέλων (M/M/1), αναδεικνύοντας την ευελιξία της MGM στον χειρισμό των ειδικών μεταβάσεων στην αρχή της αλυσίδας. Τέλος, η ενότητα ολοκληρώνεται με το M/Cox/1, ένα θεμελιώδες μοντέλο τύπου M/G/1.

1.4.1 M/M/1 με Λειτουργία Set-Up/Turn-Off

Το μοντέλο αυτό περιγράφει ένα σύστημα όπου ο εξυπηρετητής (μηχανή) δεν είναι μόνιμα ενεργός, αλλά απενεργοποιείται όταν το σύστημα αδειάσει ($n = 0$) και χρειάζεται χρόνο για να επανέλθει σε λειτουργία.

Περιγραφή Χώρου Καταστάσεων

Αφίξεις: Poisson με ρυθμό λ

Εξυπηρέτηση: Εκθετική με ρυθμό μ

Χρόνος Ενεργοποίησης (Set-Up Time): Εκθετικός με ρυθμό θ

Χρόνος Απενεργοποίησης (Turn-Off Time): Μηδενικός χρόνος κλεισίματος μηχανής

Καταστάσεις ($X(t)$): Το σύστημα περιγράφεται από το ζεύγος $(X_1(t), X_2(t))$, όπου η διαδικασία είναι QBD, με:

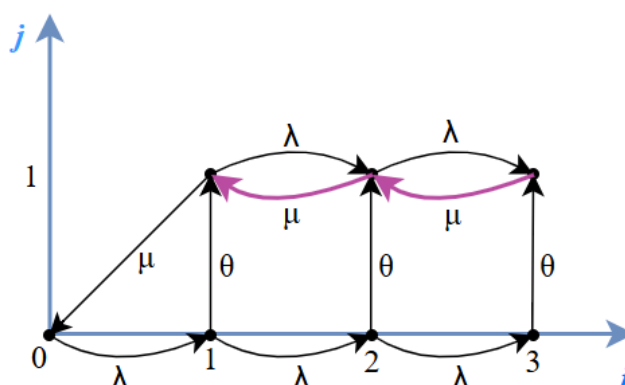
- $X_1(t), (Level\ i)$: ο αριθμός των προϊόντων προς επεξεργασία.
- $X_2(t), (Phase\ j)$: η κατάσταση της μηχανής τη χρονική στιγμή t , με 0 όταν είναι απενεργοποιημένη και 1 όταν είναι σε λειτουργία.

Ο χώρος καταστάσεων διαμερίζεται σε επίπεδα ως εξής: $E = \{0\} \cup \{(i, j); i \in \mathbf{N}, j \in \{0, 1\}\}$, με $\mathbf{N} = \{1, 2, \dots\}$.

Στο σημείο αυτό, να παρατηρήσουμε ότι στο επίπεδο L_0 οι καταστάσεις (0,0) (μηδέν πελάτες, μηχανή απενεργοποιημένη) και (0,1) (μηδέν πελάτες, μηχανή σε λειτουργία) που αναμενόταν να υπάρχουν έχουν συγχωνευτεί σε μία ενιαία κατάσταση $\{0\}$. Αυτό συμβαίνει διότι αν το σύστημα βρεθεί στην κατάσταση (0,1), πρέπει άμεσα (σε χρόνο $t \rightarrow 0$) να μεταβεί στην (0,0) και αυτό εξαιτίας του μηδενικού χρόνου κλεισίματος της μηχανής. Εφόσον, λοιπόν το turn-off της μηχανής είναι ακαριαίο, δηλαδή η (0,1) δεν είναι μια κατάσταση στην οποία το σύστημα μπορεί να παραμείνει, η μόνη στάσιμη κατανομή στο επίπεδο 0 είναι όταν η μηχανή είναι απενεργοποιημένη, που εκλαμβάνεται ως $\pi(0)$.

Διάγραμμα Μεταβάσεων

Το διάγραμμα μεταβάσεων της αλυσίδας απεικονίζεται σε ένα διδιάστατο καρτεσιανό σύστημα με τον οριζόντιο άξονα να μετρά τα Επίπεδα και τον κατακόρυφο τις Φάσεις.



Σχήμα 1.3: Διάγραμμα μεταβάσεων του μοντέλου $M/M/1$ με Set-Up/Turn-Off

Συνδεσιμότητα Καταστάσεων

Ακολουθεί πίνακας συνδεδεμένων καταστάσεων και αντίστοιχων χρόνων των μεταξύ τους μεταβάσεων.

Κατάσταση	Επόμενη Κατάσταση	Χρόνος Μετάβασης σε Επόμενη Κατάσταση
0	(1,0)	$T_{0,(1,0)} \sim \exp(\lambda)$
$(i,0), i \geq 1$	$(i+1,0)$	$T_{(i,0),(i+1,0)} \sim \exp(\lambda)$
	$(i,1)$	$T_{(i,0),(i,1)} \sim \exp(\theta)$
$(i,1), i \geq 2$	$(i+1,1)$	$T_{(i,1),(i+1,1)} \sim \exp(\lambda)$
	$(i-1,1)$	$T_{(i,1),(i-1,1)} \sim \exp(\mu)$
(1,1)	(2,1)	$T_{(1,1),(2,1)} \sim \exp(\lambda)$
	0	$T_{(1,1),0} \sim \exp(\mu)$

Πίνακας 1.3: Συνδεσιμότητα καταστάσεων και χρόνοι μετάβασης

Πίνακας Ρυθμών Μετάβασης

Η πινακική μορφή των εξισώσεων ολικής ισορροπίας δίνεται από τις σχέσεις (1.1), (1.2), (1.3), που παρουσιάζονται στην 1.3.2. Η μορφή του γεννήτορα πίνακα της $(X(t))$ είναι αυτή που προβάλλεται στην ίδια υποπαράγραφο, εξασφαλίζοντας την block – τριδιαγώνια μορφή του, με υποπίνακες διάστασης 2×2 για $i \geq 1$ και ειδικές διαστάσεις για την αρχική κατάσταση $i = 0$.

Οι υποπίνακες του επαναληπτικού τμήματος είναι:

$$\Lambda_{-1} = \begin{bmatrix} 0 & 0 \\ 0 & \mu \end{bmatrix}, \Lambda_0 = \begin{bmatrix} -(\lambda + \theta) & \theta \\ 0 & -(\lambda + \mu) \end{bmatrix}, \Lambda_1 = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}.$$

Ενώ, οι υποπίνακες του οριακού τμήματος είναι:

$$\Lambda_{1,0} = \begin{bmatrix} 0 \\ \mu \end{bmatrix}, \Lambda_{0,0} = [-\lambda], \Lambda_{0,1} = [\lambda \quad 0].$$

Συνθήκη Ευστάθειας

Η συνθήκη ευστάθειας προκύπτει έπειτα από την επίλυση του συστήματος των $\mathbf{x}\Lambda = \mathbf{0}$ και $\mathbf{x}\mathbf{e} = 1$, όπου $\Lambda = \Lambda_{-1} + \Lambda_0 + \Lambda_1$. Άρα:

$$\Lambda = \begin{bmatrix} 0 & 0 \\ 0 & \mu \end{bmatrix} + \begin{bmatrix} -(\lambda + \theta) & \theta \\ 0 & -(\lambda + \mu) \end{bmatrix} + \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -\theta & \theta \\ 0 & 0 \end{bmatrix}.$$

Το στάσιμο διάνυσμα προκύπτει ως εξής:

$$\left\{ \begin{array}{l} \mathbf{x}\Lambda = \mathbf{0} \Rightarrow \begin{bmatrix} x(0) & x(1) \end{bmatrix} \begin{bmatrix} -\theta & \theta \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \theta x(0) = 0 \Rightarrow x(0) = 0 \\ \mathbf{x}\mathbf{e} = 1 \Rightarrow x(0) + x(1) = 1 \end{array} \right\} \Rightarrow x(0) = 0, x(1) = 1.$$

Συνεπώς, για να ικανοποιηθεί η συνθήκη ευστάθειας του συστήματος, απαιτείται:

$$\mathbf{x}\Lambda_1\mathbf{e} < \mathbf{x}\Lambda_{-1}\mathbf{e} \Rightarrow (0 \quad 1) \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} < (0 \quad 1) \begin{bmatrix} 0 & 0 \\ 0 & \mu \end{bmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow [0 \quad \lambda] \begin{pmatrix} 1 \\ 1 \end{pmatrix} < [0 \quad \mu] \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow \lambda < \mu.$$

Υπολογισμός του $\mathbf{R}$

Επειδή ο πίνακας έχει ειδική μορφή $\Lambda_{-1} = \mathbf{a} \times \mathbf{b}$, με $\mathbf{a} = \begin{pmatrix} 0 \\ \mu \end{pmatrix}$, $\mathbf{b} = (0 \quad 1)$ και το άθροισμα των στοιχείων του $\mathbf{b}$ ισούται με 1, στηριζόμενοι στο Λήμμα 1.1, ο πίνακας $\mathbf{R}$ δίνεται από τον τύπο:

$$\mathbf{R} = -\Lambda_1 (\Lambda_0 + \Lambda_1 \mathbf{e} \mathbf{b})^{-1}.$$

Οπότε, αναλυτικά έχουμε:

$$\begin{aligned} \mathbf{R} &= -\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \left(\begin{bmatrix} -(\lambda + \theta) & \theta \\ 0 & -(\lambda + \mu) \end{bmatrix} + \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} (0 \quad 1) \right)^{-1} = \begin{bmatrix} -\lambda & 0 \\ 0 & -\lambda \end{bmatrix} \begin{bmatrix} -\lambda - \theta & \lambda + \theta \\ 0 & -\mu \end{bmatrix}^{-1} \\ &= \begin{bmatrix} -\lambda & 0 \\ 0 & -\lambda \end{bmatrix} \cdot \frac{1}{(\lambda + \theta)\mu} \cdot \begin{bmatrix} -\mu & -\lambda - \theta \\ 0 & -\lambda - \theta \end{bmatrix} = \frac{1}{(\lambda + \theta)\mu} \cdot \begin{bmatrix} \lambda\mu & \lambda(\lambda + \theta) \\ 0 & \lambda(\lambda + \theta) \end{bmatrix} = \begin{bmatrix} \frac{\lambda}{\lambda + \theta} & \frac{\lambda}{\mu} \\ 0 & \frac{\lambda}{\mu} \end{bmatrix}. \end{aligned}$$

Σχόλιο: Η άνω τριγωνική μορφή του R οφείλεται στο ότι το στοιχείο στην θέση $(1,0)$ περιγράφει την πιθανότητα μετάβασης από την $(i-1, 1)$ στην $(i,0)$ κατάσταση, το οποίο είναι αδύνατον να συμβεί, καθώς η μηχανή δεν μπορεί να απενεργοποιηθεί –δηλαδή να υπάρξει μετάβαση από το 1 (on) στο 0 (off)– όταν βρίσκεται έστω και ένας πελάτης στο σύστημα.

Επίλυση Εξισώσεων Ισορροπίας

Πλέον, σκοπός είναι να βρεθούν τα $\pi(0)$, $\pi_1 = (\pi(1,0), \pi(1,1))$, μέσω των (1.1), (1.2) και της συνθήκης κανονικοποίησης.

Με αντικατάσταση των πινάκων και διαδοχικές πράξεις στην (1.1), προκύπτει :

$$\pi(0)[- \lambda] + \pi_1 \begin{bmatrix} 0 \\ \mu \end{bmatrix} = 0 \Rightarrow \lambda \pi(0) = \pi_1 \begin{bmatrix} 0 \\ \mu \end{bmatrix} \Rightarrow \pi(0) = \frac{1}{\lambda} (\pi(1,0) \quad \pi(1,1)) \begin{bmatrix} 0 \\ \mu \end{bmatrix} \Rightarrow \pi(0) = \frac{\mu}{\lambda} \pi(1,1).$$

Βάζοντας στην σχέση (1.4) όπου i το $i-1$, παίρνουμε $\pi_2 = \pi_1 R$ και εφαρμόζοντας στην (1.2), εξαιτίας του 2×2 συστήματος, έχουμε τις εξής δύο σχέσεις:

$$\begin{cases} \lambda \pi(0) - (\lambda + \theta) \pi(1,0) = 0, \\ (\lambda + \theta) \pi(1,0) - \mu \pi(1,1) = 0. \end{cases}$$

Με άθροισμα κατά μέλη των δυο τελευταίων σχέσεων έχουμε:

$$\lambda \pi(0) - \mu \pi(1,1) = 0 \Leftrightarrow \pi(1,1) = \frac{\lambda}{\mu} \pi(0).$$

Τώρα, λύνοντας την πρώτη από τις εξισώσεις του συστήματος ως προς $\pi(1,0)$, προκύπτει:

$$\pi(1,0) = \frac{\lambda}{\lambda + \theta} \pi(0).$$

Έχοντας, πλέον, προσδιορίσει τα $\pi(1,0)$ και $\pi(1,1)$, συναρτήσει της $\pi(0)$, τα εφαρμόζουμε πάνω στην συνθήκη κανονικοποίησης (1.7). Έπειτα από διαδοχικές πράξεις αναγωγής, απλοποίησης και αναδιάταξης των όρων, καταλήγουμε στο συμπέρασμα ότι:

$$\pi(0) = \frac{\theta}{\lambda + \theta} \left(1 - \frac{\lambda}{\mu} \right).$$

Επιστρέφοντας στην $\pi(1,0) = \frac{\lambda}{\lambda + \theta} \pi(0)$ και στην $\pi(1,1) = \frac{\lambda}{\mu} \pi(0)$, καθορίζουμε την τελική μορφή των συνιστωσών του π_1 . Έτσι, στο εξής, μπορούμε να υπολογίσουμε κάθε οριακή πιθανότητα π_i , μέσω του $\pi_i = \pi_1 R^{i-1}$, για $i \geq 2$.

1.4.2 M/M/1 με Διακοπές Λειτουργίας και Επιδιόρθωση (Breakdown/Repair)

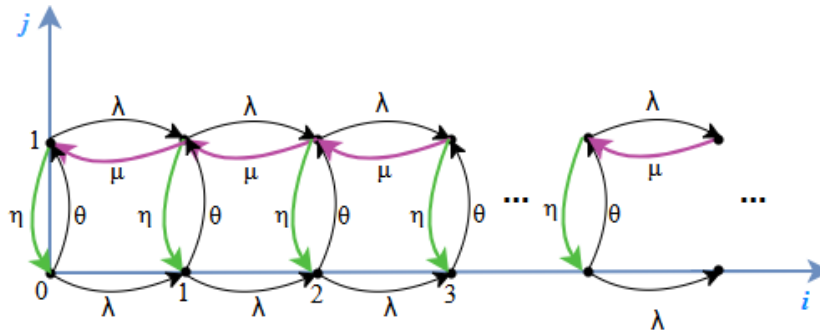
Το παρόν μοντέλο περιγράφει ένα σύστημα όπου ο εξυπηρετητής εναλλάσσεται μεταξύ λειτουργίας Up και Down. Αυτό που διαφοροποιείται στην περίπτωση αυτή, συγκριτικά με την παραπάνω περίπτωση 1.4.1, που μελετήθηκε προηγουμένως, έγκειται στην διαφορετική μορφή των καταστάσεων, στο boundary portion της αλυσίδας. Η ειδοποιός διαφορά έχει να κάνει με το ότι ο χώρος καταστάσεων στο επίπεδο 0 διατηρεί τις δύο φάσεις $L_0 = \{(0,0), (0,1)\}$, αφού το σύστημα μπορεί να παραμείνει είτε σε λειτουργία, είτε σε επισκευή, ακόμα και όταν είναι άδειο.

Χώρος Καταστάσεων

Ο χώρος καταστάσεων είναι $E = \{(i, j) : i \in N_0, j \in \{0,1\}\}$, $N_0 = \{0,1,2,\dots\}$.

Διάγραμμα Μεταβάσεων και Συνδεσιμότητα Καταστάσεων

Ακολουθεί το διάγραμμα μεταβάσεων και ο πίνακας συνδεδεμένων καταστάσεων με τους αντίστοιχους χρόνους μεταβάσεων.



Σχήμα 1.4: Διάγραμμα μεταβάσεων του μοντέλου M/M/1 με Breakdown/Repair

Κατάσταση	Επόμενη Κατάσταση	Χρόνος Μετάβασης σε Επόμενη Κατάσταση
$(i,0), i \geq 0$	$(i+1,0)$	$T_{(i,0),(i+1,0)} \sim \exp(\lambda)$
	$(i,1)$	$T_{(i,0),(i,1)} \sim \exp(\theta)$
$(i,1), i \geq 0$	$(i+1,1)$	$T_{(i,1),(i+1,1)} \sim \exp(\lambda)$
	$(i-1,1), i \geq 1$	$T_{(i,1),(i-1,1)} \sim \exp(\mu)$
	$(i,0)$	$T_{(i,1),(i,0)} \sim \exp(\eta)$

Πίνακας 1.4: Συνδεσιμότητα καταστάσεων και χρόνοι μετάβασης

Πίνακας Ρυθμών Μετάβασης

Η ανάλυση γίνεται με την MGM, αφού ο γεννήτορας Q διαθέτει block-τριδιαγώνια μορφή και οι υποπίνακες Λ_{-1} (εξυπηρέτηση), Λ_0 (εντός επιπέδου), Λ_1 (αφίξεις), του repetitive portion, είναι:

$$\Lambda_{-1} = \begin{bmatrix} 0 & 0 \\ 0 & \mu \end{bmatrix}, \Lambda_0 = \begin{bmatrix} -(\lambda + \theta) & \theta \\ \eta & -(\lambda + \mu + \eta) \end{bmatrix}, \Lambda_1 = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}.$$

Παρατηρούμε ότι οι Λ_{-1} , Λ_1 παραμένουν ίδιοι με αυτούς του προηγούμενου παραδείγματος. Οι υποπίνακες του boundary portion διαμορφώνονται ως εξής:

$$\Lambda_{1,0} = \Lambda_{-1}, \Lambda_{0,0} = \begin{bmatrix} -(\lambda + \theta) & \theta \\ \eta & -(\lambda + \eta) \end{bmatrix}, \Lambda_{0,1} = \Lambda_1.$$

Η γεννήτρια Λ του συστήματος, που δείχνει την κίνηση μεταξύ των καταστάσεων της φάσης σε κάθε επίπεδο, λαμβάνει την μορφή: $\Lambda = \begin{bmatrix} -\theta & \theta \\ \eta & -\eta \end{bmatrix}$. Η μορφή αυτή είναι αναμενόμενη, καθώς, πλέον, ο χρόνος κλεισίματος της μηχανής δεν είναι αμελητέος.

Συνθήκη Ευστάθειας

Το διάνυσμα $\mathbf{x} = (x(0), x(1))$, μετά από την επίλυση των $\mathbf{x}\Lambda = \mathbf{0}$, $\mathbf{x}\mathbf{e} = 1$, είναι:

$$\mathbf{x} = \left(\frac{\eta}{\theta + \eta}, \frac{\theta}{\theta + \eta} \right).$$

Η συνθήκη ευστάθειας του συστήματος, με την απαίτηση $\mathbf{x}\Lambda_1\mathbf{e} < \mathbf{x}\Lambda_{-1}\mathbf{e}$, οδηγεί στην σχέση:

$$\lambda < \mu \frac{\theta}{\theta + \eta}.$$

Υπολογισμός του $\mathbf{R}$

Εξαιτίας της ειδικής μορφής του Λ_{-1} , ο υπολογισμός του $\mathbf{R}$ δίνεται και πάλι από τον τύπο

$\mathbf{R} = -\Lambda_1 (\Lambda_0 + \Lambda_1 \mathbf{e} \mathbf{b})^{-1}$, με το τελικό αποτέλεσμα, μετά από διαδοχικές πράξεις να είναι:

$$\mathbf{R} = \frac{\lambda}{\mu} \begin{bmatrix} \frac{\eta + \mu}{\lambda + \theta} & 1 \\ \frac{\eta}{\lambda + \theta} & 1 \end{bmatrix}.$$

Επίλυση Εξισώσεων Ισορροπίας

Η επίλυση των εξισώσεων ισορροπίας οριακού τμήματος, σε συνδυασμό με τη συνθήκη κανονικοποίησης, δίνει την αναλυτική έκφραση του διδιάστατου $\boldsymbol{\pi}_0$, δηλαδή προσδιορίζει τα $\pi(0,0)$ και $\pi(0,1)$, συναρτήσει των ρυθμών $\lambda, \mu, \theta, \eta$ του συστήματος.

Οπότε, έχουμε:

$$\pi(0,0) = \frac{\eta}{\eta + \theta} \left(1 - \frac{\lambda}{\mu} \cdot \frac{\eta + \mu + \theta}{\lambda + \theta} \right),$$

$$\pi(0,1) = \frac{\lambda + \theta}{\eta + \theta} \left(1 - \frac{\lambda}{\mu} \cdot \frac{\eta + \mu + \theta}{\lambda + \theta} \right).$$

Τέλος, αφού υπολογίστηκε ο πίνακας $\mathbf{R}$, αλλά και το $\boldsymbol{\pi}_0$, ο υπολογισμός των υπόλοιπων οριακών πιθανοτήτων καθορίζεται από την $\boldsymbol{\pi}_i = \boldsymbol{\pi}_0 \mathbf{R}^i$, για κάθε $i \geq 1$.

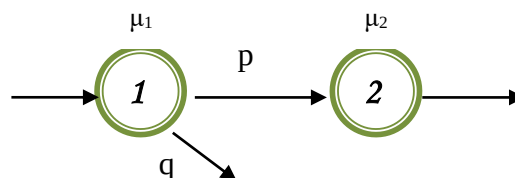
1.4.3 M/Cox/1

Περιγραφή της Κατανομής Coxian (Phase – Type)

Η Coxian Κατανομή είναι ένας τύπος Κατανομής Φάσης και χρησιμοποιείται για την αναπαράσταση οποιασδήποτε κατανομής χρόνου εξυπηρέτησης. Στο συγκεκριμένο μοντέλο 2 φάσεων, έχουμε:

- Φάση 1: ο χρόνος είναι εκθετικός $\exp(\mu_1)$.
Με την ολοκλήρωση της φάσης αυτής, ο πελάτης προχωράει στην επόμενη φάση (με πιθανότητα p) ή ολοκληρώνει την εξυπηρέτηση (με πιθανότητα $q = 1 - p$).
- Φάση 2: ο χρόνος είναι εκθετικός $\exp(\mu_2)$.
Με την ολοκλήρωση της φάσης αυτής, η εξυπηρέτηση φτάνει στο τέλος της.

Οι μεταβάσεις μεταξύ των φάσεων φαίνονται στο ακόλουθο διάγραμμα.



Σχήμα 1.5: Διάγραμμα μεταβάσεων των φάσεων της Coxian (Phase-Type)

Περιγραφή του Μοντέλου

Το μοντέλο M/Cox/1 περιγράφει ένα σύστημα όπου οι αφίξεις ακολουθούν διαδικασία Poisson με ρυθμό λ , ενώ ο χρόνος εξυπηρέτησης ακολουθεί την Coxian Κατανομή $k=2$ φάσεων.

Χώρος Καταστάσεων

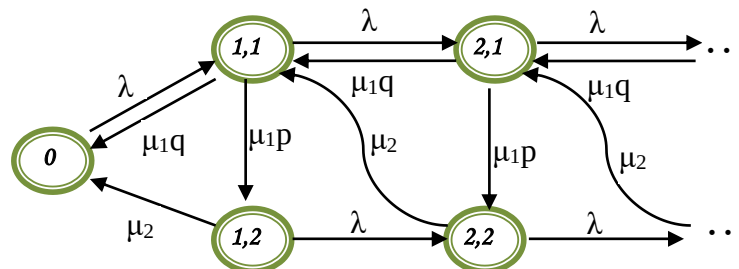
Η κατάσταση του συστήματος περιγράφεται από ένα ζεύγος $(N(t), C(t))$, όπου η διαδικασία είναι QBD, με:

- $N(t), (Level\ n)$: ο αριθμός των πελατών στο σύστημα.
- $C(t), (Phase\ k)$: η φάση εξυπηρέτησης του τρέχοντος πελάτη τη χρονική στιγμή t .

Ο χώρος καταστάσεων E ορίζεται ως $E = \{0\} \cup \{(n, k); n \geq 1, k = 1, 2\}$.

Διάγραμμα Μεταβάσεων

Το διάγραμμα μεταβάσεων του συστήματος είναι το ακόλουθο:



Σχήμα 1.6: Διάγραμμα μεταβάσεων του μοντέλου M/Cox/1

Πίνακας Ρυθμών Μετάβασης

Ο πίνακας Q έχει την δομή QBD, με επαναληπτικό τμήμα για $n \geq 2$ και οριακό για $n = 0, 1$.

Άρα έχει διασφαλιστεί η block – τριδιαγώνια μορφή του, όπως απεικονίζεται παρακάτω:

$$Q = \begin{bmatrix} L_0 & F_0 & 0 & 0 & 0 \\ B_0 & L & F & 0 & 0 \\ 0 & B & L & F & 0 \\ 0 & 0 & \ddots & \ddots & \ddots \end{bmatrix}.$$

Οι υποπίνακες του επαναληπτικού τμήματος ($n \geq 2$) είναι οι 2×2 πίνακες:

$$\text{Μετάβαση προς τα κάτω: } B = \begin{bmatrix} q\mu_1 & 0 \\ \mu_2 & 0 \end{bmatrix}.$$

$$\text{Μετάβαση εντός του κάθε επιπέδου: } L = \begin{bmatrix} -(\lambda + \mu_1) & p\mu_1 \\ 0 & -(\lambda + \mu_2) \end{bmatrix}.$$

$$\text{Μετάβαση προς τα πάνω: } F = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}.$$

Οι υποπίνακες του οριακού τμήματος είναι οι πίνακες:

$$B_0 = \begin{bmatrix} \mu_1 q \\ \mu_2 \end{bmatrix}, L_0 = [-\lambda], F_0 = \begin{bmatrix} \lambda & 0 \end{bmatrix}.$$

Συνθήκη Ευστάθειας

Έχουμε $\Lambda = \Lambda_{-1} + \Lambda_0 + \Lambda_1 = B + L + F$, την γεννήτρια του επαναληπτικού τμήματος της QBD, να διαμορφώνεται ως εξής:

$$\Lambda = \begin{bmatrix} q\mu_1 & 0 \\ \mu_2 & 0 \end{bmatrix} + \begin{bmatrix} -(\lambda + \mu_1) & p\mu_1 \\ 0 & -(\lambda + \mu_2) \end{bmatrix} + \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -(1-q)\mu_1 & p\mu_1 \\ \mu_2 & -\mu_2 \end{bmatrix} = \begin{bmatrix} -p\mu_1 & p\mu_1 \\ \mu_2 & -\mu_2 \end{bmatrix}.$$

Για την εύρεση του διανύσματος $\mathbf{x}$, έχουμε:

$$\mathbf{x}\Lambda = \mathbf{0} \Rightarrow x(1)p\mu_1 - x(2)\mu_2 = 0$$

και

$$\mathbf{x}\mathbf{e} = 1 \Rightarrow x(1) + x(2) = 1.$$

Τα οποία, μέσω επίλυσης γραμμικού συστήματος, μας οδηγούν σε:

$$\mathbf{x} = \left(\frac{\mu_2}{p\mu_1 + \mu_2}, \frac{p\mu_1}{p\mu_1 + \mu_2} \right).$$

Επειδή ένα σύστημα είναι ευσταθές αν και μόνο αν ισχύει $\mathbf{x}\Lambda_1\mathbf{e} < \mathbf{x}\Lambda_{-1}\mathbf{e}$, τότε διαδοχικά έχουμε:

$$\begin{aligned} \mathbf{xFe} < \mathbf{xBe} &\Leftrightarrow \begin{bmatrix} \frac{\mu_2}{p\mu_1 + \mu_2} & \frac{p\mu_1}{p\mu_1 + \mu_2} \end{bmatrix} \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} < \begin{bmatrix} \frac{\mu_2}{p\mu_1 + \mu_2} & \frac{p\mu_1}{p\mu_1 + \mu_2} \end{bmatrix} \begin{bmatrix} q\mu_1 & 0 \\ \mu_2 & 0 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &\Leftrightarrow \frac{\lambda\mu_2}{p\mu_1 + \mu_2} + \frac{\lambda p\mu_1}{p\mu_1 + \mu_2} < \frac{q\mu_1\mu_2}{p\mu_1 + \mu_2} + \frac{p\mu_1\mu_2}{p\mu_1 + \mu_2} \\ &\Leftrightarrow \lambda\mu_2 + \lambda p\mu_1 < q\mu_1\mu_2 + p\mu_1\mu_2 \\ &\Leftrightarrow \lambda(\mu_2 + p\mu_1) < \mu_1\mu_2(q + p) \\ &\Leftrightarrow \lambda < \frac{\mu_1\mu_2}{\mu_2 + p\mu_1}. \end{aligned}$$

Υπολογισμός του $\mathbf{R}$

Ο πίνακας ρυθμών $\mathbf{R}$ είναι η ελάχιστη μη αρνητική ρίζα της Πινακικής Δευτεροβάθμιας Εξίσωσης:

$$\mathbf{R}^2\Lambda_{-1} + \mathbf{R}\Lambda_0 + \Lambda_1 = \mathbf{0} \Leftrightarrow \mathbf{BR}^2 + \mathbf{LR} + \mathbf{F} = \mathbf{0}.$$

Ο πίνακας αυτός υπολογίζεται αριθμητικά μέσω της εξίσωσης σταθερού σημείου:

$$\mathbf{R}_{n+1} = -(\Lambda_1 + \mathbf{R}_n^2\Lambda_{-1})\Lambda_0^{-1} \Leftrightarrow \mathbf{R}_{n+1} = -(F + \mathbf{BR}_n^2)L^{-1}.$$

με $\mathbf{R}_0 = \mathbf{0}$, έτσι ώστε ο $\mathbf{R}_n$ να συγκλίνει στο $\mathbf{R}$, καθώς $n \rightarrow \infty$.

Επίλυση Εξισώσεων Ισορροπίας

Αφού υπολογιστεί ο $\mathbf{R}$, οι π_0 , π_1 προκύπτουν από τις εξισώσεις ισορροπίας στην αρχή της αλυσίδας:

$$\begin{aligned} \pi_0\Lambda_{0,0} + \pi_1\Lambda_{1,0} &= \mathbf{0} \Leftrightarrow \pi_0L_0 + \pi_1B_0 = 0, \\ \pi_0\Lambda_{0,1} + \pi_1\Lambda_0 + \pi_2\Lambda_{-1} &= \mathbf{0} \Leftrightarrow \pi_0F_0 + \pi_1(L + RB) = \mathbf{0}, \end{aligned}$$

δηλαδή μέσω της επίλυσης του συστήματος:

$$\begin{pmatrix} \pi_0 & \pi_1 \end{pmatrix} \begin{pmatrix} L_0 & F_0 \\ B_0 & L + RB \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix},$$

συνδυαστικά με την συνθήκης κανονικοποίησης:

$$\pi_0 + \pi_1(I - R)^{-1}\mathbf{e} = 1.$$

Παράδειγμα Αριθμητικής Εφαρμογής

Προκειμένου να αποδειχθεί η πρακτική εφαρμογή και η υπολογιστική αποδοτικότητα της MGM, γίνεται χρήση του λογισμικού Wolfram Mathematica. Επιλέγονται ενδεικτικές τιμές για τις παραμέτρους του συστήματος M/Cox/1 και ο πλήρης κώδικας αυτής που υλοποιεί όλες τις αναγκαίες υπολογιστικές διαδικασίες, όπως ο υπολογισμός του πίνακα R και των κατανομών ισορροπίας, παρατίθεται στο Παράρτημα Α.

Για την επιλογή των παραμέτρων $\lambda = 1, \mu_1 = 2, \mu_2 = 3$ και $p = 0,4$, εξήχθησαν τα παρακάτω αποτελέσματα:

$$R = \begin{bmatrix} 0,333333 & 0,0666667 \\ 0 & 0,25 \end{bmatrix},$$
$$\pi_0 = 0,511848,$$
$$\pi_1 = (0,236967 \quad 0,0758294).$$

Κεφάλαιο 2^ο: Στοιχεία Θεωρίας Παιγνίων και Στρατηγική Συμπεριφορά στις Ουρές Αναμονής

Το παρόν κεφάλαιο παρουσιάζει τις βασικές αρχές της θεωρίας παιγνίων και της στρατηγικής συμπεριφοράς. Το υλικό και οι έννοιες που αναπτύσσονται στην πορεία βασίζονται στο σύγγραμμα *To Queue Or Not To Queue: Equilibrium Behavior In Queueing Systems* των Rafael Hassin και Moshe Haviv, καθώς και στο σύγγραμμα *An Introduction To Game Theory* του Martin J. Osborne. Ο σκοπός αυτού του κεφαλαίου είναι να λειτουργήσει ως το απαραίτητο θεωρητικό υπόβαθρο, παρέχοντας τη βάση πάνω στην οποία θα στηριχθεί η ανάλυση των επόμενων κεφαλαίων. Έρχεται έτσι να συμπληρώσει το προηγούμενο κεφάλαιο, εξηγώντας πώς οι πελάτες λαμβάνουν αποφάσεις, όταν βρίσκονται σε ουρά αναμονής και πώς αυτή η συμπεριφορά επηρεάζει τελικά την απόδοση του συστήματος.

2.1 Εισαγωγή στην Θεωρία Παιγνίων

Η Θεωρία Παιγνίων αποσκοπεί στο να μας βοηθήσει να κατανοήσουμε καταστάσεις στις οποίες τα άτομα που λαμβάνουν αποφάσεις αλληλεπιδρούν μεταξύ τους. Αν και η λέξη «παίγνιο» ίσως παραπέμπει σε παιχνίδι διασκέδασης, η πραγματικότητα είναι διαφορετική, καθώς το πεδίο αυτό εστιάζει στην ανάλυση οικονομικών, κοινωνικών, τεχνολογικών αλλά και άλλων φαινομένων.

Οι εφαρμογές της εκτείνονται σε ένα ευρύ φάσμα καταστάσεων όπου οι αποφάσεις του καθενός επηρεάζουν τους γύρω τους. Τις συναντάμε παντού στην καθημερινότητα: από την αγορά προϊόντων και υπηρεσιών, όπως καταστήματα που ανταγωνίζονται για πελάτες ή ουρές εξυπηρέτησης όπου ο καθένας ζυγίζει την αναμονή του, μέχρι την κίνηση στους δρόμους, όπου οδηγοί επιλέγουν διαδρομές για να αποφύγουν την κίνηση ή ψάχνουν θέση σε ένα γεμάτο πάρκινγκ. Ακόμη, εμφανίζονται σε πιο σύνθετους τομείς, όπως εταιρείες τηλεπικοινωνιών που καθορίζουν τα τιμολόγια τους με βάση τον ανταγωνισμό, πλατφόρμες διαδικτύου που διαχειρίζονται τη διαφημιστική τους προβολή ή ακόμα και πολιτικούς που διεκδικούν ψήφους. Σε όλες αυτές τις περιπτώσεις, αλλά και σε πληθώρα άλλων, ο καθένας προσπαθεί να κάνει την καλύτερη δυνατή κίνηση υπολογίζοντας τις κινήσεις των άλλων.

Όπως συμβαίνει σε κάθε επιστήμη, έτσι και η θεωρία παιγνίων βασίζεται στη χρήση μοντέλων. Ένα μοντέλο είναι ουσιαστικά μία απλοποιημένη αναπαράσταση της πραγματικότητας που μας βοηθάει να κατανοήσουμε πώς λειτουργούν διάφορα φαινόμενα, εστιάζοντας στις βασικές αρχές ενός προβλήματος.

Τα μοντέλα αυτά δεν αξιολογούνται με βάση κάποιο απόλυτο κριτήριο, ούτε χαρακτηρίζονται απλώς ως «σωστά» ή «λάθος», αλλά με το αν είναι χρήσιμα ανάλογα με τον σκοπό που τα

χρησιμοποιούμε. Κάθε φορά επιλέγουμε την κατάλληλη προσέγγιση που εξυπηρετεί καλύτερα αυτό που επιθυμούμε να μελετήσουμε.

Τελικά, ο λόγος που μελετάμε αυτά τα μοντέλα δεν είναι μόνο για να καταλάβουμε τον κόσμο γύρω μας, αλλά για να εξετάσουμε δυνατές επιλογές και να καταλήγουμε στις καλύτερες δυνατές αποφάσεις. Η γνώση αυτή μας βοηθάει να βελτιώνουμε την καθημερινότητά μας και να διορθώνουμε τον τρόπο που αντιδρούμε στις διάφορες καταστάσεις με τις οποίες ερχόμαστε αντιμέτωποι. Ειδικά σε οικονομικά και κοινωνικά ζητήματα, η ανάλυση των κινήτρων και των στρατηγικών μας δείχνει ποια κίνηση συμφέρει περισσότερο, επιτρέποντάς μας να ζυγίζουμε τα υπέρ και τα κατά πριν αποφασίσουμε.

Η θεωρία της ορθολογικής επιλογής (rational choice theory) αποτελεί βασικό θεμέλιο για πολλά παιγνοθεωρητικά μοντέλα και ορίζει ότι ένα άτομο επιλέγει πάντα την καλύτερη δυνατή ενέργεια σύμφωνα με τις προτιμήσεις του, ανάμεσα σε όσα έχει διαθέσιμα. Αξίζει να σημειωθεί ότι δεν μπαίνει κανένας περιορισμός στο τι του αρέσει ή τι προτιμάει. Η λεγόμενη «ορθολογικότητα» (rationality) δεν αφορά το περιεχόμενο των επιθυμιών του, αλλά αποκλειστικά το πόσο συνεπής είναι στις αποφάσεις του, όταν αλλάζουν κάθε φορά οι διαθέσιμες επιλογές του.

Για να μπορέσουμε να αναλύσουμε το πώς λαμβάνεται μία απόφαση, η θεωρία βασίζεται σε δύο κύρια στοιχεία: στις διαθέσιμες επιλογές, δηλαδή στο σύνολο των ενεργειών που μπορεί γενικά να πραγματοποιήσει κάποιος σε μία κατάσταση, και στις προτιμήσεις του, δηλαδή στην προσωπική του κρίση και τις προτεραιότητες που έχει θέσει.

Στην πράξη, όταν φτάνει η στιγμή της απόφασης, το άτομο δεν έχει στη διάθεση του ολόκληρο το εύρος των πιθανών επιλογών, αλλά περιορίζεται σε ένα συγκεκριμένο υποσύνολο δυνατοτήτων. Για παράδειγμα, ενώ θεωρητικά θα μπορούσε να εξετάσει πολλές διαφορετικές κινήσεις, οι κανόνες του γενικότερου πλαισίου μίας κατάστασης ή κάποιοι περιορισμοί του περιβάλλοντος τον αναγκάζουν να διαλέξει αποκλειστικά ανάμεσα σε συγκεκριμένες προκαθορισμένες επιλογές.

Σχετικά με τις προτιμήσεις, θεωρούμε δεδομένο ότι όταν δίνεται σε ένα άτομο ένα ζεύγος ενεργειών, το ίδιο είναι σε θέση να ξεχωρίζει ποια από τις δύο προτιμά ή να αναγνωρίζει ότι του είναι εξίσου αποδεκτές, άρα αδιάφορες. Παράλληλα, οι επιλογές αυτές διέπονται από σταθερότητα και μεταβατικότητα: αν η ενέργεια a προτιμάται έναντι της b , και η b έναντι της c , τότε η a προτιμάται απαραίτητα έναντι της c . Πέρα από αυτόν τον κανόνα, δεν μπαίνουν άλλοι περιορισμοί. Βασικός σκοπός των θεωριών ορθολογικής επιλογής είναι να καταλήγουν σε συμπεράσματα που δεν εξαρτώνται από τα ιδιαίτερα ποιοτικά χαρακτηριστικά των προτιμήσεων.

Για την αποτύπωση των προτιμήσεων ενός ατόμου, ακολουθούνται κυρίως δύο οδοί:

- Είτε εξετάζεται ξεχωριστά κάθε πιθανό ζεύγος ενεργειών, καταγράφοντας ποια ενέργεια προτιμάται ή αν υπάρχει αδιαφορία σε σχέση με τη σειρά προτίμησης.
- Είτε αποτυπώνονται οι προτιμήσεις με μία συνάρτηση αμοιβής, γνωστή και ως συνάρτηση χρησιμότητας, η οποία αποδίδει έναν αριθμό σε κάθε ενέργεια, φροντίζοντας οι ενέργειες με υψηλότερους αριθμούς να είναι αυτές που προτιμώνται.

Πιο συγκεκριμένα, μία συνάρτηση αμοιβής u εκφράζει τις προτιμήσεις ενός ατόμου όταν, για οποιοσδήποτε ενέργειες a και b στο σύνολο A , ισχύει:

$$u(a) > u(b), \text{ αν και μόνο αν το άτομο προτιμά το } a \text{ από το } b.$$

Στο σημείο αυτό, να διευκρινίσουμε ότι οι προτιμήσεις περιέχουν καθαρά διατακτική πληροφορία. Μας δείχνουν ποια ενέργεια προτιμάται έναντι μίας άλλης, χωρίς να ποσοτικοποιούν την απόσταση ή την ένταση της προτίμησης. Συνεπώς, μία συνάρτηση αμοιβής μεταφέρει αποκλειστικά πληροφορία διάταξης και όχι έντασης, πράγμα που σημαίνει ότι οι προτιμήσεις ενός ατόμου αναπαριστώνται εξίσου καλά από οποιαδήποτε άλλη γνησίως αύξουσα συνάρτηση.

2.2 Στρατηγικές και Ισορροπία Nash στα Συστήματα Ουρών

Στο πλαίσιο των συστημάτων ουρών αναμονής, η θεωρία παιγνίων αξιοποιείται για να περιγράψει πώς οι αυτόνομοι και ορθολογικοί πελάτες λαμβάνουν αποκλίνουσες αποφάσεις. Αξίζει να σημειωθεί ότι, στο σύνολο της ανάλυσης του παρόντος κεφαλαίου, ο όρος «ισορροπία» ταυτίζεται με την έννοια της «ισορροπίας Nash». Δεδομένου ότι η απόδοση του συστήματος διέπεται από Μαρκοβιανές Διαδικασίες, η συμπεριφορά κάθε χρήστη επηρεάζει, αλλά και επηρεάζεται, από τη συνολική κατάσταση του συστήματος. Όταν όλοι οι πελάτες υιοθετούν μία κοινή στρατηγική (όπως η απόφαση εισόδου, βάσει του μήκους της ουράς ή του αριθμού προτεραιότητας), το σύστημα τείνει σε μία κατάσταση ισορροπίας, στην οποία κανένας μεμονωμένος πελάτης δεν έχει κίνητρο να αποκλίνει μονομερώς. Η αξιολόγηση αυτής της ισορροπίας βασίζεται στις στάσιμες πιθανότητες, επιτρέποντας μας να συνδέσουμε άμεσα τις ατομικές στρατηγικές επιλογές με τα μέτρα απόδοσης της ουράς.

Πιο συγκεκριμένα, ένα μη συνεργατικό παίγνιο, σε αυτό το πλαίσιο, ορίζεται από ένα πεπερασμένο σύνολο παικτών $N = \{1, 2, \dots, n\}$, οι οποίοι αντιστοιχούν στους πελάτες του συστήματος, και από ένα σύνολο διαθέσιμων ενεργειών A_i , για κάθε παίκτη $i \in N$, ενέργειες που αφορούν για παράδειγμα την απόφαση εισόδου ή αποχώρησης από την ουρά. Μια καθαρή στρατηγική (pure strategy) για έναν πελάτη αντιστοιχεί σε μία συγκεκριμένη, καθορισμένη ενέργεια από το σύνολο A_i , δηλαδή μία ξεκάθαρη απόφαση, όπως να μείνει ή όχι στην ουρά. Αντίθετα, μία μικτή στρατηγική (mixed strategy) περιγράφει μία κατανομή πιθανότητας επί του συνόλου των ενεργειών, που σημαίνει ότι ο πελάτης παίρνει την απόφασή του με βάση κάποια πιθανότητα, επιλέγοντας τυχαία ανάμεσα στις διαθέσιμες ενέργειες, σύμφωνα με ορισμένα ποσοστά.

Το σύνολο των στρατηγικών που έχει στη διάθεσή του ο παίκτης i συμβολίζεται με S_i . Για να μπορέσουμε να μελετήσουμε τη συνολική εικόνα ορίζουμε το προφίλ στρατηγικών $s = (s_1, \dots, s_n)$, το οποίο συγκεντρώνει όλες τις επιλογές ταυτόχρονα. Επιπλέον, κάθε άτομο συνδέεται με μία συνάρτηση αμοιβής $F_i(s)$. Η συνάρτηση αυτή εκφράζει το καθαρό όφελος που αποκομίζει ο

πελάτης, λαμβάνοντας υπόψη τόσο τη δική του ενέργεια, όσο και τις στρατηγικές που ακολουθούν οι υπόλοιποι γύρω του. Για να διευκολυνθεί η ανάλυση, συμβολίζουμε με s_{-i} τις στρατηγικές όλων των άλλων πελατών, εκτός του πελάτη i , οπότε και η συνάρτηση αμοιβής γράφεται $F_i(s) = F_i(s_i, s_{-i})$. Η συνάρτηση αυτή θεωρείται γραμμική ως προς τη στρατηγική s_i , που σημαίνει ότι για οποιοδήποτε s_{-i} , ισχύει ότι $F_i(s_i, s_{-i}) = \alpha F_i(s_i^1, s_{-i}) + (1 - \alpha) F_i(s_i^2, s_{-i})$, δείχνοντας ότι το s_i είναι γραμμικός συνδυασμός με πιθανότητες $\alpha, 1 - \alpha$, των s_i^1, s_i^2 , αντίστοιχα.

Αναφορικά με τις επιλογές ενός πελάτη, μία στρατηγική s_i^1 λέγεται ότι κυριαρχεί ασθενώς μίας άλλης στρατηγικής s_i^2 εάν ότι κι αν επιλέξουν οι λοιποί πελάτες, δηλαδή για κάθε τιμή του s_{-i} , το σύστημα του αποφέρει τουλάχιστον το ίδιο ή και μεγαλύτερο όφελος, που σημαίνει $F_i(s_i^1, s_{-i}) \geq F_i(s_i^2, s_{-i})$, με την προϋπόθεση να υπάρχει μία περίπτωση όπου η ανισότητα να είναι αυστηρή. Αν μία στρατηγική κυριαρχεί με αυτόν τον τρόπο έναντι όλων των υπολοίπων διαθέσιμων επιλογών στο S_i , τότε χαρακτηρίζεται ως ασθενώς κυρίαρχη.

Προχωρώντας περαιτέρω, ορίζεται η έννοια της καλύτερης απόκρισης για έναν παίκτη i απέναντι στις επιλογές των υπολοίπων s_{-i} . Πρόκειται για εκείνη τη στρατηγική s_i^* που μεγιστοποιεί τη συνάρτηση αμοιβής του:

$$s_i^* \in \arg \max_{s_i \in S_i} F_i(s_i, s_{-i}).$$

Με βάση αυτό, ένα προφίλ στρατηγικών s^c χαρακτηρίζεται ως ισορροπία, δηλαδή ισορροπία Nash, όταν η στρατηγική κάθε ξεχωριστού πελάτη αποτελεί την καλύτερη δυνατή απόκριση απέναντι στις επιλογές όλων των άλλων:

$$s^c \in \arg \max_{s_i \in S_i} F_i(s_i, s_{-i}^c), \quad i \in N.$$

Να αναφερθεί ότι αν μια καλύτερη απόκριση είναι ένας συνδυασμός στρατηγικών, τότε όλες οι επιμέρους στρατηγικές αυτού του μείγματος αποτελούν εξίσου καλύτερες αποκρίσεις. Ωστόσο, αυτή η ιδιότητα δεν ισχύει γενικά όταν περνάμε στην έννοια της ισορροπίας.

Στα συστήματα ουρών αναμονής, η ανάλυση επικεντρώνεται κυρίως σε παίγνια με άπειρους και μη διακριτούς πελάτες, αφού συνηθίζεται να καταφθάνουν διαρκώς νέοι πελάτες στο σύστημα. Σε αυτή την περίπτωση, όλοι οι πελάτες μοιράζονται το ίδιο σύστημα στρατηγικών S και την ίδια συνάρτηση αμοιβής F . Έτσι, αν συμβολίσουμε με $F(a, b)$ το όφελος ενός πελάτη που ακολουθεί τη στρατηγική a , όταν όλοι οι υπόλοιποι επιλέγουν τη στρατηγική b , τότε μία συμμετρική ισορροπία ορίζεται ως μία στρατηγική $s^c \in S$, τέτοια ώστε:

$$s^c \in \arg \max_{s \in S} F(s, s^c).$$

Που σημαίνει ότι, μία συμμετρική ισορροπία είναι απλώς μία στρατηγική που αποτελεί την καλύτερη δυνατή απόκριση απέναντι στον ίδιο της τον εαυτό, δείχνοντας ότι κανένας πελάτης δεν έχει λόγο να παρεκκλίνει από τη γενική συμπεριφορά του πληθυσμού.

Επιπλέον, αξίζει να σημειωθεί ότι η ύπαρξη μίας τέτοιας ισορροπίας δεν είναι πάντα δεδομένη, καθώς υπάρχουν ειδικά μοντέλα στα οποία δεν εμφανίζεται κατάσταση ισορροπίας.

Συνήθως, τα συστήματα ουρών αναμονής ταξινομούνται ανάλογα με το αν ο πελάτης έχει ορατότητα στο μήκος της παρατηρούμενης ουράς, προτού πάρει την απόφασή του. Ως εκ τούτου, διακρίνουμε τις ορατές ουρές και τις μη ορατές ουρές. Στην πρώτη κατηγορία οι πελάτες έρχονται αντιμέτωποι με καταστάσεις που αντιστοιχούν στην τρέχουσα κατάσταση του συστήματος και καλούνται να επιλέξουν μία ενέργεια από ένα προκαθορισμένο σύνολο. Οι βασικές έννοιες που ορίσαμε προγενέστερα (ενέργειες, στρατηγικές, αμοιβές, ισορροπία) μπορούν να επεκταθούν και σε μοντέλα που εξαρτώνται από την κατάσταση.

Στην πράξη, μία τέτοια κατάσταση μπορεί να εκφράζεται από τον ακριβή αριθμό των ατόμων που βρίσκονται στο σύστημα, ενώ το σύνολο των ενεργειών μπορεί να περιλαμβάνει τις εξής επιλογές: κανονική είσοδος πελάτη, είσοδος με προτεραιότητα, απόφαση να μην υπάρξει είσοδος. Μία καθαρή στρατηγική ορίζει μία συγκεκριμένη ενέργεια για κάθε ξεχωριστή κατάσταση. Ο συνδυασμός των στρατηγικών όλων των ατόμων, μαζί με την αρχική κατάσταση του συστήματος διαμορφώνει μία κατανομή πιθανότητας. Έτσι, η αμοιβή ενός πελάτη i εξαρτάται από την κατάσταση του συστήματος, τη δική του ενέργεια, αλλά και τις επιλογές των υπολοίπων. Επομένως, ο πελάτης δεν γνωρίζει εκ των προτέρων το ακριβές αποτέλεσμα της επιλογής του, γι' αυτό και λαμβάνει τις αποφάσεις του με βάση το αναμενόμενο όφελος. Ο υπολογισμός αυτός συνυπολογίζει τόσο τις πιθανότητες εμφάνισης των διαφόρων καταστάσεων στο σύστημα, όσο και τις ενέργειες που υπαγορεύει η στρατηγική του για καθεμία από αυτές τις καταστάσεις.

2.3 Μονοτονία των Συναρτήσεων Βέλτιστης Απόκρισης και Μορφές Ισορροπίας

Σε πολλά μοντέλα ουρών αναμονής, οι επιλεγόμενες στρατηγικές μπορούν να αποτυπωθούν μέσω μίας μοναδικής αριθμητικής τιμής. Σε αυτές τις περιπτώσεις ιδιαίτερο ενδιαφέρον παρουσιάζει η διερεύνηση του κατά πόσο η βέλτιστη αντίδραση ενός μεμονωμένου ατόμου αποτελεί μία μονότονα αύξουσα ή φθίνουσα συνάρτηση της στρατηγικής που επέλεξαν οι υπόλοιποι πελάτες.

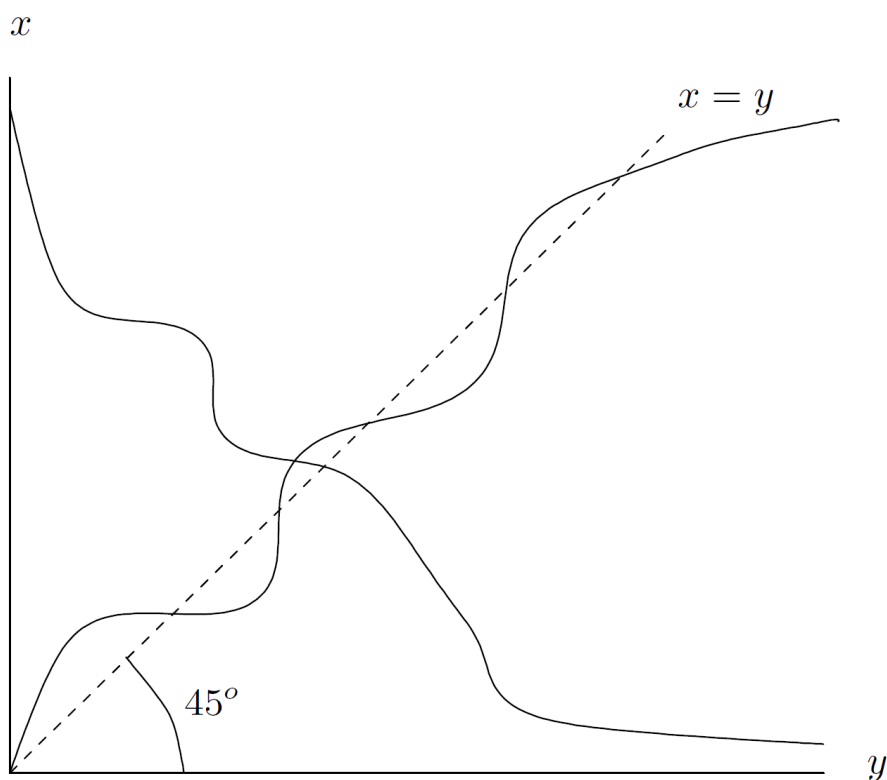
Έστω $F(x, y)$ το κέρδος για έναν πελάτη ο οποίος επιλέγει τη στρατηγική x , όταν όλοι οι υπόλοιποι επιλέγουν τη στρατηγική y . Θεωρούμε ότι για κάθε y υπάρχει μία μοναδική βέλτιστη αντίδραση $x(y)$, τέτοια ώστε να ισχύει:

$$x(y) = \arg \max_x F(x, y).$$

Η ανάλυση εστιάζει σε περιπτώσεις όπου η συνάρτηση βέλτιστης απόκρισης $x(y)$ είναι συνεχής και αυστηρά μονότονη. Στο ακόλουθο σχήμα απεικονίζεται μία κατάσταση στην οποία μία στρατηγική αντιστοιχεί σε ένα μη αρνητικό αριθμό και παρουσιάζονται δύο περιπτώσεις για την βέλτιστη αντίδραση $x(y)$, που η κάθε μία από αυτές αντιπροσωπεύεται και από μία καμπύλη. Στην μία καμπύλη η $x(y)$ είναι μονότονα φθίνουσα και η κατάσταση αυτή καλείται «αποφυγή του πλήθους» (avoid the crowd – ATC), ενώ, η άλλη καμπύλη είναι μονότονα αύξουσα και η κατάσταση αυτή καλείται «προσχώρηση στο πλήθος» (follow the crowd - FTC).

Μάλιστα, η σύνδεση της ονομασίας αυτών των καταστάσεων προέρχεται από την χαρακτηριστική ιδιότητα ότι, σε μία περίπτωση ATC, όσο μεγαλύτερες είναι οι τιμές της στρατηγικής του πλήθους, τόσο μικρότερη είναι η τιμή της βέλτιστης απόκρισης του ατόμου, άρα απομακρύνεται η προτίμηση του ενός από αυτή των πολλών. Κατ' αντιστοιχία, σε μία περίπτωση FTC, όσο το πλήθος μεγαλώνει το όριο εισόδου, τόσο η τιμή της βέλτιστης απόκρισης του ατόμου αυξάνεται, δείχνοντας ότι το άτομο είναι προτιμότερο να ακολουθεί την επιλογή των πολλών.

Σαφέστατα, είναι σημαντικό να διατυπωθεί ότι ο χαρακτηρισμός του προς μελέτη μοντέλου ως ATC ή FTC είναι καθοριστικός. Αυτό οφείλεται στο ότι, μια στρατηγική ισορροπίας y ικανοποιεί την σχέση $x(y) = y$, που σημαίνει ότι πρόκειται για ένα σταθερό σημείο της συνάρτησης x . Έτσι, εάν το μοντέλο ανήκει στην περίπτωση ATC, τότε υπάρχει το πολύ μία ισορροπία, ενώ, στην περίπτωση FTC είναι πιθανή η ύπαρξη πολλαπλών ισορροπιών.



Σχήμα 2.1: Καμπύλες βέλτιστης απόκρισης και σημεία ισορροπίας για τα μοντέλα ATC και FTC

Θέλοντας να καταγράψουμε τι παρατηρούμε στο παραπάνω Σχήμα 2.1 και να κάνουμε μία πιο προσεκτική ανάγνωση των ιδιοτήτων των διαφορετικών περιπτώσεων, θα λέγαμε ότι τα σημεία στα οποία οι καμπύλες τέμνουν την διακεκομμένη γραμμή $x = y$, η οποία αποτελεί ευθεία διχοτόμο του τεταρτημορίου, αντιπροσωπεύουν τις καταστάσεις ισορροπίας. Σε αυτά τα σημεία, η βέλτιστη επιλογή του ατόμου ταιριάζει ακριβώς με αυτό που κάνουν όλοι οι υπόλοιποι.

Έπειτα, η φθίνουσα καμπύλη αντιστοιχεί στην περίπτωση ATC και ξεκινάει ψηλά στον άξονα, κατεβαίνοντας προς τα κάτω και δεξιά. Εδώ, καθώς η στρατηγική των άλλων y αυξάνεται, η βέλτιστη αντίδραση του ατόμου x μειώνεται. Όπως φαίνεται χαρακτηριστικά στο διάγραμμα, η συγκεκριμένη καμπύλη τέμνει την $x = y$ σε ένα ακριβώς σημείο, γεγονός που επιβεβαιώνει ότι στην περίπτωση ATC υπάρχει μία το πολύ ισορροπία.

Αντιθέτως, η αύξουσα καμπύλη απεικονίζει την περίπτωση FTC ξεκινώντας από χαμηλά στον άξονα και ανεβαίνοντας προς τα πάνω και δεξιά. Εδώ, καθώς η στρατηγική των άλλων y αυξάνεται, η βέλτιστη αντίδραση του ατόμου x αυξάνεται και αυτή. Λόγω του χαρακτηριστικού σχήματος που λαμβάνει η συγκεκριμένη καμπύλη, είναι δυνατόν να διασταυρωθεί με την $x = y$ σε περισσότερα από ένα σημεία, αποδεικνύοντας ότι στην περίπτωση FTC ευνοείται η ύπαρξη πολλαπλών ισορροπιών.

2.4 Στρατηγικές Κατωφλιού

Στην ενότητα αυτή περιγράφεται μία κατηγορία στρατηγικών στα συστήματα ουρών αναμονής, γνωστές ως «στρατηγικές κατωφλιού» (threshold strategies). Θεωρούμε ότι κατά την άφιξη του ο πελάτης καλείται να επιλέξει ανάμεσα σε δύο διαθέσιμες ενέργειες, A_1 και A_2 , όπως για παράδειγμα η απόφαση συμμετοχής (join) ή μη στο σύστημα (balk). Η απόφαση αυτή λαμβάνεται αφότου ο πελάτης παρατηρήσει μία μη αρνητική ακέραη μεταβλητή που αποτυπώνει την τρέχουσα κατάσταση του συστήματος, όπως για παράδειγμα το ακριβές μήκος της ουράς ή ο αριθμός προτεραιότητας που αναγράφεται σε ένα εκδοθέν εισιτήριο. Με άλλα λόγια, ο πελάτης, φτάνοντας στον χώρο, αξιολογεί άμεσα το πόσο φορτωμένο είναι το σύστημα, μετατρέποντας την εικόνα που αντικρίζει σε ένα συγκεκριμένο ακέραιο αριθμό των ατόμων που αναμένουν.

Μία «καθαρή» στρατηγική κατωφλιού με οριακή τιμή n ορίζει την επιλογή μίας εκ των δύο ενεργειών, για κάθε κατάσταση στο σύνολο $\{0, 1, \dots, n-1\}$, και την επιλογή της εναλλακτικής ενέργειας για κάθε κατάσταση που υπερβαίνει ή ισούται με αυτό το κρίσιμο όριο, κατώφλι.

Συχνά, η αναζήτηση μίας στρατηγικής κατωφλιού ισορροπίας επικεντρώνεται στις καθαρές στρατηγικές κατωφλιού. Ωστόσο, είναι δυνατό να προκύψουν περιπτώσεις όπου μία καθαρή στρατηγική κατωφλιού, που να ορίζει ισορροπία, ενδέχεται να μην υπάρχει. Ενδεχομένως, μπορεί να παρατηρηθεί το φαινόμενο όπου, εάν όλος ο πληθυσμός υιοθετήσει ένα συγκεκριμένο κατώφλι, η

βέλτιστη απόκριση για ένα μεμονωμένο πελάτη να είναι διαφορετική, και αντίστροφα. Αυτό συμβαίνει διότι η συλλογική συμπεριφορά του πληθυσμού έρχεται σε αντίθεση με την ατομική βέλτιστη απόκριση. Για παράδειγμα, μπορεί να παρατηρηθεί ότι αν όλοι υιοθετήσουν το κατώφλι με τιμή 4, η βέλτιστη απόκριση για έναν ορισμένο πελάτη είναι το όριο 5, ενώ, αν όλοι υιοθετήσουν την τιμή 5, η βέλτιστη απόκριση είναι το 4. Κατά συνέπεια, ο ορισμός της στρατηγικής κατωφλιού επεκτείνεται προκειμένου να καλύψει αυτές τις περιπτώσεις.

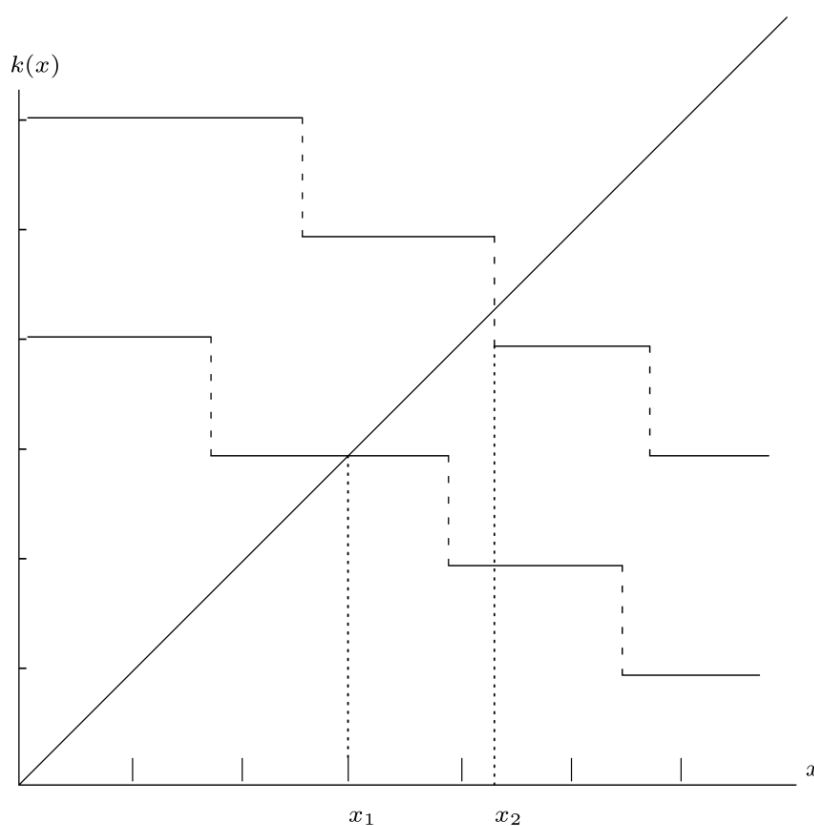
Έτσι, μια στρατηγική κατωφλιού με κατώφλι $x = n + p$, όπου $n \in \mathbb{N}$ και $p \in [0, 1]$, ορίζει έναν συνδυασμό μεταξύ δύο διαδοχικών καθαρών στρατηγικών κατωφλιού, με τιμές n και $n + 1$, αποδίδοντας βάρος $1 - p$ στην πρώτη και p στην δεύτερη. Η διαδικασία που συμβαίνει στο σύστημα περιγράφεται από τον ακόλουθο συλλογισμό. Όλοι οι πελάτες επιλέγουν μία δεδομένη ενέργεια, έστω A_1 , όταν η κατάσταση του συστήματος παίρνει τιμές από 0 έως και $n - 1$. Όταν η κατάσταση ισούται ακριβώς με n , οι πελάτες δρουν τυχαioκρατικά, επιλέγοντας την ενέργεια A_1 με πιθανότητα p (την ενέργεια που αντιστοιχεί στην στρατηγική $n + 1$) και την εναλλακτική ενέργεια A_2 , με πιθανότητα $1 - p$ (την ενέργεια που αντιστοιχεί στην στρατηγική n). Ενώ, για τιμές μεγαλύτερες του ορίου n οι πελάτες οδηγούνται στη επιλογή της ενέργειας A_2 . Στην ιδιαίτερη περίπτωση όπου το x είναι ακέραιος αριθμός ($p = 0$), η στρατηγική χαρακτηρίζεται ως «καθαρή», ενώ, σε κάθε άλλη περίπτωση θεωρείται «μικτή».

Εξετάζονται μοντέλα στα οποία η βέλτιστη απόκριση ενός μεμονωμένου πελάτη απέναντι σε μία στρατηγική x λαμβάνει τη μορφή καθαρού κατωφλιού, δηλαδή, για κάποιον ακέραιο $k(x)$, οι πελάτες επιλέγουν την ενέργεια A_1 , όταν η κατάσταση της ουράς παίρνει τιμές στο $\{0, \dots, k(x) - 1\}$, και την ενέργεια A_2 , διαφορετικά. Η συνάρτηση βέλτιστης απόκρισης $k(x)$ χαρακτηρίζεται τυπικά από σημεία ασυνέχειας με άλματα μοναδιαίου μεγέθους, τα οποία μπορεί να είναι είτε ανοδικά είτε καθοδικά. Σε κάθε σημείο ασυνέχειας x , οι δύο εμπλεκόμενες καθαρές στρατηγικές συνιστούν εξίσου βέλτιστες αποκρίσεις, με αποτέλεσμα οποιαδήποτε μικτή στρατηγική που προκύπτει από την ανάμειξή τους να αποτελεί επίσης βέλτιστη απόκριση. Κατά συνέπεια, μία στρατηγική κατωφλιού x ορίζει μία στρατηγική ισορροπίας όταν αποτελεί λύση της εξίσωσης $k(x) = x$ ή όταν βρίσκεται ενδιάμεσα στις τιμές $k(x-)$ και $k(x+)$. Γενικότερα, είναι χρήσιμο να θεωρούμε ότι και οι δύο περιπτώσεις ικανοποιούν την εξίσωση $k(x) = x$, λειτουργώντας ουσιαστικά ως σταθερά σημεία της συνάρτησης $k(x)$. Υπό αυτές τις συνθήκες, εάν όλοι οι πελάτες υιοθετήσουν το κατώφλι x , η στρατηγική αυτή παραμένει η βέλτιστη επιλογή για καθέναν ξεχωριστά. Επομένως, κανένας δεν έχει κίνητρο να αλλάξει την συμπεριφορά του, επιβεβαιώνοντας πράγματι ότι πρόκειται για μία στρατηγική ισορροπίας.

Τώρα, όπως αναλύθηκε και στην προηγούμενη ενότητα και ως σύντομη υπενθύμιση και σε αυτό το σημείο, θα λέγαμε ότι η συμπεριφορά που αποτυπώνεται από μία μονότονη μη αύξουσα συνάρτηση βέλτιστης απόκρισης $k(x)$ χαρακτηρίζεται ως «αποφυγή του πλήθους» (avoid the crowd - ATC). Σύμφωνα με αυτήν, όσο υψηλότερο είναι το κατώφλι που υιοθετείται από το υπόλοιπο πλήθος, τόσο

χαμηλότερο είναι το κατώφλι που συνιστά βέλτιστη απόκριση για έναν μεμονωμένο πελάτη. Αντίστοιχα, η περίπτωση κατά την οποία η συνάρτηση $k(x)$ είναι μονότονη μη φθίνουσα χαρακτηρίζεται ως «προσχώρηση στο πλήθος» (follow the crowd - FTC), γεγονός που σημαίνει ότι υψηλότερα κατώφλια από την πλευρά του πληθυσμού συνεπάγονται αντίστοιχα υψηλότερα βέλτιστα κατώφλια για τον πελάτη.

Συνεπώς, μεταξύ των δύο περιπτώσεων, ναι μεν διατηρούνται τα ίδια ποιοτικά συμπεράσματα, εντούτοις, υφίστανται ορισμένες σημαντικές διαφορές. Στην περίπτωση του ATC, υπάρχει το πολύ ένα σταθερό σημείο, το οποίο μπορεί να αντιστοιχεί είτε σε καθαρή είτε σε μικτή στρατηγική. Αντίθετα, η περίπτωση του FTC παρουσιάζει μεγαλύτερη πολυπλοκότητα και ενδέχεται να επιδέχεται πολλαπλές ισορροπίες, καθώς η συνάρτηση μπορεί να διαθέτει πολυάριθμα σταθερά σημεία.



Σχήμα 2.2: Καθαρή και μικτή στρατηγική ισορροπίας σε μοντέλο ATC

Τώρα, παρατηρώντας το παραπάνω σχήμα διακρίνουμε δύο μη αύξουσες συναρτήσεις περίπτωσης ATC. Στις καμπύλες αυτές, τα οριζόντια τμήματα αντιστοιχούν σε ακέραιες τιμές του κατωφλιού, όπου η στρατηγική είναι καθαρή. Αντίθετα, οι θέσεις (άλματα) που απεικονίζονται ως διακεκομμένα κατακόρυφα τμήματα και μοιάζουν με σκαλοπάτια αποτελούν τα σημεία ασυνέχειας της γραφικής. Αυτά τα άλματα αντιπροσωπεύουν τη μετάβαση μεταξύ διαφορετικών επιπέδων. Πιο συγκεκριμένα, στη μία περίπτωση η στρατηγική ισορροπίας που προκύπτει στο σημείο x_1 είναι καθαρή, αφού το

σημείο τομής της συνάρτησης με την διχοτόμο ευθεία βρίσκεται επάνω σε οριζόντιο τμήμα της γραφικής, λαμβάνοντας ακέραια τιμή κατωφλιού, ενώ στην άλλη περίπτωση η στρατηγική ισορροπίας στο σημείο x_2 είναι μικτή.

Τέλος, τα δεδομένα χαρακτηρίζονται ως «μη εκφυλισμένα» (non-degenerate) εάν κανένα από τα σημεία άλματος της συνάρτησης $k(x)$ δεν συμπίπτει με ακέραιη τιμή του x . Έστω $x_1, x_2, \dots$ οι τιμές που αντιστοιχούν στα σταθερά σημεία. Για $k=1, 2, \dots$, το x_{2k+1} αντιστοιχεί σε καθαρή στρατηγική ισορροπίας, ενώ το x_{2k} αντιστοιχεί σε μικτή στρατηγική ισορροπίας. Στην περίπτωση που επιτρέπονται εκφυλισμένα δεδομένα, ενδέχεται να εμφανιστούν διαδοχικές καθαρές στρατηγικές ισορροπίες. Συμπερασματικά, εάν η ισορροπία του συστήματος είναι μοναδική, τότε αυτή είναι αναγκαστικά καθαρή.

Κεφάλαιο 3^ο: Μοντελοποίηση και Ανάλυση Ουράς Εισιτηρίων με Πελάτες που Αποχωρούν

Το παρόν κεφάλαιο εστιάζει στην αναλυτική περιγραφή και τον μαθηματικό προσδιορισμό ενός συστήματος εξυπηρέτησης με έκδοση εισιτηρίων και συμπεριφορά αποχώρησης (balking). Η ανάλυση βασίζεται στο έργο *Service Performance Analysis and Improvement for a Ticket Queue with Balking Customers* των Susan H. Xu, Long Gao και Jihong Ou. Η πολυπλοκότητα του προβλήματος έγκειται στην ανάγκη ταυτόχρονης παρακολούθησης τόσο του αριθμού των φυσικών πελατών όσο και της θέσης που αναγράφεται στα εισιτήρια, στοιχεία που συχνά αποκλίνουν λόγω των αποχωρήσεων.

3.1 Περιγραφή και Λειτουργία του Συστήματος Εισιτηρίων – Ανάλυση των Ρυθμών Μετάβασης

Το σύστημα λειτουργεί με βάση έναν μηχανισμό ψηφιακής αναγγελίας, όπου κάθε νέος πελάτης κατά την άφιξή του λαμβάνει έναν μοναδικό αύξοντα αριθμό προτεραιότητας. Η επικοινωνία του συστήματος με τους πελάτες γίνεται μέσω ενός κεντρικού πίνακα, ο οποίος προβάλλει τον αριθμό που εξυπηρετείται την δεδομένη στιγμή. Όταν η εξυπηρέτηση ενός ατόμου ολοκληρωθεί, το σύστημα καλεί τον επόμενο διαθέσιμο αριθμό. Εάν δεν υπάρχουν άλλοι πελάτες στην αναμονή, ο πίνακας εμφανίζει τον επόμενο αριθμό που πρόκειται να εκδοθεί, ώστε ο μελλοντικός πελάτης που θα εισέλθει στο σύστημα να εξυπηρετηθεί αμέσως.

Μια κρίσιμη παραδοχή του μοντέλου είναι ότι ενώ ο πάροχος της υπηρεσίας γνωρίζει ανά πάσα στιγμή ποιο είναι το τελευταίο εισιτήριο που εκδόθηκε, εντούτοις, ούτε αυτός ούτε και οι πελάτες γνωρίζουν ποιοι από τους κατόχους εισιτηρίων παραμένουν ενεργοί στην αναμονή. Η πραγματική εικόνα αποκαλύπτεται μόνο την στιγμή που καλείται ένας αριθμός και διαπιστώνεται εάν ο κάτοχός του είναι παρών ή έχει εγκαταλείψει το σύστημα.

Το κεντρικό ζήτημα, λοιπόν, από την σκοπιά του πελάτη, είναι η λήψη απόφασης για το αν θα παραμείνει ή θα αποχωρήσει (balk), βασιζόμενος αποκλειστικά στη διαφορά μεταξύ του αριθμού προτεραιότητας που έλαβε και της τρέχουσας ένδειξης του πίνακα. Σαφώς, η απόφαση αυτή λαμβάνεται υπό συνθήκες αβεβαιότητας, αφού ο κάθε πελάτης δεν γνωρίζει πόσοι από τους προπορευόμενους αριθμούς αντιστοιχούν σε άτομα που βρίσκονται ακόμη στην ουρά και σε άτομα που έχουν ήδη εγκαταλείψει το σύστημα.

Για την μαθηματική τυποποίηση του μοντέλου, θεωρούμε τις εξής παραδοχές:

- Θεωρούμε D την διαφορά μεταξύ αριθμού εισιτηρίου και ένδειξης του πίνακα, γνωστή και ως «θέση εισιτηρίου» (ticket position).
- Θεωρούμε N το πλήθος των πελατών που βρίσκονται στο σύστημα, γνωστό και ως «θέση στην ουρά» (queueing position).

- Οι πελάτες καταφθάνουν σύμφωνα με διαδικασία Poisson με ρυθμό αφίξεων λ .
- Ο χρόνος εξυπηρέτησης για τους πελάτες που παραμένουν στην ουρά ακολουθεί εκθετική κατανομή παραμέτρου μ .
- Ο χρόνος εξυπηρέτησης για τους πελάτες που εγκαταλείπουν είναι 0.
- Υπάρχει συγκεκριμένο όριο ανοχής K , το οποίο ορίζει την μέγιστη αριθμητική απόσταση που είναι διατεθειμένος να αποδεχτεί ένας πελάτης μεταξύ του αριθμού εισιτηρίου και της ένδειξης του πίνακα. Συγκεκριμένα, εάν η θέση εισιτηρίου είναι ίση ή μεγαλύτερη από το όριο αυτό, δηλαδή $D \geq K$, ο πελάτης θεωρεί την αναμονή υπερβολική και αποχωρεί αμέσως. Αντίθετα, η είσοδος στο σύστημα πραγματοποιείται μόνο όταν η απόσταση αυτή είναι αυστηρά μικρότερη του K .
- Την απόφαση για εγκατάλειψη του συστήματος, ο κάθε πελάτης την λαμβάνει αποκλειστικά κατά την στιγμή άφιξής του και καθορίζεται από την σταθερά K , όπως προαναφέρθηκε.

Η ιδιαιτερότητα του συστήματος καθιστά τις μεταβλητές N (αριθμός πελατών) και D (θέση εισιτηρίου) ανεπαρκείς για μία πλήρη στοχαστική περιγραφή. Σε μία κλασική ουρά, αν γνωρίζουμε το N έχουμε πλήρη γνώση όλων των στοιχείων που μας ενδιαφέρουν. Στην περίπτωση μας, αν γνωρίζουμε το N , χάνεται η πληροφορία των εισιτηρίων που λείπουν εξαιτίας της αποχώρησης, ενώ, αν γνωρίζουμε το D , χάνεται η πληροφορία των πραγματικών χρόνων εξυπηρέτησης που υπολείπονται, αφού δεν γνωρίζουμε πόσοι από αυτούς που προηγούνται θα παραμείνουν στο σύστημα.

Το ουσιαστικό πρόβλημα που αντιμετωπίζουμε είναι ότι η καθυστέρηση που θα βιώσει ένας πελάτης δεν εξαρτάται από το πόσα εισιτήρια έχουν εκδοθεί, αλλά από το πού ακριβώς βρίσκονται οι πελάτες, ανάμεσα στα εκδοθέντα εισιτήρια.

Για να γεφυρωθεί αυτό το κενό πληροφορίας, το σύστημα μοντελοποιείται μέσω ενός διανύσματος $x = (n_1, n_2, \dots, n_L)$, το οποίο δεν καταγράφει απλώς πλήθη, αλλά την σχετική δομή της ουράς. Στη δομή αυτή ο δείκτης L αντιπροσωπεύει τον πραγματικό αριθμό των πελατών που βρίσκονται στο σύστημα και λαμβάνει τιμές από 1 έως K και κάθε στοιχείο $n_l \geq 1, l = 1, \dots, L$, περιγράφει το πλήθος των εισιτηρίων που έχουν εκδοθεί από τον παρόντα πελάτη στη θέση l , μέχρι και τον αμέσως επόμενο στη θέση $l+1$. Ουσιαστικά, τα n_l λειτουργούν ως δείκτες απόστασης μεταξύ των πραγματικών πελατών του συστήματος. Επιπλέον, ορίζεται η κατάσταση 0, για να περιγράψει την περίπτωση του εντελώς κενού συστήματος.

Το διάνυσμα είναι, λοιπόν, ο πλήρης χάρτης της ουράς, ο οποίος αποκαλύπτει ποιος είναι ακριβώς στην ουρά και πόσοι έχουν αποχωρήσει, μετά από κάθε άτομο που παραμένει.

Ο πλήρης χώρος καταστάσεων του συστήματος συμβολίζεται με $\tilde{J}$ και ορίζεται από το σύνολο:

$$\tilde{J} = \left\{ x = 0 \text{ ή } x = (n_1, n_2, \dots, n_L) : \sum_{l=1}^{L-1} n_l < K, \quad n_l \geq 1, \quad l = 1, \dots, L, \quad L = 1, \dots, K \right\}.$$

Ο περιορισμός $\sum_{l=1}^{L-1} n_l < K$, λειτουργεί ως ένας μηχανισμός ελέγχου για την είσοδο των πελατών στο σύστημα. Ουσιαστικά, περιγράφει την διαδικασία λήψης απόφασης ενός πελάτη κατά την άφιξη του, όπου ο αφιγχείς πελάτης διατηρεί την θέση του στην ουρά μόνο εάν ο αριθμός προτεραιότητας που λαμβάνει δεν ξεπερνά το προκαθορισμένο όριο ανοχής του. Με τον τρόπο αυτό, η συνθήκη διασφαλίζει ότι όλες οι καταστάσεις που περιλαμβάνονται στο σύνολο $\tilde{J}$ αφορούν πελάτες που επέλεξαν να παραμείνουν και να εξυπηρετηθούν, αποκλείοντας περιπτώσεις άμεσης αποχώρησης, λόγω υπερβολικής αναμονής.

Για την πληρέστερη περιγραφή των καταστάσεων, εισάγονται δυο βασικοί δείκτες:

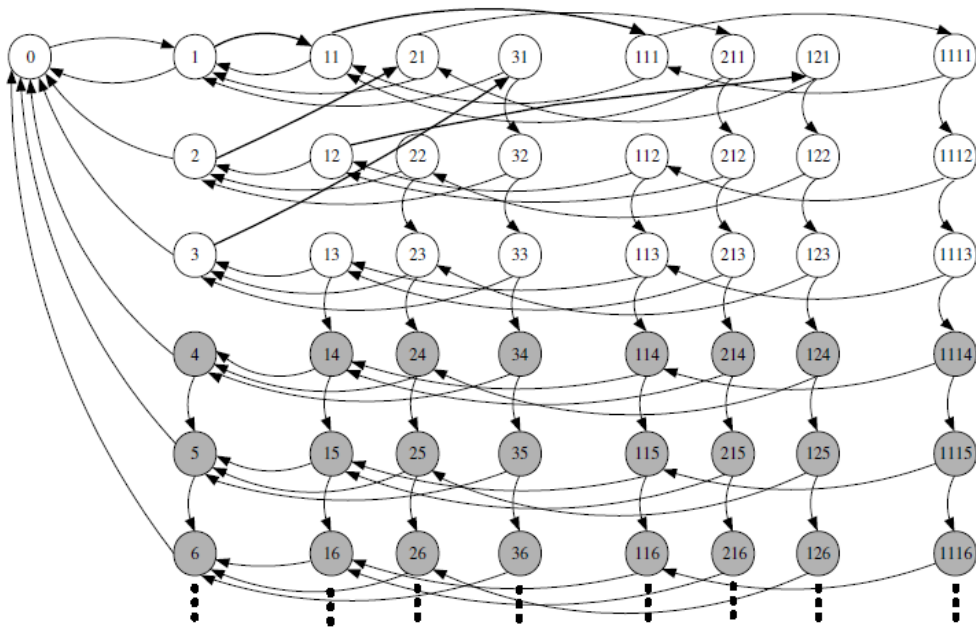
- Η διάσταση της κατάστασης $|x| = L$, που αντιπροσωπεύει το εύρος του διανύσματος, δηλαδή το πλήθος των στοιχείων που συνθέτουν την τρέχουσα κατάσταση.
- Το συνολικό μέγεθος του συστήματος $\|x\| = \sum_{j=1}^L n_j$, που εκφράζει το άθροισμα όλων των πελατών που βρίσκονται στην ουρά την δεδομένη στιγμή.

Η εξέλιξη του συστήματος διέπεται από την Μαρκοβιανή Ιδιότητα, πράγμα που σημαίνει ότι η μετάβαση από μία κατάσταση x σε μία μελλοντική εξαρτάται αποκλειστικά από τις τρέχουσες παραμέτρους και όχι από το ιστορικό των προηγούμενων μεταβολών.

Συγκεκριμένα, η εμφάνιση μίας νέας άφιξης, που πραγματοποιείται με ρυθμό λ , προκαλεί διαφορετική μεταβολή στο διάνυσμα ανάλογα με τον τρέχοντα φόρτο του συστήματος. Εάν το άθροισμα των πελατών είναι μικρότερο του ορίου ανοχής, η νέα άφιξη δημιουργεί μια επιπλέον διάσταση στο σύστημα, οδηγώντας στην μορφή $(n_1, n_2, \dots, n_L, 1)$. Αντιθέτως, εάν το σύστημα έχει ήδη φτάσει ή ξεπεράσει το όριο K , ο νέος πελάτης ενσωματώνεται στην τελευταία υπάρχουσα ομάδα, μεταβάλλοντας την κατάσταση σε $(n_1, n_2, \dots, n_L + 1)$.

Από την άλλη πλευρά, η διαδικασία αποσυμφόρησης του συστήματος πραγματοποιείται μέσω των ολοκληρώσεων εξυπηρέτησης, με ρυθμό μ . Κατά τη στιγμή εξυπηρέτησης, πέραν του πελάτη που αποχωρεί λόγω ολοκλήρωσης της εξυπηρέτησής του, συγχρόνως, συμβαίνει ακαριαία απελευθέρωση όλων των ατόμων που βρίσκονται μεταξύ αυτού και του αμέσως επόμενου πελάτη, που είχε επιλέξει να εισέλθει στο σύστημα. Αυτός ο μηχανισμός εκκαθάρισης έχει ως αποτέλεσμα την πλήρη απομάκρυνση του πρώτου στοιχείου n_1 του διανύσματος. Συνεπώς, η νέα κατάσταση στην οποία οδηγείται το σύστημα, στην περίπτωση αυτή, είναι η $(n_2, n_3, \dots, n_L)$.

Προκειμένου να γίνει κατανοητός ο τρόπος λειτουργίας του συστήματος, ακολουθεί απεικόνιση του διαγράμματος μεταβάσεων του συστήματος, για την περίπτωση όπου το όριο αποχώρησης παίρνει την τιμή $K = 4$.



Σχήμα 3.1: Διάγραμμα μεταβάσεων ουράς εισιτηρίων με $K = 4$

Η περιγραφή ενός συγκεκριμένου στοιχείου του χώρου καταστάσεων είναι πολύ βοηθητική για την αποσαφήνιση των εννοιών του μοντέλου.

Ας υποθέσουμε ότι το σύστημα βρίσκεται στη θέση $(2,1,5)$. Σε αυτή την περίπτωση, η διάσταση είναι $L=3$, άρα $N=3$ άτομα περιμένουν πραγματικά στο σύστημα. Το άθροισμα $\|x\|=2+1+5=8$, το οποίο αντιστοιχεί στη φαινομενική ουρά D , αποτελεί το συνολικό μέγεθος που αντικρίζει ο επόμενος αφικνούμενος πελάτης και περιλαμβάνει τόσο αυτούς που έμειναν, όσο και αυτούς που έφυγαν. Σύμφωνα με τον ορισμό του διανύσματος x , κάθε στοιχείο αντιπροσωπεύει το πλήθος των εισιτηρίων που εκδόθηκαν για την ομάδα που ξεκινά από l -οστό πελάτη. Εφόσον ο ίδιος ο πελάτης που παρέμεινε καταλαμβάνει το 1^ο εισιτήριο της ομάδας του, το πλήθος των αποχωρούντων μετά από αυτόν είναι $n_l - 1$. Συγκεκριμένα, το $n_1 = 2$ υποδηλώνει ότι μετά τον 1^ο πελάτη που παρέμεινε, υπήρξε ένα άτομο που αποχώρησε πριν από την άφιξη του 2^{ου} πραγματικού πελάτη. Το $n_2 = 1$ υποδηλώνει ότι κανένας πελάτης δεν αποχώρησε μεταξύ του 2^{ου} και του 3^{ου} πραγματικού πελάτη, αφού η τιμή 1 αντιστοιχεί, μόνο στον ίδιο τον πραγματικό πελάτη. Τέλος, το $n_3 = 5$ υποδηλώνει ότι μετά τον 3^ο πραγματικό πελάτη, υπήρξαν 4 άτομα που αποχώρησαν.

Στο σημείο αυτό, είναι σημαντικό να παρατηρήσουμε ότι στην κατάσταση $(2,1,5)$, το σύστημα, ως προς το όριό του, βρίσκεται σε φάση υπερπλήρωσης. Εφόσον το συνολικό μέγεθος $\|x\|=8$ είναι ήδη μεγαλύτερο από το επιτρεπτό όριο $K=4$, το σύστημα δεν μπορεί να αυξήσει περαιτέρω την διάστασή του με νέες αφίξεις. Αυτό σημαίνει ότι οποιοσδήποτε νέος πελάτης αφιχθεί, παρατηρώντας την κατάσταση αυτή, θα αποχωρήσει αναγκαστικά (balking). Οπότε, ο επόμενος πελάτης ενσωματώνεται στην τελευταία ομάδα n_3 , προκαλώντας την αύξησή της, χωρίς όμως να παραμένει ο ίδιος στο σύστημα. Για τον λόγο αυτό, το n_3 μπορεί να μεγαλώσει απεριόριστα, απορροφώντας

όλους τους επόμενους αφικνούμενους πελάτες, όπως προκύπτει και από τις τελίτσες του σχήματος. Η μόνη περίπτωση να ξεμπλοκάρει το σύστημα και να αλλάξει αυτή η διάταξη είναι να ολοκληρωθεί μια εξυπηρέτηση, που θα συμβεί με ρυθμό μ .

Πριν προχωρήσουμε στον υπολογισμό των πιθανοτήτων, είναι σημαντικό να επισημανθεί η ευστάθεια του συστήματος. Η Μαρκοβιανή Αλυσίδα που περιγράφει το σύστημα παρουσιάζει οριακή κατανομή, υπό την προϋπόθεση ότι η τιμή του ορίου ανοχής K παραμένει πεπερασμένη και ο ρυθμός εξυπηρέτησης μ υπερβαίνει τον ρυθμό αφίξεων λ . Υπό αυτές τις συνθήκες, το σύστημα δεν δύναται να αυξάνεται απεριόριστα, καθώς το φαινόμενο των αποχωρήσεων (balking) περιορίζει τον αριθμό των πελατών που παραμένουν, εξασφαλίζοντας έτσι ότι το σύστημα καταλήγει πάντα σε μία κατάσταση ισορροπίας.

Για τη μελέτη του συστήματος και τον προσδιορισμό της στάσιμης κατανομής του για γενικές τιμές του K , εφαρμόζεται μια τεχνική δύο σταδίων:

Στάδιο 1: Όλες οι καταστάσεις $x \in \tilde{J}$, για τις οποίες ισχύει ότι $|x| = L$ και των οποίων το τελευταίο στοιχείο n_L είναι ίσο ή μεγαλύτερο του K , ομαδοποιούνται σε μία υπερ-κατάσταση S_L , όπου $L = 1, 2, \dots, K$. Πλέον, η Μαρκοβιανή Αλυσίδα των υπερ-καταστάσεων διαθέτει πεπερασμένο πλήθος καταστάσεων. Αυτή η δομή επιτρέπει τη μοντελοποίηση του συστήματος ως μια Διαδικασία QBD, την οποία και μπορούμε να επιλύσουμε.

Στάδιο 2: Στη συνέχεια, αναστρέφεται η διαδικασία της ομαδοποίησης, ώστε να υπολογιστούν αναδρομικά οι στάσιμες κατανομές κάθε μεμονωμένης κατάστασης, χρησιμοποιώντας τα αποτελέσματα του προηγούμενου σταδίου.

Τα στάδια αυτά πρόκειται να αναλυθούν διεξοδικά στις επόμενες παραγράφους του παρόντος κεφαλαίου.

3.2 Ομαδοποίηση Καταστάσεων και Υπολογισμός της Στάσιμης Κατανομής τους

Η πολυπλοκότητα του χώρου καταστάσεων $\tilde{J}$, που αποκαλύπτεται και από το Σχήμα 3.1, καθιστά την εύρεση μιας αναλυτικής λύσης για τη στάσιμη κατανομή του συστήματος εξαιρετικά δυσχερή, ιδιαίτερα για μεγάλες τιμές του K . Για τον λόγο αυτόν, επιλέγεται η δημιουργία μίας υπερ-κατάστασης S_L , υπό την προϋπόθεση ικανοποίησης κάποιων προδιαγραφών. Η S_L δεν είναι μια απλή κατάσταση, αλλά ένα σύνολο από πολλές διαφορετικές καταστάσεις, που ομαδοποιούνται επειδή συμπεριφέρονται με τον ίδιο ακριβώς τρόπο όσον αφορά τις αφίξεις. Πιο συγκεκριμένα, μέσα στην S_L βρίσκονται όλες οι καταστάσεις που έχουν τον ίδιο αριθμό πραγματικών πελατών ($|x| = L$) και, συγχρόνως, οδηγούνται σε εγκατάλειψη του συστήματος ($n_L \geq K$). Αυτό σημαίνει ότι όταν το σύστημα βρίσκεται σε οποιαδήποτε κατάσταση που ανήκει στην S_L , ο επόμενος πελάτης με

βεβαιότητα θα αποχωρήσει, καθώς ο αριθμός του εισιτηρίου του δεν είναι μικρότερος από K , ενώ το μήκος της πραγματικής ουράς δεν θα μεγαλώνει, καθώς κάθε νέα άφιξη θα απορρίπτεται, διατηρώντας το πλήθος των πραγματικών πελατών L σταθερά.

Έτσι, ορίζουμε την υπέρ-κατάσταση ως:

$$S_L = \{x \in \tilde{J} : |x| = L, n_L \geq K\}, \quad L = 1, 2, \dots, K-1, \quad (3.1)$$

$$S_K = \{x = (e_{K-1}, n_K), n_K \geq K\}, \quad (3.2)$$

όπου $e_L = (1, 1, \dots, 1)$ είναι το L -διάστατο μοναδιαίο διάνυσμα, συνεπώς $(e_{K-1}, n_K) = (1, 1, \dots, 1, n_K)$.

Η ανάγκη για διαχωρισμό των σχέσεων (3.1) και (3.2) πηγάζει από τη δομική διαφορά των υπερ-καταστάσεων. Στη σχέση (3.1), για $L < K$, η υπερ-κατάσταση S_L περιλαμβάνει ένα σύνολο συνδυασμών, όπου οι ενδιάμεσες καταστάσεις n_l μπορούν να ποικίλλουν, επιτρέποντας την ύπαρξη κενών εισιτηρίων από πελάτες που αποχώρησαν ανάμεσα στους παρόντες πελάτες. Αντίθετα, η σχέση (3.2) περιγράφει την κατάσταση όπου η ουρά έχει φτάσει στην μέγιστη χωρητικότητά της $L = K$. Σε αυτή την περίπτωση, για να παραμείνει ο K -οστός πελάτης, δεν πρέπει να μεσολαβήσει κανένα κενό εισιτήριο μεταξύ των προηγούμενων πελατών. Ως εκ τούτου, οι συνθήκες του συστήματος επιβάλλουν τη μοναδική δυνατή κατάσταση όπου όλοι οι προπορευόμενοι πελάτες βρίσκονται σε διαδοχικές θέσεις εισιτηρίων ($n_1 = n_2 = \dots = n_{K-1} = 1$), γεγονός το οποίο αποτυπώνεται μέσω του διανύσματος μονάδων e_{K-1} .

Η βαθύτερη σκέψη πίσω από την επιλογή των υπερ-καταστάσεων απορρέει από την προσπάθεια αποφυγής των άπειρων καταστάσεων της αλυσίδας. Αν παρατηρήσουμε το Σχήμα 3.1 της προηγούμενης παραγράφου, θα διαπιστώσουμε ότι υπάρχουν άπειρες καταστάσεις που η πελάτης εγκαταλείπουν την ουρά. Αντί, λοιπόν, το μοντέλο να διατηρήσει άπειρους κύκλους μεταβάσεων, επιλέγεται η συγχώνευση όλων των καταστάσεων που έχουν L άτομα και δεν δέχονται νέους πελάτες, σε ένα ενιαίο σύνολο.

Με την εφαρμογή αυτής της στρατηγικής, ο αρχικά άπειρος χώρος $\tilde{J}$, αναδιαμορφώνεται στο πεπερασμένο σύνολο J . Στη νέα αυτή μορφή, οι επιμέρους καταστάσεις περιορίζονται από τη συνθήκη $n_L < K$, ενώ όλες οι περιπτώσεις όπου $n_L \geq K$ απορροφώνται από τις αντίστοιχες υπερ-καταστάσεις.

Έτσι, ο χώρος καταστάσεων αναδιαμορφώνεται στην εξής απλούστερη μορφή:

$$J = \{0\} \cup \left\{ x : \sum_{l=1}^{L-1} n_l < K, \quad n_L < K, \quad n_l \geq 1, \quad l = 1, \dots, L, \quad L = 1, \dots, K \right\} \cup \{S_1, S_2, \dots, S_K\}.$$

Για την ανάλυση του συστήματος ως μια Διαδικασία QBD, ο χώρος καταστάσεων J διαμερίζεται σε $K+1$ ξένα μεταξύ τους υποσύνολα $\{E_0, E_1, \dots, E_K\}$.

Το πρώτο υποσύνολο $E_0 = \{0\}$ αντιπροσωπεύει το κενό σύστημα, ενώ για κάθε $L=1,2,...,K$, το αντίστοιχο υποσύνολο E_L , ορίζεται ως:

$$E_L = \{x \in J : |x| = L, \quad n_L < K\} \cup \{S_L\}, \quad L=1,2,...,K. \quad (3.3)$$

Τώρα, η δυναμική του συστήματος επιτρέπει μεταβάσεις αποκλειστικά μεταξύ γειτονικών επιπέδων. Συγκεκριμένα, η διαμέριση αυτή εισάγει τη δομή επιπέδου και φάσης, που χαρακτηρίζει τις QBD διεργασίες. Κάθε σύνολο E_L αντιστοιχεί σε ένα επίπεδο της αλυσίδας και προσδιορίζεται από τον αριθμό L των πραγματικά παρόντων πελατών που περιμένουν στην ουρά τη δεδομένη στιγμή και λαμβάνει τιμές από 0 έως K .

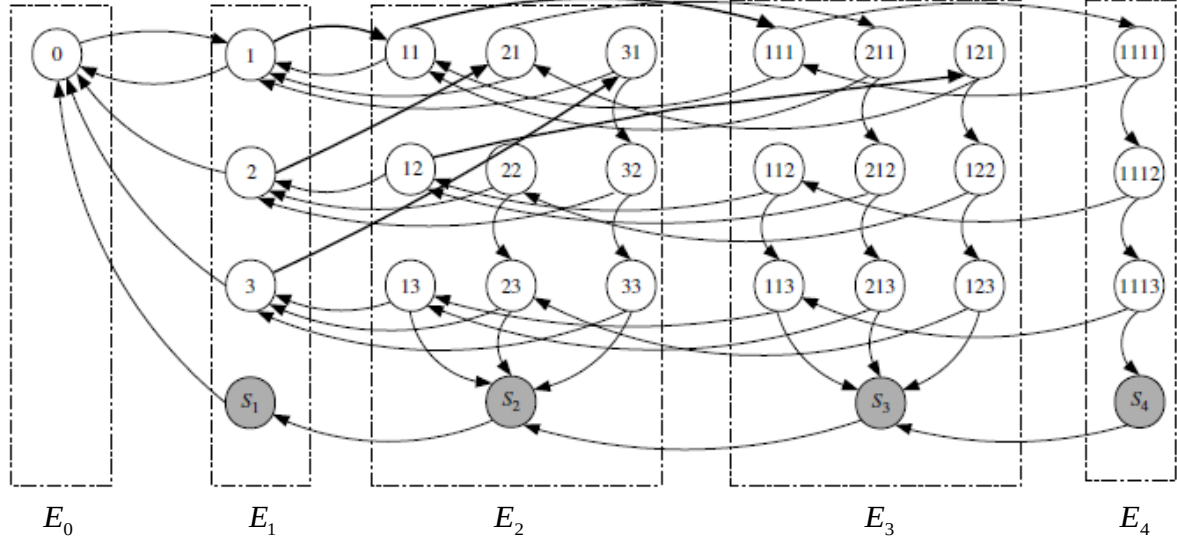
Στο πλαίσιο αυτό, η μετάβαση μεταξύ των επιπέδων καθορίζεται από τις αφίξεις και τις εξυπηρετήσεις. Η άφιξη προκαλεί μετάβαση στο επόμενο επίπεδο ($L \rightarrow L+1$), εφόσον ο πελάτης εισέλθει στην ουρά (επιτυχής άφιξη πελάτη), ενώ αν ο πελάτης αποχωρήσει αμέσως, λόγω εγκατάλειψης του συστήματος, το σύστημα παραμένει στο επίπεδο L . Αντιστοίχως, η ολοκλήρωση εξυπηρέτησης οδηγεί το σύστημα στο προηγούμενο επίπεδο ($L \rightarrow L-1$).

Παράλληλα, μέσα σε κάθε επίπεδο E_L , οι καταστάσεις περιγράφουν τη συγκεκριμένη διάταξη των εισιτηρίων, δηλαδή τις αποστάσεις n_l , μεταξύ των παρόντων πελατών, που αποτελούν τις φάσεις. Έτσι, οι μεταβάσεις εντός του ίδιου επιπέδου ($L \rightarrow L$) υποδηλώνουν αλλαγές στη φάση του συστήματος, διατηρώντας σταθερό τον αριθμό των πελατών. Όταν το σύστημα βρίσκεται σε ένα επίπεδο L , η φάση καθορίζεται από τις αποστάσεις μεταξύ των πελατών που αναμένουν.

Προκειμένου να υπάρχει μία σαφώς ορισμένη εσωτερική δομή σε κάθε επίπεδο, οι καταστάσεις (φάσεις) εντός κάθε συνόλου E_L ταξινομούνται με λεξικογραφική σειρά. Σύμφωνα με αυτή τη διάταξη, μία κατάσταση $x = (n_1, n_2, n_3, ..., n_L)$ προηγείται μίας άλλης $x' = (n'_1, n'_2, n'_3, ..., n'_L)$ εάν η θέση του τελευταίου πελάτη στο σύστημα είναι μικρότερη, $n_L < n'_L$. Στην περίπτωση που $n_L = n'_L$, η προτεραιότητα καθορίζεται από το άθροισμα των θέσεων όλων των προηγούμενων. Έτσι, το x θα προηγείται εφόσον $\sum_{l=1}^{L-1} n_l \leq \sum_{l=1}^{L-1} n'_l$, για όλα τα $1 \leq l \leq L-1$. Η επιλογή του αθροίσματος $\sum_{l=1}^{L-1} n_l$, για $l=1,2,...,L-1$, εξυπηρετεί τον συστηματικό έλεγχο των θέσεων των πελατών, ξεκινώντας από τον τελευταίο και προχωρώντας σταδιακά προς την αρχή της ουράς, όπου ο δείκτης μετατοπίζει το σημείο αναφοράς του αθροίσματος.

Σε συνέχεια του Σχήματος 3.1, της προηγούμενης παραγράφου, το ακόλουθο Σχήμα 3.2 απεικονίζει πλέον τη δομή της Ομαδοποιημένης Μαρκοβιανής Αλυσίδας, για την περίπτωση $K=4$. Στο διάγραμμα αυτό, που αποτελεί ένα στιγμιότυπο της γενικής δομής, φαίνεται ο διαχωρισμός του χώρου καταστάσεων J σε $K+1$ ξένα μεταξύ τους επίπεδα E_L , $L=1,2,...,K$, τα οποία οριοθετούνται από τις διακεκομμένες γραμμές. Κάθε επίπεδο περιλαμβάνει τις αντίστοιχες φάσεις

και την υπερ-κατάσταση S_L , η οποία λειτουργεί ως κατάσταση απορρόφησης των άπειρων καταστάσεων του αρχικού μοντέλου.



Σχήμα 3.2: Διάγραμμα μεταβάσεων με υπερ-καταστάσεις για $K = 4$

Για την περίπτωση $K = 4$, η διαμέριση του χώρου J στα επιμέρους E_L , οργανωμένα βάσει της λεξικογραφικής διάταξης, διαμορφώνεται ως:

$$\begin{aligned} E_0 &= \{0\}, \\ E_1 &= \{1, 2, 3, S_1\}, \\ E_2 &= \{11, 21, 31, 12, 22, 32, 13, 23, 33, S_2\}, \\ E_3 &= \{111, 211, 121, 112, 212, 122, 113, 213, 123, S_3\}, \\ E_4 &= \{1111, 1112, 1113, S_4\}. \end{aligned}$$

Στην παραπάνω διάταξη, η υπερ-κατάσταση S_L τοποθετείται τελευταία σε κάθε επίπεδο, καθώς ορίζει το σημείο πέρα από το οποίο η θέση του τελευταίου πελάτη υπερβαίνει το K .

Θεωρώντας, πλέον, τα σύνολα E_L ως block-στοιχεία του συστήματος, παρατηρούμε ότι κάθε μετάβαση στην αλυσίδα περιορίζεται αυστηρά μεταξύ των γειτονικών επιπέδων. Επισημαίνουμε, για ακόμα μία φορά, ότι το σύστημα μπορεί είτε να παραμείνει στο τρέχον επίπεδο, είτε να μεταβεί στο αμέσως επόμενο $L+1$ λόγω άφιξης, είτε στο προηγούμενο $L-1$ λόγω εξυπηρέτησης.

Αυτή ακριβώς η ιδιότητα της περιορισμένης μετάβασης μεταξύ των γειτονικών επιπέδων προσδίδει στο σύστημα τη δομή μίας QBD Διαδικασίας. Κατά συνέπεια, ο αντίστοιχος πίνακας ρυθμών μετάβασης Q λαμβάνει μια block-τριδιαγώνια μορφή, όπου οι διαστάσεις των υποπινάκων του καθορίζονται από την πληθικότητα $|E_L|$ κάθε επιπέδου.

Συγκεκριμένα, ο γεννήτορας πίνακας Q διαμορφώνεται ως εξής:

$$Q = \begin{pmatrix} -\lambda & A_{0,1} & & & & & \\ \mu \mathbf{e}'_K & A_{1,1} & A_{1,2} & & & & \\ & A_{2,1} & A_{2,2} & A_{2,3} & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & & A_{K-1,K-2} & A_{K-1,K-1} & A_{K-1,K} \\ & & & & & A_{K,K-1} & A_{K,K} \end{pmatrix}, \quad (3.4)$$

όπου $\mathbf{e}'_K$ είναι το ανάστροφο διάνυσμα του $\mathbf{e}_K$. Επιπλέον, οι $A_{L+1,L}, L=1,2,\dots,K-1$ και $A_{L,L}, A_{L-1,L}, L=1,2,\dots,K$, είναι πίνακες διαστάσεων $|E_{L+1}| \times |E_L|$ και $|E_L| \times |E_L|, |E_{L-1}| \times |E_L|$, αντίστοιχα.

Γενικά, οι υποπίνακες $A_{i,j}$ περιγράφουν τους ρυθμούς μετάβασης από το επίπεδο i στο επίπεδο j . Ειδικότερα, οι πίνακες $A_{L-1,L}$ αντιστοιχούν σε αφίξεις πελατών, οι πίνακες $A_{L+1,L}$ σε εξυπηρετήσεις, ενώ, οι πίνακες $A_{L,L}$, επί της κύριας διαγωνίου, περιλαμβάνουν τις μεταβάσεις μεταξύ των φάσεων του ίδιου επιπέδου. Τέλος, το πρώτο block $(-\lambda)$ αφορά το άδαιο σύστημα E_0 .

Είναι σημαντικό να επισημανθεί η μεταβλητή διάσταση των υποπινάκων του γεννήτορα Q . Σε μία τυπική QBD, όλοι οι πίνακες στο επαναλαμβανόμενο τμήμα του Q είναι πανομοιότυποι, συνεπώς έχουν την ίδια διάσταση. Εδώ, επειδή το πλήθος των πιθανών συνδυασμών αυξάνεται όσο αυξάνεται ο αριθμός των πελατών, οι διαστάσεις των υποπινάκων μεγαλώνουν, καθώς προχωράμε κάτω και δεξιά, στον πίνακα Q . Έτσι, ο πίνακας Q τερματίζει στο K και δεν συνεχίζεται επ' άπειρον προς τα κάτω με την ίδια μορφή.

Έστω $\mathbf{p}_L$ η στάσιμη κατανομή για τις καταστάσεις που ανήκουν στο E_L . Θεωρούμε $\mathbf{p}_0 = p(0)$, πιθανότητα το σύστημα να είναι άδαιο, και $\mathbf{p}_L$ διάνυσμα γραμμής, με διάσταση $|E_L|$. Συνολικά, η πλήρης κατανομή των πιθανοτήτων για κάθε κατάσταση $\{p(x), x \in J\}$ εκφράζεται μέσω του διανύσματος $\mathbf{p} = (\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_K)$.

Οι πιθανότητες αυτές προκύπτουν από την επίλυση των εξισώσεων ισορροπίας:

$$\mathbf{p}Q = \mathbf{0},$$

έτσι, οδηγούμαστε σε:

$$-\lambda \mathbf{p}_0 + \mu \mathbf{p}_1 \mathbf{e}'_K = 0, \quad (3.5)$$

$$\mathbf{p}_{L-1} A_{L-1,L} + \mathbf{p}_L A_{L,L} + \mathbf{p}_{L+1} A_{L+1,L} = 0, \quad L=1,2,\dots,K-1, \quad (3.6)$$

$$\mathbf{p}_{K-1} A_{K-1,K} + \mathbf{p}_K A_{K,K} = 0. \quad (3.7)$$

Η (3.6) καταλήγει σε ένα σύστημα $|E_L|$ εξισώσεων, κάθε μία από τις οποίες σχετίζεται με το αντίστοιχο E_L .

Επιλύοντας, αναδρομικά, έχουμε:

$$\mathbf{p}_K = -\mathbf{p}_{K-1}A_{K-1,K} \left(A_{K,K} \right)^{-1} = \mathbf{p}_{K-1}R_{K-1,K}, \quad (3.8)$$

$$\mathbf{p}_L = \mathbf{p}_{L-1}R_{L-1,L}, \quad L=1,2,\dots,K-1, \quad (3.9)$$

όπου ο πίνακας $R_{L-1,L}$, με διαστάσεις $|E_{L-1}| \times |E_L|$, υπολογίζεται σύμφωνα με:

$$R_{K-1,K} = -A_{K-1,K} \left(A_{K,K} \right)^{-1}, \quad (3.10)$$

$$R_{L-1,L} = -A_{L-1,L} \left(A_{L,L} + R_{L,L+1}A_{L+1,L} \right)^{-1}, \quad L=1,2,\dots,K-1. \quad (3.11)$$

Να σημειώσουμε ότι η (3.8) προκύπτει από την (3.7) και η τελευταία της ισότητα αξιοποιεί τον $R_{K-1,K}$ από του τύπου (3.10). Όσο για την (3.9), η μετάβαση σε αυτή έχει αφετηρία την (3.6). Όμως, απαιτούνται ελαφρώς περισσότερες διαδικασίες.

Λεπτομερέστερα, για τον καθορισμό του $\mathbf{p}_{L+1}$ επαναλαμβάνουμε δύο αντικαταστάσεις της (3.9) και προκύπτει:

$$\mathbf{p}_{L+1} = \mathbf{p}_L R_{L,L+1} = \mathbf{p}_{L-1} R_{L-1,L} R_{L,L+1}.$$

Για την απόδειξη του τύπου (3.10), που αφορά το οριακό επίπεδο, έχουμε την ακόλουθη διαδικασία:

Από την εξίσωση (3.7) $\mathbf{p}_{K-1}A_{K-1,K} + \mathbf{p}_K A_{K,K} = 0$ αντικαθιστώντας την (3.8) $\mathbf{p}_K = \mathbf{p}_{K-1}R_{K-1,K}$, προκύπτει:

$$\mathbf{p}_{K-1}A_{K-1,K} + (\mathbf{p}_{K-1}R_{K-1,K})A_{K,K} = 0 \Leftrightarrow \mathbf{p}_{K-1}(A_{K-1,K} + R_{K-1,K}A_{K,K}) = 0.$$

Επειδή αυτό πρέπει να ισχύει για κάθε $\mathbf{p}_{K-1}$, τότε υποχρεωτικά πρέπει:

$$A_{K-1,K} + R_{K-1,K}A_{K,K} = 0.$$

Τώρα με επίλυση αυτής ως προς $R_{K-1,K}$ προκύπτει άμεσα η (3.10).

Έπειτα, για την απόδειξη του τύπου (3.11), που αφορά κάθε ενδιάμεσο επίπεδο, έχουμε την ακόλουθη διαδικασία:

Από την (3.6), αντικαθιστώντας την (3.9) $\mathbf{p}_L = \mathbf{p}_{L-1}R_{L-1,L}$, μαζί με $\mathbf{p}_{L+1} = \mathbf{p}_{L-1}R_{L-1,L}R_{L,L+1}$, προκύπτει:

$$\mathbf{p}_{L-1}A_{L-1,L} + \mathbf{p}_{L-1}R_{L-1,L}A_{L,L} + \mathbf{p}_{L-1}R_{L-1,L}R_{L,L+1}A_{L+1,L} = 0.$$

Παραγοντοποιούμε το $\mathbf{p}_{L-1}$ και για να ισχύει για κάθε περίπτωση, πρέπει:

$$\mathbf{p}_{L-1} (A_{L-1,L} + R_{L-1,L} A_{L,L} + R_{L-1,L} R_{L,L+1} A_{L+1,L}) = 0 \Leftrightarrow A_{L-1,L} + R_{L-1,L} A_{L,L} + R_{L-1,L} R_{L,L+1} A_{L+1,L} = 0.$$

Λύνουμε ως προς το $R_{L-1,L}$ και άμεσα καταλήγουμε στην (3.11).

Τώρα, με διαδοχική εφαρμογή της (3.9), για τις διάφορες τιμές του L , μπορούμε να οδηγηθούμε σε μία απλούστερη μορφή για την $\mathbf{p}_L$. Συγκεκριμένα, με αφετηρία το επίπεδο $L=1$, η (3.9), δίνει $\mathbf{p}_1 = \mathbf{p}_0 R_{0,1}$. Εν συνεχεία, για το επίπεδο $L=2$, η (3.9) μας δίνει ότι $\mathbf{p}_2 = \mathbf{p}_1 R_{1,2}$ και εάν αντικαταστήσουμε το $\mathbf{p}_1$ από το προηγούμενο βήμα, έχουμε $\mathbf{p}_2 = (\mathbf{p}_0 R_{0,1}) R_{1,2} = \mathbf{p}_0 (R_{0,1} R_{1,2})$. Συνεχίζοντας, με όμοιο αλυσιδωτό τρόπο, για κάθε επίπεδο L , το $\mathbf{p}_L$ λαμβάνει την μορφή του γινομένου:

$$\mathbf{p}_L = \mathbf{p}_0 (R_{0,1} \cdot R_{1,2} \cdot \dots \cdot R_{L-1,L}).$$

Όμως, χάριν απλότητας, ορίζουμε τον πίνακα R_L ως το γινόμενο όλων των επιμέρους πινάκων $R_{l-1,l}, l=1,2,\dots,L$:

$$R_L = \prod_{l=1}^L R_{l-1,l}.$$

Μέσω του ορισμού αυτού, η $\mathbf{p}_L$, για κάθε επίπεδο L , μπορεί πλέον να εκφραστεί ως:

$$\mathbf{p}_L = \mathbf{p}_0 R_L, \quad L=1,2,\dots,K, \quad (3.12)$$

όπου $\mathbf{p}_0 = p(0)$ αριθμός.

Η (3.12) αναδεικνύει την πινακογεωμετρική μορφή της λύσης, σε μια διαφοροποιημένη εκδοχή από την κλασική μορφή που μελετήθηκε στο Κεφάλαιο 1, της παρούσας εργασίας. Εδώ, ο σταθερός πίνακας R έχει αντικατασταθεί από τον R_L , εξαρτώμενο πίνακα από το επίπεδο L . Ο πίνακας αυτός ενσωματώνει όλη την απαραίτητη πληροφορία των μεταβάσεων, επιτρέποντας την επιτυχή εφαρμογή της MGM.

Έπειτα, χρησιμοποιώντας την συνθήκη κανονικοποίησης $\mathbf{p} \mathbf{e}'_{|J|} = 1$, όπου $|J|$ το πλήθος των καταστάσεων του συνόλου J , διασφαλίζεται ότι το άθροισμα των πιθανοτήτων όλων των καταστάσεων του χώρου ισούται με τη μονάδα. Η συνθήκη αυτή λαμβάνει τη μορφή:

$$\mathbf{p}_0 \left(1 + \sum_{L=1}^K R_L \mathbf{e}'_{|E_L|} \right) = 1.$$

Το οποίο, λύνοντάς το ως προς p_0 , δίνει ισοδύναμα:

$$p_0 = \frac{1}{1 + \sum_{L=1}^K R_L e^{|E_L|}}. \quad (3.13)$$

Τέλος, αντικαθιστώντας την (3.13) στην (3.12), λαμβάνουμε την τελική έκφραση για τις πιθανότητες κάθε επιπέδου L :

$$p_L = \frac{R_L}{1 + \sum_{L=1}^K R_L e^{|E_L|}}, \quad L=1, 2, \dots, K. \quad (3.14)$$

Η σχέση (3.14) επιτρέπει τον πλήρη προσδιορισμό της στάσιμης κατανομής του συστήματος, χρησιμοποιώντας αποκλειστικά τους αναδρομικά υπολογισμένους πίνακες R_L .

Ολοκληρώνοντας την υποενότητα αυτή, είναι σημαντικό να επισημανθούν ορισμένες παρατηρήσεις σχετικά με την φύση της λύσης και την υπολογιστική της προσέγγιση. Αρχικά, σημειώνεται ότι η Ομαδοποιημένη Μαρκοβιανή Αλυσίδα που μελετήθηκε, περιέχει την πλήρη πληροφορία για τον αριθμό των πελατών N στο σύστημα, καθώς η πιθανότητα $P(N=L), L=1, 2, \dots, K$, προκύπτει από το άθροισμα $\sum_{x \in E_L} p(x)$. Επιπλέον, από τη στάσιμη κατανομή $\{p(x), x \in J\}$ μπορούν να εξαχθούν και άλλα κρίσιμα μεγέθη, όπως η πιθανότητα άρνησης εισόδου των πελατών.

Ωστόσο, η τρέχουσα ανάλυση δεν προσφέρει επακριβώς την περιθώρια κατανομή της φάσης D , ούτε την από κοινού κατανομή (N, D) , καθώς και οι δύο εξαρτώνται από τις επιμέρους πιθανότητες των καταστάσεων που είναι «κρυμμένες» μέσα στις υπερ-καταστάσεις. Ο υπολογισμός αυτών των μεμονωμένων πιθανοτήτων, εντός των υπερ-καταστάσεων, αποτελεί αντικείμενο του δεύτερου σταδίου της διαδικασίας μας, το οποίο εξετάζεται στην επόμενη ενότητα.

Επιπλέον, αξίζει να αναφερθεί ότι αν και υπάρχουν εναλλακτικοί μέθοδοι κατάτμησης του χώρου καταστάσεων, η επιλογή της παρούσας δομής, οδηγεί σε μια Διαδικασία QBD. Η προσέγγιση αυτή προτιμάται καθώς προσφέρει μια πιο άμεση και υπολογιστικά αποτελεσματική αναδρομική διαδικασία για τον προσδιορισμό της στάσιμης κατανομής του συστήματος.

3.3 Προσδιορισμός των Στάσιμων Πιθανοτήτων Εντός των Υπερ-Καταστάσεων

Έχοντας προσδιορίσει στην προηγούμενη ενότητα τις συνολικές πιθανότητες των υπερ-καταστάσεων S_L , το επόμενο κρίσιμο βήμα στην διαδικασία επίλυσης είναι η αποδόμηση αυτών των πιθανοτήτων στις επιμέρους καταστάσεις που τις απαρτίζουν. Συνεπώς, στόχος είναι να

εξεταστεί το εσωτερικό των ήδη διαμορφωμένων αυτών ομάδων, ώστε να προσδιοριστεί η στάσιμη πιθανότητα κάθε μεμονωμένης κατάστασης x . Για να επιτευχθεί αυτό, πραγματοποιούμε μια νέα αναδιανομή του χώρου καταστάσεων $\tilde{J}$ σε ξένα μεταξύ τους υποσύνολα $\{T_n, n \geq 0\}$, όπου $T_0 = \{0\}$ αντιστοιχεί στην κενή κατάσταση και κάθε σύνολο T_n , για $n = 1, 2, \dots$, ορίζεται ως:

$$T_n = \{x \in \tilde{J} : n_L = n, \quad L = 1, 2, \dots, K\}, \quad n = 1, 2, \dots$$

Μπορούμε να καταλάβουμε ότι το T_n ομαδοποιεί όλες εκείνες τις καταστάσεις x όπου η τιμή του τελευταίου ψηφίου του διανύσματος, το οποίο αντιστοιχεί στη θέση του τελευταίου εισιτηρίου, είναι ακριβώς n . Έτσι, εάν μια κατάσταση τελειώνει σε 1 (για παράδειγμα η 1, η 11, η 111, κ.τ.λ.) ανήκει στο T_1 , εάν μια κατάσταση τελειώνει σε 2 (για παράδειγμα η 2, η 12, η 22, κ.τ.λ.) ανήκει στο T_2 , και ούτω καθεξής.

Η δομή αυτής της διαμέρισης γίνεται άμεσα αντιληπτή μέσω του Σχήματος 3.1, όπου το T_n αντιστοιχεί στην n -οστή σειρά του διαγράμματος μεταβάσεων. Ενώ η προηγούμενη διαμέριση σε υπερ-καταστάσεις S_L λειτουργούσε κάθετα, ομαδοποιώντας καταστάσεις με βάση τον αριθμό L των πελατών που βρίσκονται μέσα στο σύστημα και λαμβάνοντας υπόψη την συμπεριφορά αποχώρησης, όπως φαίνεται στις διαμορφωμένες στήλες-ομάδες E_L του Σχήματος 3.2, η νέα διαμέριση T_n λειτουργεί οριζόντια, ανά επίπεδο τελευταίου ψηφίου.

Να σημειωθεί ότι, εντός κάθε συνόλου T_n , οι καταστάσεις διατάσσονται με λεξικογραφική σειρά, δηλαδή μία κατάσταση x προηγείται μίας άλλης x' εάν έχει μικρότερο αριθμό πελατών, $|x| < |x'|$. Ενώ, στην περίπτωση που ο αριθμός των πελατών είναι ίσος, $|x| = |x'|$, η ιεράρχηση γίνεται σύμφωνα με το άθροισμα των προηγούμενων ψηφίων της κατάστασης, κατά αύξουσα σειρά, που σημαίνει ότι όταν $|x| = |x'|$, η x προηγείται της x' εάν $\sum_{l=1}^{L-1} n_l \leq \sum_{l=1}^{L-1} n'_l$, για όλα τα l .

Όσον αφορά την ανάλυση της δομής και του πλήθους των καταστάσεων του συνόλου T_n , αξίζει να επισημανθεί ότι το μέγεθος κάθε συνόλου T_n υπολογίζεται από τον τύπο:

$$|T_n| = \sum_{L=1}^K \sum_{x: n_L = n} n_l = 2^{K-1}, \quad n \geq 1.$$

Ο τύπος αυτός μαρτυρά ότι το μέγεθος κάθε συνόλου T_n παραμένει σταθερό για $n \geq 1$ και ίσο με 2^{K-1} . Η σταθερότητα αυτή δεν είναι ένα τυχαίο γεγονός, αλλά απορρέει από τους περιορισμούς που θέτει το όριο ανοχής K στην δομή του διανύσματος κατάστασης. Συγκεκριμένα, για τον υπολογισμό των στοιχείων του T_n αθροίστηκαν όλοι οι δυνατοί συνδυασμοί για κάθε αριθμό L , από 1 έως K , με την προϋπόθεση το τελευταίο ψηφίο να είναι n . Αυτό συμβαίνει διότι το εξωτερικό άθροισμα $\sum_{L=1}^K$ προσδίδει στο σύνολο όλες τις καταστάσεις με διαφορετικό πλήθος

πελατών, από όλο το εύρος του L , ενώ το εσωτερικό άθροισμα $\sum_{x:n_L=n} n_i$, περιορίζει έτσι ώστε να

λαμβάνονται μονάχα καταστάσεις όπου ο τελευταίος πελάτης να διαθέτει το εισιτήριο με αριθμό n . Το γεγονός ότι το αποτέλεσμα των αθροισμάτων ισούται με 2^{K-1} έγκειται στην ελευθερία των επιλογών που επιτρέπει το σύστημα στα υπόλοιπα ψηφία του διανύσματος, εφόσον το τελευταίο ψηφίο του x λαμβάνει υποχρεωτικά την τιμή n . Έτσι, η διαφοροποίηση εντός του T_n εξαρτάται αποκλειστικά από τις τιμές των $L-1$ σε πλήθος, προηγούμενων στοιχείων του, πλην του τελευταίου n_L . Λόγω της συνθήκης εισόδου, το άθροισμα των στοιχείων αυτών πρέπει να μην ξεπερνά τον επιτρεπτό αριθμό K , καθώς σε αντίθετη περίπτωση ο αφιχθείς πελάτης θα έβλεπε κορεσμένο το σύστημα και δεν θα μπορούσε να εισέλθει. Στην Συνδυαστική Ανάλυση, το πρόβλημα του προσδιορισμού όλων των δυνατών τρόπων με τους οποίους μπορούν να σχηματισθούν αθροίσματα θετικών ακεραίων, που δεν ξεπερνούν ένα όριο, οδηγεί στο άθροισμα των διωνυμικών συντελεστών $\sum_{j=1}^{K-1} \binom{K-1}{j}$, το οποίο ισούται με 2^{K-1} .

Για να γίνει πιο κατανοητό, στην περίπτωση ορίου $K=4$, αναζητούμε όλους τους συνδυασμούς n_i , $i=1,2,...,L-1$, με άθροισμα μικρότερο ή ίσο του 3. Όπως φαίνεται από το Σχήμα 3.1, για $L=1$ έχουμε 1 περίπτωση, για $L=2$ έχουμε 3 περιπτώσεις, για $L=3$ έχουμε 3 περιπτώσεις και για $L=4$ έχουμε 1 περίπτωση. Το άθροισμα $1+3+3+1=8$ επιβεβαιώνει τον τύπο $2^{4-1}=8$.

Αυτή η ομοιομορφία των διαστάσεων είναι καθοριστικής σημασίας, καθώς διασφαλίζει ότι η εσωτερική πολυπλοκότητα κάθε T_n παραμένει σταθερή, ανεξάρτητη από την τιμή του n .

Στο σημείο αυτό, πρέπει να επισημανθεί ότι οι στάσιμες πιθανότητες των καταστάσεων στα T_n , $n \leq K-1$, θεωρούνται ήδη γνωστές, καθώς έχουν προκύψει από τους υπολογισμούς του πρώτου σταδίου της διεργασίας. Επομένως, στην συνέχεια της ανάλυσης επικεντρωνόμαστε στον προσδιορισμό των στάσιμων πιθανοτήτων των καταστάσεων στα σύνολα $T_n, n \geq K$.

Αν βασιστούμε στη σχέση $\bigcup_{n \geq K} T_n = \bigcup_{L=1}^K S_L$, η οποία μαρτυρά την ισοδυναμία μεταξύ των δύο διαφορετικών τρόπων οργάνωσης του χώρου καταστάσεων, διασφαλίζεται ότι η μετάβαση από την κάθετη διαμέριση (ανά αριθμό πελατών) στην οριζόντια διαμέριση (ανά επίπεδο τελευταίου ψηφίου) δεν αφήνει καμία κατάσταση εκτός μελέτης. Η ισότητα αυτή επικυρώνει ότι η νέα δομή του χώρου καταστάσεων είναι πλήρης, χωρίς απώλεια πληροφορίας για το σύστημα.

Στο πλαίσιο αυτό, με την αντιμετώπιση των συνόλων $\{T_n, n \geq 0\}$ ως ομάδες (blocks) και των δεικτών $n \geq 1$ ως επιπέδων, διευκολύνεται η παρακολούθηση της δυναμικής του συστήματος. Από τους δείκτες εξαιρέθηκε η αναφορά στον $n=0$ καθώς από το επίπεδο 1 και πάνω συναντάμε στο σώμα της αλυσίδας την επαναλαμβανόμενη λογική των μεταβάσεων.

Ως εκ τούτου, παρατηρούμε ότι με αφετηρία κάθε επίπεδο T_n , εντός της περιοχής κορεσμού $n \geq K$, η Μαρκοβιανή Αλυσίδα ακολουθεί την εξής διαδικασία:

- είτε μεταπηδά στο αμέσως επόμενο επίπεδο T_{n+1} , λόγω της άφιξης ενός νέου πελάτη, η οποία προκαλεί την έκδοση ενός νέου εισιτηρίου αυξάνοντας τον δείκτη n , παρόλο που ο ίδιος επιλέγει να μην εισέλθει στην ουρά.

Για παράδειγμα, εάν βρισκόμαστε στην κατάσταση $x=(1,1,4)$ (επίπεδο T_4 , αφού $n_L=4$), μια νέα άφιξη που οδηγεί σε αποχώρηση θα μεταβάλει την κατάσταση σε $x'=(1,1,5)$, οδηγώντας την αλυσίδα στο επίπεδο T_5 .

- είτε επιστρέφει στην αρχική κατάσταση T_0 , μέσω μίας εξυπηρέτησης, η οποία οδηγεί σε πλήρη κένωση του συστήματος, καθώς όλοι οι ενδιάμεσοι αριθμοί εισιτηρίων, που αντιστοιχούσαν σε αποχωρούντες πελάτες, απελευθερώνονται ακαριαία, μόλις ολοκληρωθεί η τρέχουσα εξυπηρέτηση.

Αν η αλυσίδα βρίσκεται στην κατάσταση $x=(4)$ (μόνο ένας πελάτης με τελευταίο εισιτήριο 4, άρα επίπεδο T_4) και ολοκληρωθεί η εξυπηρέτησή του, το σύστημα μεταβαίνει απευθείας στην κατάσταση 0, οδηγώντας την αλυσίδα στο επίπεδο T_0 .

- είτε, τέλος, παραμένει εντός του ίδιου επιπέδου T_n , μεταβάλλοντας μόνο την εσωτερική δομή του διανύσματος x , το οποίο μπορεί να συμβεί όταν υπάρξει εξυπηρέτηση πελατών που προηγούνται του τελευταίου εισιτηρίου n .

Αν βρισκόμαστε στην κατάσταση $x=(1,1,4)$ και ο πρώτος πελάτης ολοκληρώσει την εξυπηρέτησή του, η νέα κατάσταση γίνεται $x'=(1,4)$. Εδώ παραμένουμε στο επίπεδο T_4 , καθώς το τελευταίο στοιχείο n_L παραμένει ίσο με 4.

Αυτή η διαδοχή των μεταβάσεων, που επαναλαμβάνεται με τον ίδιο τρόπο από το επίπεδο K και μετά, επιτρέπει τον χαρακτηρισμό της αλυσίδας ως GI/M/1.

Το σύστημά μας διαθέτει μια περιοδικότητα στην συμπεριφορά της ουράς, που οδηγεί σε έναν γεννήτορα πίνακα, ανοίγοντας τον δρόμο για τον προσδιορισμό των στάσιμων πιθανοτήτων μέσω της MGM.

Ως εκ τούτου, η οργάνωση του χώρου καταστάσεων σε οριζόντια επίπεδα επιτρέπει την διαμόρφωση του block-διαμερισμένου απειροστού γεννήτορα $\tilde{Q}$. Ο πίνακας αυτός, εστιάζοντας αποκλειστικά στην περιοχή κορεσμού, η οποία διαθέτει την επαναληπτική δομή των μεταβάσεων, λαμβάνει την εξής χαρακτηριστική μορφή:

$$\tilde{Q} = \begin{pmatrix} B & \lambda I^0 & & \\ & B & \lambda I & \\ & & B & \lambda I \end{pmatrix}. \quad (3.15)$$

Στη δομή αυτή, πέρα από τις θέσεις που δεν εμφανίζονται (οι οποίες είτε είναι μηδενικές, είτε δεν εξυπηρετούν στον υπολογισμό των στάσιμων πιθανοτήτων), κάθε γραμμή και στήλη αντιστοιχεί σε ένα ολόκληρο επίπεδο T_n , ξεκινώντας από το T_{K-1} , προχωρώντας στο T_K και συνεχίζοντας διαδοχικά για τα επόμενα επίπεδα. Οι πίνακες B , λI και λI^0 αποτελούν τα δομικά στοιχεία που περιγράφουν τις μεταβάσεις που καταγράφηκαν παραπάνω. Πιο αναλυτικά:

- Ο πίνακας B είναι ένας κάτω τριγωνικός πίνακας διαστάσεων $2^{K-1} \times 2^{K-1}$, που περιγράφει τις εσωτερικές μεταβάσεις, εντός κάθε επιπέδου. Δηλαδή, περιλαμβάνει όλες τις μεταβολές λόγω εξυπηρέτησεων, οι οποίες δεν αλλάζουν τον αριθμό του τελευταίου εισιτηρίου.
- Οι πίνακες λI και λI^0 αντιπροσωπεύουν τις μεταβολές λόγω αφίξεων, προς το αμέσως επόμενο ανώτερο επίπεδο. Ο I είναι ο ταυτοτικός πίνακας διαστάσεων $2^{K-1} \times 2^{K-1}$, ενώ ο I^0 διαφοροποιείται μόνο ως προς το πρώτο στοιχείο της κύριας διαγωνίου του, το οποίο είναι μηδέν.

Όπως είναι αντιληπτό, εάν διαθέτουμε την πληροφορία για το όριο ανοχής K , τότε ο B πίνακας είναι εύκολο να υπολογιστεί.

Για να κατανοήσουμε τη μορφή του πίνακα B , εάν θεωρήσουμε για παράδειγμα την περίπτωση $K=4$, το κάθε επίπεδο για $n \geq K-1$ περιλαμβάνει $2^{K-1} = 2^3 = 8$ καταστάσεις, οι οποίες αντιπροσωπεύουν όλους τους δυνατούς συνδυασμούς πελατών που μπορεί να προηγούνται του τελευταίου εισιτηρίου n . Συνεπώς, το κάθε επίπεδο αποτελείται από το σύνολο καταστάσεων $T_n = \{n, 1n, 2n, 3n, 11n, 21n, 12n, 111n\}$. Σε αυτήν την περίπτωση, ο πίνακας B διαμορφώνεται ως εξής:

$$B = \begin{pmatrix} -(\lambda + \mu) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu & -(\lambda + \mu) & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu & 0 & -(\lambda + \mu) & 0 & 0 & 0 & 0 & 0 \\ \mu & 0 & 0 & -(\lambda + \mu) & 0 & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 & -(\lambda + \mu) & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 & 0 & -(\lambda + \mu) & 0 & 0 \\ 0 & 0 & \mu & 0 & 0 & 0 & -(\lambda + \mu) & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 & 0 & -(\lambda + \mu) \end{pmatrix}.$$

Εδώ, οι τιμές του μ κάτω από την διαγώνιο επιβεβαιώνουν την ικανότητα του συστήματος να μεταβάλλει την εσωτερική του κατάσταση μέσω εξυπηρέτησεων, παραμένοντας στο ίδιο block της αλυσίδας. Συγκεκριμένα, ο ρυθμός μ υποδηλώνει τη μετάβαση από μια εσωτερική κατάσταση $x = (n_1, n_2, n_3, \dots, n_L)$ σε μία νέα κατάσταση $x' = (n_2, n_3, \dots, n_L)$, δείχνοντας ότι το σύστημα προχωρά στον επόμενο πραγματικό πελάτη, χωρίς να αλλάζει ο αριθμός του τελευταίου εκδοθέντος εισιτηρίου n . Δηλαδή, τα στοιχεία αυτά απεικονίζουν τις εσωτερικές μεταβάσεις που προκύπτουν αποκλειστικά από την ολοκλήρωση μιας εξυπηρέτησης.

Όσον αφορά τα διαγώνια στοιχεία $-(\lambda + \mu)$, είναι φανερό ότι κάθε ένα από αυτά αντιπροσωπεύει τον συνολικό ρυθμό εξόδου από την αντίστοιχη κατάσταση. Ο ρυθμός λ υποδηλώνει την έξοδο από το τρέχον επίπεδο T_n προς το T_{n+1} , λόγω της άφιξης ενός νέου πελάτη, ενώ ο ρυθμός μ υποδηλώνει την έξοδο λόγω της εξυπηρέτησης του πελάτη που προηγείται στην ουρά. Το αρνητικό πρόσημο οφείλεται στη ιδιότητα του απειροστού γεννήτορα πίνακα να έχει άθροισμα στοιχείων κάθε γραμμής ίσο με μηδέν.

Έχοντας ορίσει τη δομή του γεννήτορα πίνακα, μπορούμε να προχωρήσουμε στον υπολογισμό των στάσιμων πιθανοτήτων. Έστω, $\bar{p}_n$ το διάνυσμα των στάσιμων πιθανοτήτων που αντιστοιχεί στο επίπεδο T_n . Για την περιοχή κορεσμού $n \geq K$, οι εξισώσεις ισορροπίας σε μορφή πινάκων διαμορφώνονται ως εξής:

$$\lambda \bar{p}_{K-1} I^0 + \bar{p}_K B = 0, \quad (3.16)$$

$$\lambda \bar{p}_n + \bar{p}_{n+1} B = 0, \quad n = K, K+1, \dots \quad (3.17)$$

Η σχέση (3.16) περιγράφει την ισορροπία μεταξύ του επιπέδου T_{K-1} και του πρώτου επιπέδου κορεσμού T_K . Ενώ, η σχέση (3.17) αναδεικνύει την επαναληπτική φύση του συστήματος για κάθε επόμενο επίπεδο $n \geq K$.

Για την επίλυση αυτών των αναδρομικών σχέσεων, ορίζουμε τους πίνακες ρυθμού $\bar{R}_0$ και $\bar{R}$ ως εξής:

$$\bar{R}_0 = -\lambda I^0 B^{-1} \quad \text{και} \quad \bar{R} = -\lambda B^{-1}.$$

Τότε, από την (3.16), προκύπτει:

$$\bar{p}_K = \bar{p}_{K-1} (-\lambda I^0 B^{-1}) = \bar{p}_{K-1} \bar{R}_0. \quad (3.18)$$

Τώρα, από την (3.17), λαμβάνοντας αντί n το $n-1$ και επιλύοντας ως προς $\bar{p}_n$, οδηγούμαστε στην ακόλουθη σχέση, στην οποία σιωπηλά επαναλαμβάνοντας $n-1-k$ φορές, παίρνουμε την αμέσως επόμενη μορφή. Έπειτα, εφαρμόζοντας την (3.18), καθώς και τους πίνακες ρυθμού $\bar{R}_0$, $\bar{R}^{n-K}$, έχουμε την τελική απλοποιημένη μορφή. Ακολουθούν όλες οι μορφές, που περιγράφηκαν, διαδοχικά:

$$\begin{aligned} \bar{p}_n &= \bar{p}_{n-1} (-\lambda B^{-1}) = \bar{p}_K (-\lambda B^{-1})^{n-K} \\ &= \bar{p}_{K-1} \bar{R}_0 \bar{R}^{n-K}, \quad n = K+1, K+2, \dots \end{aligned} \quad (3.19)$$

Η σχέση (3.19) είναι καθοριστική, καθώς δείχνει ότι οι πιθανότητες των ανώτερων επιπέδων μειώνονται γεωμετρικά, βάσει του πίνακα $\bar{R}$, όσο αυξάνεται το μέγεθος της ουράς. Η δομή της

σχέσης αυτής υπαγορεύει τη σταδιακή μείωση των πιθανοτήτων $\bar{p}_{K-1}$, καθώς το διάνυσμα $\bar{p}_{K-1}\bar{R}_0$ πολλαπλασιάζεται διαδοχικά με δυνάμεις του πίνακα $\bar{R}$, που λόγω της μορφής του λειτουργεί ως συντελεστής συρρίκνωσης των πιθανοτήτων. Αυτό είναι αναγκαίο, διότι εάν οι τιμές δεν μειώνονταν, όσο αυξάνεται ο εκθέτης $n-K$, το άθροισμα των πιθανοτήτων δεν θα μπορούσε να συγκλίνει στην μονάδα.

Εφόσον το διάνυσμα $\bar{p}_{K-1}$ είναι ήδη γνωστό από την ανάλυση της συνολικής Μαρκοβιανής Αλυσίδας που προηγήθηκε, μπορούμε πλέον να υπολογίσουμε τις ακριβείς στάσιμες πιθανότητες για κάθε μεμονωμένη κατάσταση εντός των επιπέδων κορεσμού.

3.4 Βασικά Μέτρα Απόδοσης της Ουράς Εισιτηρίων

Μετά τον προσδιορισμό της στάσιμης κατανομής $\{p(x): x \in \tilde{J}\}$, της Μαρκοβιανής Αλυσίδας, είναι πλέον εφικτός ο υπολογισμός των βασικών μεγεθών, που περιγράφουν την λειτουργία του συστήματος.

Η μελέτη επικεντρώνεται στις δύο θεμελιώδεις μεταβλητές: τον αριθμό N των πραγματικών πελατών που βρίσκονται στην ουρά, ο οποίος καλείται και θέση στην ουρά, και τη διαφορά μεταξύ αριθμού εισιτηρίου και ένδειξης πίνακα D , δηλαδή τη θέση του τελευταίου εκδοθέντος εισιτηρίου, η οποία αναφέρεται και ως θέση εισιτηρίου. Στο σύστημά μας, οι δύο αυτές μεταβλητές δεν ταυτίζονται υποχρεωτικά, καθώς η θέση εισιτηρίου D επηρεάζεται και από τους αριθμούς εκείνων που αποχώρησαν κατά την άφιξή τους, ενώ η θέση στην ουρά N καταγράφει μόνο όσους παρέμειναν για εξυπηρέτηση. Είμαστε σε θέση να υπολογίσουμε την από κοινού κατανομή των N και D , αλλά και τις περιθώριες κατανομές αυτών.

Ακολουθούν οι τύποι που ορίζουν τις περιθώριες κατανομές των μεγεθών αυτών:

$$P(N=0)=p(0) \text{ και } P(N=L)=\sum_{x \in E_L} p(x), \quad L=1,2,\dots,K, \quad (3.20)$$

$$P(D=0)=p(0) \text{ και } P(D=d)=\sum_{x: \|x\|=d} p(x), \quad d=1,2,\dots \quad (3.21)$$

Οι παραπάνω σχέσεις δείχνουν τον τρόπο με τον οποίο οι πιθανότητες των μεμονωμένων καταστάσεων $p(x)$ χρησιμοποιούνται για τον προσδιορισμό των κατανομών των κεντρικών μεταβλητών του συστήματος. Αρχικά, στην περίπτωση που το σύστημα είναι κενό, οι δύο μεταβλητές ταυτίζονται, με την πιθανότητα $P(N=0)$ και $P(D=0)$ να ισούται με την πιθανότητα $p(0)$ της μηδενικής κατάστασης.

Για τις υπόλοιπες περιπτώσεις, η διαφοροποίηση των δύο μεγεθών γίνεται σαφής μέσω του τρόπου άθροισης των πιθανοτήτων. Αφού η θέση εισιτηρίου D ορίζεται ως το άθροισμα όλων των στοιχείων του διανύσματος κατάστασης, δηλαδή $\|x\| = \sum_{j=1}^L n_j$, τότε η πιθανότητα $P(D=d)$ προκύπτει από το άθροισμα των πιθανοτήτων όλων εκείνων των καταστάσεων του x που οδηγούν στο ίδιο συνολικά άθροισμα εισιτηρίων d , ενσωματώνοντας με αυτόν τον τρόπο την επίδραση των ατόμων που αποχώρησαν από το σύστημα.

Αντιθέτως, η θέση στην ουρά N υπολογίζεται μέσω του συνόλου E_L , το οποίο λειτουργεί ως ομαδοποίηση των καταστάσεων με βάση τον αριθμό των πραγματικών πελατών του συστήματος. Καθώς ο αριθμός των πελατών L αντιστοιχεί στην διάσταση του διανύσματος x , δηλαδή $|x| = L$, η δεύτερη σχέση της (3.20) αθροίζει τις πιθανότητες όλων εκείνων των καταστάσεων που περιλαμβάνουν ακριβώς L εναπομείναντα άτομα στην ουρά, ανεξάρτητα από την αριθμητική τιμή των εισιτηρίων τους.

Με βάση τις κατανομές που ορίστηκαν προηγουμένως, μπορούμε να εξάγουμε συμπεράσματα για την συμπεριφορά των πελατών κατά την άφιξή τους, καθώς και για το ποσοστό του χρόνου που το σύστημα παραμένει απασχολημένο.

Αρχικά, ορίζεται η πιθανότητα αποχώρησης (balking probability) P_b , η οποία αντιπροσωπεύει το ποσοστό των πελατών που επιλέγουν να μην εισέλθουν στην ουρά επειδή η θέση του εισιτηρίου που λαμβάνουν υπερβαίνει το όριο ανοχής τους. Η πιθανότητα αυτή προκύπτει εάν αφαιρέσουμε από τη μονάδα το άθροισμα των πιθανοτήτων όλων των καταστάσεων που αφορούν τους πελάτες που θα παραμείνουν στην ουρά, δηλαδή αυτών των οποίων η θέση εισιτηρίου είναι μικρότερη από το όριο K , όπως περιγράφεται από τη σχέση:

$$P_b = 1 - \sum_{x: \|x\| < K} p(x). \quad (3.22)$$

Στη συνέχεια, ο συντελεστής αξιοποίησης (utilization factor) του συστήματος ρ_e προσδιορίζει την πραγματική απασχόληση του εξυπηρετητή, λαμβάνοντας υπόψη αποκλειστικά τους πελάτες που τελικά παρέμειναν στην ουρά. Ο δείκτης αυτός προκύπτει από το γινόμενο του συντελεστή $\rho = \lambda / \mu$ επί την πιθανότητα ένας πελάτης να παραμείνει στο σύστημα, αποτυπώνοντας τον πραγματικό φόρτο εργασίας, μετά την αφαίρεση του πλήθους των ατόμων που επέλεξαν να αποχωρήσουν:

$$\rho_e = \frac{\lambda}{\mu} (1 - P_b) = \rho (1 - P_b). \quad (3.23)$$

Ιδιαίτερο ενδιαφέρον παρουσιάζει η ανάλυση του συστήματος από την πλευρά του πελάτη, για τον οποίο η κρίσιμη πληροφορία δεν είναι ο συνολικός αριθμός των εκδοθέντων εισιτηρίων, αλλά η εκτιμώμενη πραγματική του θέση στην ουρά N , δεδομένης της ένδειξης d που έλαβε κατά την

είσοδό του. Η πληροφορία αυτή προκύπτει απευθείας από το ορισμό της δεσμευμένης πιθανότητας ως ο λόγος της κοινής πιθανότητας των μεγεθών N και D προς την πιθανότητα εμφάνισης της συγκεκριμένης θέσης εισιτηρίου d , ως εξής :

$$P(N = L | D = d) = \frac{P(N = L, D = d)}{P(D = d)} = \frac{\sum_{x: x \in E_L, \|x\|=d} p(x)}{\sum_{x: \|x\|=d} p(x)}, \quad d = L, L+1, \dots \quad (3.24)$$

Βασίζόμενοι στη δεσμευμένη κατανομή που ορίστηκε με την (3.24), μπορούμε πλέον να προσδιορίσουμε τη μέση θέση στην ουρά για έναν πελάτη με εισιτήριο d . Η τιμή αυτή, $E(N = L | D = d)$, αντιπροσωπεύει την πραγματική σειρά προτεραιότητας που αναμένει να έχει ο πελάτης, με δεδομένο ότι κάποιοι από τους προπορευόμενους κατόχους εισιτηρίων ενδέχεται να έχουν ήδη αποχωρήσει από το σύστημα, και υπολογίζεται από τη σχέση:

$$E[N | D = d] = \sum_{L=0}^{d \wedge K} LP(N = L | D = d) = \frac{\sum_{L=0}^{d \wedge K} \sum_{x \in E_L, \|x\|=d} Lp(x)}{\sum_{\|x\|=d} p(x)}. \quad (3.25)$$

Στην παραπάνω σχέση, ο συμβολισμός $d \wedge K$ υποδηλώνει ότι το άθροισμα τερματίζεται μόλις εξαντληθεί είτε ο αριθμός του εισιτηρίου d , είτε η μέγιστη χωρητικότητα του συστήματος K , ανάλογα με ποιο από τα δύο συμβεί πρώτα.

Η πληροφορία αυτή επιτρέπει τον προσδιορισμό του μέσου χρόνου αναμονής $E[W | D = d]$, μέχρι την έναρξη της εξυπηρέτησης, με δεδομένη την θέση εισιτηρίου d . Ο χρόνος αυτός προκύπτει από το γινόμενο της μέσης θέσης του πελάτη επί τον μέσο χρόνο εξυπηρέτησης $1/\mu$ ανά άτομο, σύμφωνα με τον ακόλουθο τύπο:

$$E[W | D = d] = \frac{1}{\mu} E[N | D = d]. \quad (3.26)$$

Η σχέση αυτή βασίζεται στην λογική ότι ο συνολικός χρόνος αναμονής ισούται με το άθροισμα των χρόνων εξυπηρέτησης όλων των ατόμων που προηγούνται στην ουρά.

Παρόλο που ο μέσος χρόνος αναμονής, που υπολογίζεται από την (3.26), αποτελεί έναν χρήσιμο δείκτη, συχνά δεν επαρκεί για την πλήρη κατανόηση της στοχαστικής συμπεριφοράς του συστήματος. Αυτό συμβαίνει γιατί η δυσανεμία που προκαλεί η αναμονή, αρκετές φορές, τείνει να παρουσιάζει μη γραμμική συμπεριφορά. Σε τέτοιες περιπτώσεις, η ενόχληση του πελάτη δεν αυξάνεται σταθερά στον χρόνο που περνά, αλλά μπορεί να επιδεινωθεί απότομα μετά από ένα κρίσιμο χρονικό όριο.

Για τον λόγο αυτό, καθίσταται απαραίτητος ο υπολογισμός της δεσμευμένης συνάρτησης κατανομής του χρόνου αναμονής, η οποία προσδιορίζει την πιθανότητα ο χρόνος W να μην υπερβεί ένα συγκεκριμένο όριο w , δεδομένης της ένδειξης του εισιτηρίου d .

Η σχέση που περιγράφει αυτή την πιθανότητα είναι η εξής:

$$P[W \leq w | D = d] = \sum_{L=1}^{d \wedge K} P\left(\sum_{n=1}^L X_n \leq w\right) P(N = L | D = d), \quad d \geq 1, \quad (3.27)$$

όπου οι X_n είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με ρυθμό μ , οι οποίες αντιπροσωπεύουν τους χρόνους εξυπηρέτησης των πελατών, ενώ η πιθανότητα $P(N = L | D = d)$ έχει οριστεί προηγουμένως.

Το δεξί μέλος της (3.27) αναλύεται σε τρεις βασικούς όρους, που συνεργάζονται για να προσφέρουν την τελική εικόνα του χρόνου αναμονής. Η πιθανότητα θέσης συνδυάζεται με την πιθανότητα $P\left(\sum_{n=1}^L X_n \leq w\right)$, η οποία εκφράζει την πιθανότητα το ενδεχόμενο της συνολικής εξυπηρέτησης των L προπορευόμενων πελατών να ολοκληρωθεί εντός του χρονικού ορίου w . Τέλος, για άλλη μια φορά, ο τελεστής του αθροίσματος $\sum_{L=1}^{d \wedge K}$ υποδηλώνει ότι λαμβάνονται υπόψη όλες οι πιθανές θέσεις L , με το άνω όριο να καθορίζεται από το ελάχιστο μεταξύ της ένδειξης του εισιτηρίου και της μέγιστης χωρητικότητας του συστήματος.

Παρά την εύρεση της αναλυτικής έκφρασης για την κατανομή του χρόνου αναμονής, η πρακτική εφαρμογή της προσκρούει στο πρόβλημα της πολυπλοκότητας των υπολογισμών. Η υπολογιστική αυτή δυσκολία πηγάζει από το γεγονός ότι οι πληθικότητες των συνόλων E_L και T_n αυξάνονται εκθετικά συναρτήσει του ορίου ανοχής K . Συγκεκριμένα, η πληθικότητα του συνόλου T_n , το οποίο αφορά τη διαμέριση του χώρου καταστάσεων και περιλαμβάνει όλες τις καταστάσεις x όπου ο τελευταίος πελάτης έχει ένδειξη εισιτηρίου $n_L = n$, δίνεται από τη σχέση:

$$|T_n| = 2^{K-1}, \quad n \geq 1,$$

ενώ, για τα σύνολα E_L , που αντιπροσωπεύουν τη συλλογή των L -διάστατων διανυσμάτων κατάστασης για L πελάτες, η πληθικότητα ορίζεται ως :

$$|E_L| = (K-1) \binom{K-1}{L-1} + 1, \quad L = 1, \dots, K. \quad (3.28)$$

Αυτή η εκθετική αύξηση οδηγεί σε έναν συνολικό αριθμό καταστάσεων στην Ομαδοποιημένη Μαρκοβιανή Αλυσίδα. Ο συνολικός αυτός χώρος καταστάσεων προκύπτει ως το άθροισμα της κενής κατάστασης της οποίας η πληθικότητα είναι 1 ($|E_0| = 1$) και των πληθικοτήτων των επιμέρους συνόλων E_L . Δηλαδή, ισχύει $|J| = |E_0| + \sum_{L=1}^K |E_L|$, που οδηγεί σε:

$$|J| = (K-1) \cdot 2^{K-1} + K + 1, \quad (3.29)$$

αφού ικανοποιείται η σχέση: $\sum_{L=1}^K |E_L| = (K-1) \sum_{L=1}^K \binom{K-1}{L-1} + \sum_{L=1}^K 1 = (K-1) \cdot 2^{K-1} + K$.

Όπως γίνεται αντιληπτό, για ρεαλιστικά μεγάλες τιμές του K , το μέγεθος του χώρου καταστάσεων καθίσταται απαγορευτικό για άμεσους αριθμητικούς υπολογισμούς. Η διαπίστωση αυτή επιβεβαιώνει την ανάγκη για την ανάπτυξη προσεγγιστικών διαδικασιών, οι οποίες αποτελούν το αντικείμενο της επόμενης ενότητας.

3.5 Αντιμετώπιση της Υπολογιστικής Πολυπλοκότητας: Το Μοντέλο της Τροποποιημένης Ουράς

Η υπολογιστική διαδικασία που αναλύθηκε στις προηγούμενες ενότητες βασίζεται στον λεπτομερή προσδιορισμό των καταστάσεων μέσω του διανύσματος χ . Ωστόσο, η μέθοδος αυτή παρουσιάζει σημαντικούς περιορισμούς όταν το όριο ανοχής K λαμβάνει μεγάλες τιμές. Όπως παρατηρήθηκε, τα μεγέθη των συνόλων E_L και T_n αυξάνονται εκθετικά, καθιστώντας την αριθμητική επίλυση του συστήματος εξαιρετικά πολύπλοκη και, σε πολλές περιπτώσεις, πρακτικά ανέφικτη για μεγάλα συστήματα. Για τον λόγο αυτό, εισάγεται μία προσέγγιση που στοχεύει στη μείωση του πλήθους των καταστάσεων, γνωστή και ως «Τροποποιημένη Ουρά Εισιτηρίων» (Modified Ticket Queue).

Στο τροποποιημένο αυτό μοντέλο, αντί για μία ενιαία ουρά όπου οι πελάτες που παραμένουν και εκείνοι που αποχωρούν είναι αναμεμιγμένοι, εισάγεται η δομή δύο διακριτών ουρών με διαφορετική λειτουργική βαρύτητα, ανάλογα με την απόφαση που λαμβάνει ο πελάτης κατά την άφιξή του:

- Ουρά αφικνούμενων πελατών που δεν εγκαταλείπουν την θέση τους:
 - Σε αυτή την ομάδα εντάσσονται οι πελάτες που επιλέγουν να εισέλθουν στο σύστημα για εξυπηρέτηση.
 - Διαθέτουν απόλυτη προτεραιότητα, πράγμα που σημαίνει ότι η διαδικασία εξυπηρέτησης εστιάζει αποκλειστικά σε αυτούς, αγνοώντας προσωρινά την ύπαρξη των πελατών που αποχώρησαν.
 - Η ροή εντός της ουράς διέπεται από την αρχή FCFS (First Come, First Served), με κάθε νέο μέλος να καταλαμβάνει την τελευταία διαθέσιμη θέση της συγκεκριμένης ουράς.
- Ουρά αφικνούμενων πελατών που αποχωρούν:
 - Εδώ κατατάσσονται όσοι πελάτες αποφασίζουν να μην περιμένουν, βάσει του ορίου ανοχής τους, δημιουργώντας μία εικονική ουρά.
 - Τα εισιτήρια αυτών των πελατών παραμένουν ενεργά ως θέσεις στο σύστημα, αλλά τοποθετούνται σε μία δευτερεύουσα δομή αναμονής, όπου και σε αυτή την περίπτωση κάθε νέο μέλος τοποθετείται στο τέλος της ουράς.
 - Η αποδέσμευση αυτών των εικονικών θέσεων δεν γίνεται μεμονωμένα, αλλά πραγματοποιείται μαζικά και ακαριαία τη στιγμή που η ουρά των πραγματικών πελατών μηδενίζεται.

Η κρίσιμη παράμετρος που εξασφαλίζει την ακρίβεια της προσέγγισης είναι η διαδικασία λήψης απόφασης. Κάθε νέος πελάτης, κατά την άφιξή του, λαμβάνει ένα εισιτήριο η θέση του οποίου αντιστοιχεί στο συνολικό άθροισμα των ατόμων που βρίσκονται και στις δύο επιμέρους ουρές. Δεδομένου ότι ο πελάτης δεν έχει την δυνατότητα να παρατηρήσει τα επιμέρους μήκη των δύο ουρών, η απόφασή του για παραμονή ή αποχώρηση βασίζεται αποκλειστικά στην συνολική του θέση, ακριβώς όπως θα συνέβαινε και στο αρχικό μοντέλο. Μόλις ληφθεί η απόφαση, ο πελάτης εντάσσεται αμέσως στο τέλος της αντίστοιχης ουράς, διατηρώντας τη χρονική σειρά προτεραιότητας που ορίζει η στιγμή της άφιξής του.

Η αποτελεσματικότητα αυτής της προσέγγισης βασίζεται στη στοχαστική συμπεριφορά του συστήματος και στον τρόπο με τον οποίο κατανέμονται οι πελάτες στην ουρά. Αυτή η λειτουργική απλοποίηση πλαισιώνεται από την βασική παραδοχή ότι οι δύο τύποι πελατών τείνουν να δημιουργούν ομάδες, αντί να αναμειγνύονται τυχαία. Το φαινόμενο αυτό γίνεται διαισθητικά κατανοητό αν αναλογιστούμε την λειτουργία του συστήματος κάτω από ακραίες καταστάσεις φόρτου.

Σε συνθήκες χαμηλού φόρτου, η ουρά παραμένει μικρή και οι εξυπηρετήσεις ολοκληρώνονται γρήγορα, με αποτέλεσμα οι περιπτώσεις αποχώρησης να είναι σπάνιες και η ανάμειξη των πελατών σχεδόν μηδενική. Σε συνθήκες υψηλού φόρτου, το σύστημα λειτουργεί συχνά κοντά στο όριο K . Όταν η ουρά είναι ήδη επιβαρυνμένη, οι νέοι πελάτες που καταφθάνουν, βλέποντας έναν αριθμό στο εισιτήριο τους που υπερβαίνει την υπομονή τους, αποχωρούν διαδοχικά ο ένας μετά τον άλλο.

Για παράδειγμα, σε ένα σύστημα με όριο ανοχής $K=10$, αν υπάρχουν ήδη 10 άτομα στην αναμονή, οι επόμενοι πελάτες που θα φτάσουν, θα λάβουν το εισιτήριο τους και θα αποχωρήσουν αμέσως. Αυτό έχει ως αποτέλεσμα τη δημιουργία μιας συνεχόμενης σειράς, εικονικών εισιτηρίων (από πελάτες που έφυγαν) τα οποία βρίσκονται συσσωρευμένα στο τέλος της ουράς. Η παρατήρηση αυτή επιτρέπει τον διαχωρισμό των πελατών χωρίς να αλλοιώνεται η συνολική δυναμική του συστήματος.

Μέσω αυτής της αναδιάρθρωσης, ο χώρος των καταστάσεων μειώνεται σημαντικά, καθώς η περιγραφή του συστήματος απαιτεί πλέον μόνο ένα ζεύγος μεταβλητών (L, n) . Η μεταβλητή L , όπου $0 \leq L \leq K$, αντιπροσωπεύει το πλήθος των πραγματικών πελατών που παραμένουν για εξυπηρέτηση (συμπεριλαμβανομένου του εξυπηρετούμενου), όπως και στις προηγούμενες ενότητες. Ο ρόλος της διευκολύνεται υπολογιστικά, καθώς δεν απαιτείται πλέον η γνώση της ακριβούς θέσης των πελατών που αποχωρούν ανάμεσα στους παραμένοντες πελάτες. Αντίστοιχα, η μεταβλητή n , όπου $n \geq 0$, ορίζει πλέον το συνολικό πλήθος των πελατών που έχουν αποχωρήσει και των οποίων οι θέσεις παραμένουν ενεργές στο σύστημα, εντός της εικονικής ουράς.

Είναι σημαντικό να τονιστεί ότι η στρατηγική επίλυσης για το τροποποιημένο μοντέλο παραμένει πιστή στη διαδικασία των δύο σταδίων, που αναπτύχθηκαν προηγουμένως. Αρχικά, οι καταστάσεις ομαδοποιούνται σε υπερ-καταστάσεις για τον υπολογισμό της στάσιμης κατανομής (Στάδιο 1) και στη συνέχεια ακολουθεί η αποσύνθεσή τους για τον προσδιορισμό των πιθανοτήτων των επιμέρους καταστάσεων (Στάδιο 2). Με αυτόν τον τρόπο, επιτυγχάνεται η επίλυση του συστήματος με τις ίδιες μεθοδολογικές αρχές, αλλά σε υπολογιστικά διαχειρίσιμο πλαίσιο.

Για την πρακτική εφαρμογή της προσέγγισης, στο τροποποιημένο μοντέλο, προχωρούμε σε προσδιορισμό των υπερ-καταστάσεων S_L . Παρόλο που το συγκεκριμένο σύμβολο χρησιμοποιήθηκε για διαφορετικούς σκοπούς στην ανάλυση του αρχικού μοντέλου, σε προηγούμενες ενότητες, η εκ νέου χρήση του εδώ για την περιγραφή της τροποποιημένης αλυσίδας δεν προκαλεί σύγχυση, καθώς εξυπηρετεί τη νέα δομή. Για κάθε επίπεδο πραγματικών πελατών, η υπερ-κατάσταση S_L ορίζεται πλέον ως το σύνολο:

$$S_L = \{(L, n) : n \geq K-1\}, \quad L = 1, 2, \dots, K.$$

Η αναγκαιότητα αυτής της ομαδοποίησης γίνεται φανερή αν αναλύσουμε τον μηχανισμό λήψης αποφάσεων. Οποιοσδήποτε νέος πελάτης εισέρχεται στο σύστημα και το βρίσκει σε μία κατάσταση που ανήκει στο S_L , αντιλαμβάνεται ότι η σειρά προτεραιότητάς του, που για τον δεδομένο πελάτη είναι $L+n+1$, υπερβαίνει το όριο ανοχής του K , καθώς ήδη ισχύει $L+n \geq 1+(K-1) = K$, αφού $L \geq 1$ και $n \geq (K-1)$. Επειδή, λοιπόν, πέραν αυτού του ορίου, η πιθανότητα αποχώρησης είναι βέβαιη, ανεξάρτητα από την ακριβή τιμή της μεταβλητής n , δεν απαιτείται η μεμονωμένη καταγραφή των εικονικών θέσεων, επιτρέποντας την αντιμετώπιση όλων αυτών των περιπτώσεων ως μια ενιαία κατάσταση αναμονής.

Στη συνέχεια, αναδομούμε τα σύνολα, που περιγράφουν όλες τις δυνατές διαμορφώσεις για κάθε σταθερό L , συνεχίζοντας να αναφερόμαστε στο σύνολο των καταστάσεων που έχουν L παραμένοντες πελάτες ως E_L .

Στην πορεία, θα μελετήσουμε την δομή και το πλήθος των καταστάσεων του χώρου, αναλύοντας την μορφή των επιμέρους συνόλων. Για την περίπτωση του κενού συστήματος ισχύει $E_0 = \{(0, 0)\}$, ενώ για τις υπόλοιπες περιπτώσεις τα σύνολα ορίζονται ως:

$$E_L = \{x = (L, n) : 0 \leq n \leq K-2\} \cup S_L, \quad L = 1, 2, \dots, K.$$

Μέσω αυτής της μεθοδολογίας, επιτυγχάνεται μία κρίσιμη ομοιομορφία, καθώς κάθε σύνολο E_L περιλαμβάνει ακριβώς σταθερό πλήθος K στοιχείων, αφού αποτελείται από $K-1$ καταστάσεις και την επιπλέον υπερ-κατάσταση S_L , δηλαδή σημειώνεται ότι $|E_L| = K$, $L \geq 1$.

Τώρα, ο συνολικός χώρος των καταστάσεων J^a προκύπτει από την ένωση αυτών των συνόλων και εκφράζεται ως:

$$J^a = \{x = (L, n) : 0 \leq L \leq K, 0 \leq n \leq K-2\} \cup \{S_1, S_2, \dots, S_K\}. \quad (3.30)$$

Το σημαντικότερο πλεονέκτημα αυτής της αναδιάρθρωσης είναι η δραστική μεταβολή της υπολογιστικής πολυπλοκότητας του συστήματος. Ενώ στο αρχικό ομαδοποιημένο μοντέλο ο χώρος καταστάσεων παρουσίαζε εκθετική αύξηση, όπως προκύπτει και από τον τύπο (3.29) που παρουσιάστηκε στην ενότητα 3.4, στην τροποποιημένη αλυσίδα το μέγεθός του περιορίζεται σε $|J^a| = K^2 + 1$. Η τιμή αυτή προκύπτει από τη νέα δομή του χώρου καταστάσεων, ο οποίος αποτελείται από την κατάσταση ηρεμίας E_0 και τα K ισοπληθή επίπεδα E_L , $L = 1, 2, \dots, K$.

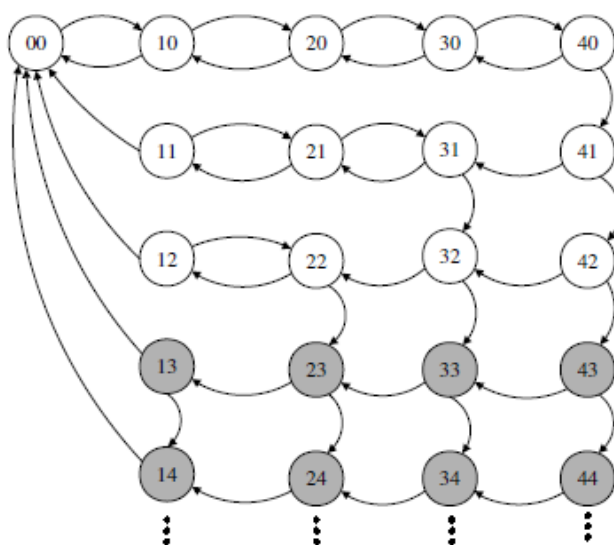
Συγκεκριμένα, όπως έχει ήδη αναφερθεί, κάθε επίπεδο E_L περιλαμβάνει K καταστάσεις και ως εκ τούτου, ο συνολικός αριθμός των καταστάσεων υπολογίζεται από το γινόμενο των K επιπέδων επί τις K καταστάσεις που περιέχει το καθένα, συν την αρχική κατάσταση, καταλήγοντας στον συγκεντρωτικό τύπο $K^2 + 1$ καταστάσεων. Επομένως, η μετάβαση από την εκθετική στην τετραγωνική πληθικότητα των καταστάσεων καθιστά την επίλυση υπολογιστικά διαχειρίσιμη, ακόμα και για μεγαλύτερες τιμές της παραμέτρου K .

Για να γίνει αντιληπτή η κλίμακα της βελτίωσης και κατ' επέκταση η σπουδαιότητα της τροποποιημένης αλυσίδας, για τη μελέτη του μοντέλου, ας θεωρήσουμε ένα σύστημα με όριο ανοχής $K = 15$. Στην αρχική διαμόρφωση του συστήματος, το μέγεθος $|J|$ του χώρου καταστάσεων θα ανερχόταν σε $(15-1) \cdot 2^{15-1} + 15 + 1 = 14 \cdot 2^{14} + 16 = 229.392$ καταστάσεις, συνολικά.

Αντίθετα, με τη νέα δομή των υπερ-καταστάσεων στην τροποποιημένη αλυσίδα, για το ίδιο K , ο αντίστοιχος χώρος $|J^a|$ συμπίπτει δραστικά σε μόλις $15^2 + 1 = 226$ καταστάσεις, συνολικά. Αυτή η τεράστια διαφορά αναδεικνύει την ανάγκη της νέας προσέγγισης μέσω της τροποποιημένης ουράς εισιτηρίων, ακόμα και για συστήματα μεγάλης ανοχής, καθώς απλοποιεί εξαιρετικά τον υπολογιστικό φόρτο.

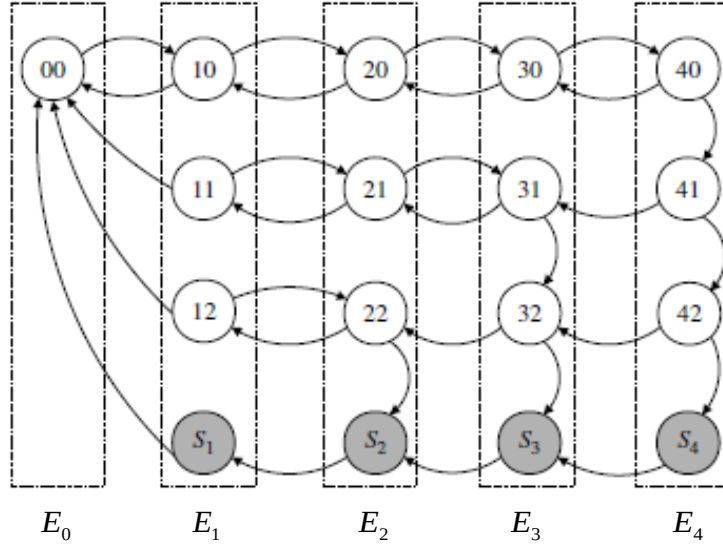
Προκειμένου να μελετήσουμε την συμπεριφορά του συστήματος, υπό το πρίσμα της νέας δομής, παρουσιάζουμε τα αντίστοιχα διαγράμματα μεταβάσεων για ένα σύστημα με όριο ανοχής $K = 4$, όπως κάναμε και στο αρχικό πλαίσιο ανάλυσης. Η οπτικοποίηση αυτή είναι καθοριστική για την κατανόηση της διαφοράς μεταξύ της αρχικής ομαδοποιημένης και της τροποποιημένης προσέγγισης.

Στο Σχήμα 3.3 απεικονίζεται η τροποποιημένη ουρά εισιτηρίων χωρίς τη χρήση των υπερ-καταστάσεων. Παρατηρείται ότι, ενώ η οριζόντια διάσταση (πραγματικοί πελάτες L) είναι πεπερασμένη, η κατακόρυφη διάσταση (εικονική αναμονή n) επεκτείνεται θεωρητικά στο άπειρο, καθώς αφήνεται περιθώριο σε νέους πελάτες να καταγράφουν άφιξη χωρίς ωστόσο να έχουν πραγματική είσοδο.



Σχήμα 3.3: Διάγραμμα μεταβάσεων της τροποποιημένης ουράς για $K = 4$

Αντίθετα, στο Σχήμα 3.4 εφαρμόζεται η στρατηγική της σύμπτυξης των καταστάσεων. Οι καταστάσεις που περιγράφουν αναμονή πέραν του ορίου ανοχής ομαδοποιούνται στις υπερ-καταστάσεις S_1, S_2, S_3, S_4 , μετατρέποντας την αλυσίδα σε ένα πεπερασμένο πλέγμα.



Σχήμα 3.4: Διάγραμμα μεταβάσεων της τροποποιημένης ουράς με υπερ-καταστάσεις για $K = 4$

Αυτή η οργάνωση επιτρέπει στην τροποποιημένη Μαρκοβιανή Αλυσίδα να αντιμετωπιστεί ως μία διαδικασία QBD. Για τον υπολογισμό της στάσιμης κατανομής $\{p^a(x) : x \in J^a\}$ του τροποποιημένου συστήματος, θεωρούμε τα σύνολα $\{E_0, E_1, \dots, E_K\}$ ως διακριτά επίπεδα.

Ο απειροστός γεννήτορας πίνακας Q επαναπροσδιορίζεται σε μία block-τριδιαγώνια μορφή, η οποία αντικατοπτρίζει τις επιτρεπτές μεταβάσεις μεταξύ των επιπέδων:

$$Q = \begin{pmatrix} -\lambda & A_{0,1} & & & \\ \mu \mathbf{e}'_K & A_{1,1} & A_{1,2} & & \\ & \mu I & A_{2,2} & A_{2,3} & \\ & & \ddots & \ddots & \ddots \\ & & & \mu I & A_{K-1,K-1} & A_{K-1,K} \\ & & & & \mu I & A_{K,K} \end{pmatrix}. \quad (3.31)$$

Η δομή του πίνακα Q ενσωματώνει όλη τη φυσική του προβλήματος, με κάθε block-στοιχείο του πίνακα να αντιστοιχεί σε συγκεκριμένες στοχαστικές μεταβολές. Ο πίνακας $A_{0,1}$ έχει διάσταση $1 \times K$ και τη μορφή $(\lambda, 0, 0, \dots, 0)$, περιγράφοντας τη μετάβαση από το κενό σύστημα $L=0$ στην πρώτη κατάσταση του επιπέδου $L=1$. Οι πίνακες $A_{L,L+1}$ είναι διαγώνιοι πίνακες με διάσταση $K \times K$, όπου τα πρώτα $K-L$ στοιχεία της κύριας διαγωνίου είναι ίσα με λ , αναπαριστώντας αφίξεις πελατών που γίνονται δεκτοί στο σύστημα, και όλα τα υπόλοιπα στοιχεία είναι μηδέν. Ο μI

πίνακας εκφράζει τις εξυπηρετήσεις, υποδηλώνοντας ότι η μετάβαση από το επίπεδο L στο $L-1$ συμβαίνει με ρυθμό μI , ανεξάρτητα από το πλήθος των εικονικών θέσεων.

Τέλος, οι πίνακες $A_{L,L}$ είναι άνω τριγωνικοί και περιγράφουν τις μεταβολές εντός του ίδιου L επιπέδου, σύμφωνα με τον τύπο:

$$A_{L,L} = \begin{pmatrix} -(\lambda + \mu) & 0 & & & & \\ & \ddots & \ddots & & & \\ & & -(\lambda + \mu) & 0 & & \\ & & & -(\lambda + \mu) & \lambda & \\ & & & & \ddots & \ddots \\ & & & & & -(\lambda + \mu) & \lambda \\ & & & & & & -\mu \end{pmatrix}. \quad (3.32)$$

Στην κύρια διαγώνιο του $A_{L,L}$ εμφανίζεται, κατά βάση, ο ρυθμός εξόδου $-(\lambda + \mu)$, ενώ τα στοιχεία λ πάνω από τη διαγώνιο αντιστοιχούν σε αφίξεις πελατών που λαμβάνουν εισιτήριο, αλλά αποχωρούν αμέσως. Η σημασία αυτών των στοιχείων έγκειται στο ότι περιγράφουν μεταβάσεις από την κατάσταση (L, n) στην $(L, n+1)$, όπου αυξάνεται μόνο η εικονική αναμονή, χωρίς να μεταβάλλεται ο αριθμός των πραγματικών πελατών στην ουρά. Ενώ, στην τελευταία γραμμή του πίνακα, η οποία αντιστοιχεί στην τελευταία κατάσταση του επιπέδου L , δεν υπάρχει πλέον χώρος για να αυξηθεί άλλο το n μέσα στο ίδιο επίπεδο. Επομένως, ο ρυθμός αφίξης λ απουσιάζει και το διαγώνιο στοιχείο παραμένει μόνο $-\mu$, εκφράζοντας τον ρυθμό ολοκλήρωσης της εξυπηρέτησης και λειτουργώντας ως εξισορροπητικός όρος, ώστε το άθροισμα των ρυθμών της γραμμής να ισούται με μηδέν.

Ορίζουμε ως $\mathbf{p}_0^a = p^a(0,0)$ την πιθανότητα το σύστημα να είναι κενό και ως $\mathbf{p}_L^a$ το διάνυσμα πιθανοτήτων που σχετίζεται με το επίπεδο E_L , για $L=1,2,\dots,K$. Βασιζόμενοι στην αναλυτική διαδικασία που παρουσιάστηκε στην Ενότητα 3.2, του παρόντος κεφαλαίου, επεκτείνουμε την χρήση των εξισώσεων ισορροπίας στον τροποποιημένο χώρο καταστάσεων, επιτρέποντας έτσι τον ακριβή προσδιορισμό της στάσιμης κατανομής, μέσω των ακόλουθων σχέσεων:

$$\mathbf{p}_0^a = \frac{1}{1 + \sum_{L=1}^K R_L \mathbf{e}'_K}, \quad (3.33)$$

$$\mathbf{p}_L^a = \frac{R_L}{1 + \sum_{L=1}^K R_L \mathbf{e}'_K}, \quad L=1,2,\dots,K. \quad (3.34)$$

Και σε αυτές τις εξισώσεις, ο όρος $\mathbf{e}'_K$ αντιπροσωπεύει ένα διάνυσμα μονάδων που χρησιμεύει για την άθροιση των πιθανοτήτων, διασφαλίζοντας την κανονικοποίηση του συστήματος.

Για τις ανάγκες της τροποποιημένης αλυσίδας, ορίζουμε τον πίνακα R_L , ώστε να ανταποκρίνεται στη νέα QBD δομή, εκφράζοντάς τον ως το γινόμενο των επιμέρους αναπροσαρμοσμένων πινάκων μετάβασης:

$$R_L = \prod_{l=1}^L R_{l-1,l} . \quad (3.35)$$

Οι πίνακες $R_{L-1,L}$, οι οποίοι εκφράζουν την πορεία των μεταβάσεων από το επίπεδο $L-1$ στο L , υπολογίζονται αναδρομικά. Η διαδικασία ξεκινά από το K επίπεδο, όπου το σύστημα φτάνει στη μέγιστη χωρητικότητα πραγματικών πελατών, και εξελίσσεται προς τα πίσω:

$$R_{K-1,K} = -\mu(A_{K,K})^{-1}, \quad (3.36)$$

$$R_{L-1,L} = -\mu(A_{L,L} - \mu R_{L,L+1})^{-1}, \quad L=1, 2, \dots, K-1. \quad (3.37)$$

Με αυτόν τον τρόπο η πληροφορία μεταφέρεται από την άκρη της αλυσίδας (επίπεδο K) προς τα πίσω, επιτρέποντας τον υπολογισμό όλων των πινάκων R και, κατ' επέκταση, των πιθανοτήτων των υπερ-καταστάσεων. Η ολοκλήρωση αυτής της αναδρομής αποτελεί το απαραίτητο θεμέλιο για το επόμενο στάδιο, όπου η ανάλυση περνά από τα συγκεντρωτικά επίπεδα στον λεπτομερή προσδιορισμό των πιθανοτήτων για κάθε μεμονωμένη κατάσταση εντός αυτών.

Στο πλαίσιο αυτό, ακολουθείται μία διαδικασία ανάλογη με εκείνη που αναπτύχθηκε στην ενότητα 3.3, όπου οι υπερ-καταστάσεις αποσυντίθενται στα επιμέρους στοιχεία τους. Ο χώρος καταστάσεων της τροποποιημένης ουράς αναδιατάσσεται με βάση το πλήθος των αποχωρούντων πελατών n .

Ορίζουμε τα νέα σύνολα καταστάσεων T_n , ως εξής:

$$T_0 = \{(0,0) \cup (L,0) : L=1, 2, \dots, K\},$$

$$T_n = \{(L,n) : L=1, 2, \dots, K\}, \quad n=1, 2, \dots .$$

Οι καταστάσεις εντός κάθε συνόλου T_n ταξινομούνται με αύξουσα σειρά ως προς τον αριθμό των πραγματικών πελατών L . Για κάθε επίπεδο αποχωρήσεων $n \geq 1$, η πληθικότητα του συνόλου T_n είναι σταθερή και ίση με K .

Είναι σημαντικό να επισημανθεί ότι οι στάσιμες πιθανότητες για τις καταστάσεις που ανήκουν στα σύνολα $T_0, T_1, \dots, T_{K-2}$ έχουν ήδη προσδιοριστεί μέσω της σχέσης 3.34, καθώς αυτές οι καταστάσεις ταυτίζονται με τα επίπεδα που αναλύθηκαν προηγουμένως.

Για τις υπόλοιπες περιπτώσεις, όπου $n \geq K-1$, ο υπολογισμός των πιθανοτήτων βασίζεται στον γεννήτορα πίνακα $\tilde{Q}$, που λαμβάνει την μορφή:

$$\tilde{Q} = \begin{pmatrix} & B & \lambda I^0 & \\ & & B & \lambda I \\ & & & B & \lambda I \end{pmatrix}. \quad (3.38)$$

Η δομή αυτή του $\tilde{Q}$ είναι ανάλογη με εκείνη του αρχικού μοντέλου, επιτρέποντας την εφαρμογή της ίδιας υπολογιστικής λογικής για τον προσδιορισμό των πιθανοτήτων στο διευρυμένο τμήμα της αλυσίδας, με τα επιμέρους στοιχεία του πίνακα να ανταποκρίνονται στις απαιτήσεις της τροποποιημένης ουράς.

Ο πίνακας I ορίζεται ως ο μοναδιαίος πίνακας διάστασης $K \times K$. Η παρουσία του στον πίνακα $\tilde{Q}$ υποδηλώνει ότι οι αφίξεις με ρυθμό λ προκαλούν μεταβάσεις από το σύνολο καταστάσεων T_n στο σύνολο T_{n+1} . Αντίστοιχα, ο πίνακας I^0 ταυτίζεται με τον I , με τη μόνη διαφορά ότι το πρώτο στοιχείο της κύριας διαγωνίου του είναι μηδενικό, μια ρύθμιση που αντανακλά την ιδιαιτερότητα των μεταβάσεων από το σύνολο T_0 .

Παράλληλα, ο πίνακας B είναι ο ακόλουθος κάτω τριγωνικός πίνακας διάστασης $K \times K$:

$$B = \begin{pmatrix} -(\lambda + \mu) & & & \\ \mu & -(\lambda + \mu) & & \\ & & \ddots & \\ & & \mu & -(\lambda + \mu) \\ & & & \mu & -(\lambda + \mu) \end{pmatrix}. \quad (3.39)$$

Στη διαγώνιο του παραπάνω πίνακα εμφανίζεται ο συνολικός ρυθμός εξόδου από κάθε κατάσταση, ενώ τα στοιχεία κάτω από τη διαγώνιο αντιπροσωπεύουν την ολοκλήρωση εξυπηρέτησης που οδηγεί σε μείωση των πραγματικών πελατών, εντός του ίδιου επιπέδου αποχωρήσεων n .

Θεωρώντας τους ακόλουθους πίνακες:

$$\bar{R}^0 = -\lambda I^0 B^{-1} \text{ και } \bar{R} = -\lambda B^{-1},$$

όπου ο πίνακας B έχει οριστεί από την (3.39) σχέση, οι πιθανότητες για κάθε σύνολο καταστάσεων δίνονται από τον τύπο:

$$\bar{p}_n^a = \bar{p}_{K-2}^a (-\lambda I^0 B^{-1}) (-\lambda B^{-1})^{n-K+1} = \bar{p}_{K-2}^a \bar{R}^0 \bar{R}^{n-K+1}, \quad n = K-1, K, \dots \quad (3.40)$$

Είναι φανερό ότι μέσω της παραπάνω σχέσης κάθε $\bar{p}_n^a$ μπορεί να υπολογιστεί άμεσα με βάση την τιμή του $\bar{p}_{K-2}^a$, που υπολογίζεται από την (3.34). Η ανάλυση αυτή ολοκληρώνει τον υπολογισμό των

στάσιμων πιθανοτήτων, επιτρέποντας, πλέον, τη μελέτη των μέτρων απόδοσης και την αξιολόγηση της ακρίβειας του συστήματος.

3.6 Μελέτη Απόδοσης και Χαρακτηριστικά της Συμπεριφοράς του Συστήματος

Μετά την ανάπτυξη του τροποποιημένου μοντέλου της ουράς εισιτηρίων, είναι απαραίτητο να αξιολογηθεί η αξιοπιστία της προσέγγισης σε σύγκριση με το αρχικό, πλήρες μοντέλο. Η αξιολόγηση αυτή επικεντρώνεται στην ικανότητα της προσέγγισης να αναπαράγει με ακρίβεια τις οριακές πιθανότητες και τα βασικά μέτρα απόδοσης, προσφέροντας ταυτόχρονα τη δραστική μείωση της υπολογιστικής πολυπλοκότητας, που αναλύθηκε στην προηγούμενη ενότητα.

3.6.1 Συγκριτική Ανάλυση Ακρίβειας και Δείκτες Απόκλισης

Βασισμένοι στη στάσιμη κατανομή $\{p^a(x): x \in J^a\}$, που υπολογίστηκε για το τροποποιημένο μοντέλο, είμαστε πλέον σε θέση να προσδιορίσουμε τα βασικά μέτρα απόδοσης του συστήματος. Είναι σημαντικό να σημειωθεί ότι οι ορισμοί των μέτρων αυτών στην τροποποιημένη ουρά παραμένουν αντίστοιχοι με εκείνους που αναλύθηκαν στην ενότητα 3.4, για το αρχικό μοντέλο. Ωστόσο, η χρήση των μέτρων αυτών στο προσεγγιστικό μοντέλο προϋποθέτει την εξασφάλιση της ακρίβειας των αποτελεσμάτων.

Για την επαλήθευση της απόδοσης της προσεγγιστικής μεθόδου, οι Susan H. Xu, Long Gao και Jihong Ou [[13]] πραγματοποίησαν μια εκτενή σειρά 130 αριθμητικών πειραμάτων, καλύπτοντας ένα ευρύ φάσμα παραμέτρων. Εξετάστηκαν διάφορες περιπτώσεις του ορίου ανοχής K για τιμές από 2 έως 50 και του ρυθμού κίνησης $\rho = \lambda / \mu$ από 0,1 έως 1,0. Η σύγκριση βασίστηκε στις οριακές κατανομές του πλήθους των πελατών N που παραμένουν στο σύστημα και του συνολικού αριθμού εκδοθέντων εισιτηρίων D .

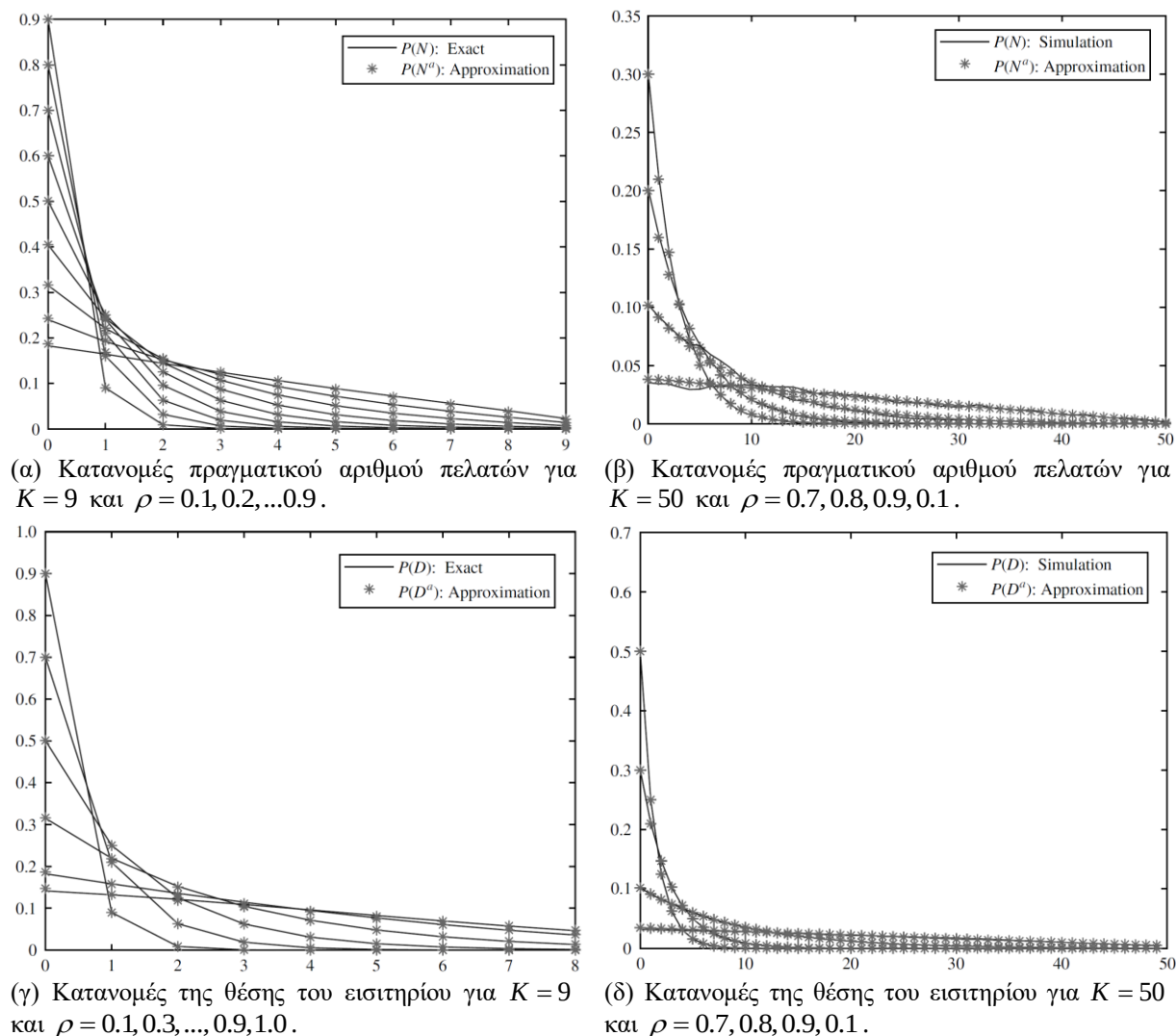
Για τις περιπτώσεις όπου το όριο ανοχής είναι μικρό, $K \leq 9$, η προσέγγιση συγκρίθηκε με τα αποτελέσματα που προέκυψαν από μελέτη του αρχικού μοντέλου. Αντίθετα, για μεγαλύτερες τιμές του ορίου, $K \geq 10$, όπου η υπολογιστική πολυπλοκότητα καθιστά την ακριβή επίλυση ανέφικτη – εξαιτίας της εκθετικής αύξησης του πλήθους των καταστάσεων στο σύστημα – χρησιμοποιήθηκαν αποτελέσματα προσομοίωσης.

Η οπτική επιβεβαίωση της ακρίβειας του προσεγγιστικού μοντέλου παρουσιάζεται στο Σχήμα 3.5, όπου συγκρίνονται οι οριακές κατανομές του αρχικού και του τροποποιημένου συστήματος.

Στο εξής, για λόγους σαφήνειας, τα αποτελέσματα που προκύπτουν από το αρχικό μοντέλο θα αναφέρονται ως “ακριβή” (exact), ενώ εκείνα που προκύπτουν από την αριθμητική αναπαράσταση του συστήματος για μεγάλες τιμές παραμέτρων ως “προσομοίωση”

(simulation). Αντιστοίχως, τα αποτελέσματα του τροποποιημένου μοντέλου θα αναφέρονται ως “προσεγγιστικά” (approximation).

Συγκεκριμένα, στα γραφήματα (α) και (β) απεικονίζεται η κατανομή των πραγματικών πελατών N, N^a για διαφορετικούς συνδυασμούς των K και ρ , με τον οριζόντιο άξονα να διατρέχει τιμές του L , ενώ στα γραφήματα (γ) και (δ) παρουσιάζεται η αντίστοιχη κατανομή για το συνολικό πλήθος D, D^a των εισιτηρίων στο σύστημα.



Σχήμα 3.5: Σύγκριση οριακών κατανομών μεταξύ της αρχικής και της τροποποιημένης ουράς

Όπως παρατηρείται στα παραπάνω διαγράμματα, σε όλες τις περιπτώσεις, τα αποτελέσματα της προσέγγισης (σημεία) ταυτίζονται σχεδόν πλήρως με αυτά της ακριβούς επίλυσης (γραμμές) ή της προσομοίωσης αυτής. Αξίζει να σημειωθεί ότι, σύμφωνα με τα ευρήματα που αναφέρονται στο έργο [[13]], η εξαιρετική αυτή ακρίβεια διατηρείται σε όλο το εύρος των παραμέτρων που εξετάστηκαν. Ακόμα και στην οριακή περίπτωση, δηλαδή σε συνθήκες

υψηλού φόρτου, όπου ο ρυθμός κίνησης φτάνει στη μονάδα, $\rho \rightarrow 1$, η προσέγγιση παραμένει αξιόπιστη, παρά την αυξημένη πίεση που δέχεται το σύστημα.

Για την ποσοτικοποίηση της απόκλισης μεταξύ των οριακών κατανομών του ακριβούς μοντέλου και της προσεγγιστικής μορφής αυτού, με το τροποποιημένο μοντέλο, χρησιμοποιούνται οι δείκτες της Μέγιστης Απόλυτης Διαφοράς – ΜΑΔ (Maximum Absolute Difference – MAD) και της Συνολικής Απόλυτης Διαφοράς – ΣΑΔ (Total Absolute Difference – TAD), οι οποίοι ορίζονται ως:

$$ΜΑΔ = \max_n \left| P_{ακριβής}(N=n) - P_{προσεγγιστική}(N=n) \right|,$$

$$ΣΑΔ = \sum_n \left| P_{ακριβής}(N=n) - P_{προσεγγιστική}(N=n) \right|.$$

Πέρα από την οπτική σύγκριση που προηγήθηκε, με την βοήθεια του Σχήματος 3.5, η ακρίβεια του μοντέλου τεκμηριώνεται ποσοτικά μέσω των παραπάνω δεικτών. Στον Πίνακα 3.1, που ακολουθεί, παρουσιάζονται οι 10 δυσμενέστερες περιπτώσεις απόκλισης των μελετημένων ποσοτήτων, από το σύνολο των 130 αριθμητικών πειραμάτων που πραγματοποιήθηκαν.

Κ	ρ	P(N)		P(D)	
		ΣΑΔ	ΜΑΔ	ΣΑΔ	ΜΑΔ
6	1.	0.02297	0.00514	0.0168	0.00514
7	1.	0.02556	0.00542	0.02037	0.00542
8	1.	0.03041	0.00592	0.02076	0.00592
9	0.9	0.02151	0.00445	0.02238	0.00449
9	1.	0.03292	0.00588	0.02397	0.0058
10	1.	0.03434	0.00591	0.02544	0.00591
20	1.	0.04667	0.00534	0.0338	0.00528
30	1.	0.06411	0.00453	0.04265	0.00451
40	1.	0.07508	0.00402	0.05729	0.00382
50	1.	0.06489	0.00307	0.04699	0.00303
Μέγιστη απόκλιση 130 δοκιμών		0.07508	0.00592	0.05729	0.00592
Μέσος όρος αποκλίσεων 130 δοκιμών		0.00812	0.00141	0.00774	0.00179

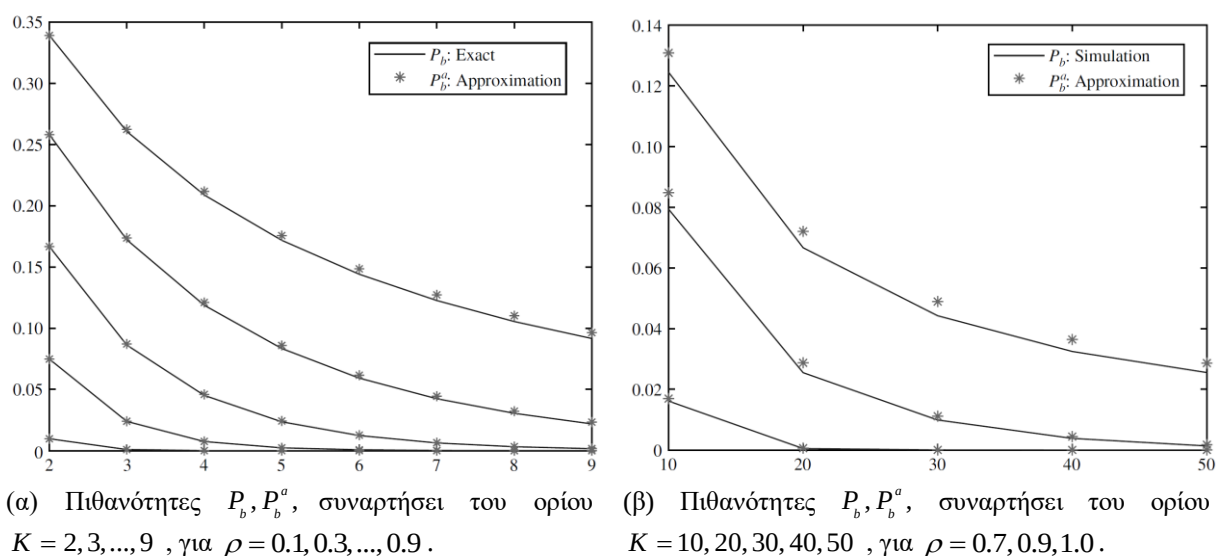
Πίνακας 3.1: Σύγκριση των N, N^a και D, D^a

Παρατηρούμε ότι ακόμη και σε ακραίες περιπτώσεις η τιμή της ΣΑΔ δεν ξεπερνά το 0,08, ενώ η τιμή της ΜΑΔ παραμένει, για όλες τις δοκιμές των K και ρ , σταθερά μικρότερη από

το 0,06. Επιπλέον, στην τελευταία γραμμή του πίνακα προβάλλονται οι μέσες τιμές των δεικτών αυτών, οι οποίες παραμένουν εξαιρετικά χαμηλές, γεγονός που επιβεβαιώνει ότι η προσέγγιση είναι στατιστικά αξιόπιστη και ασφαλής για τη μελέτη και την ανάλυση του συστήματος.

Η αξιοπιστία της προσεγγιστικής μεθόδου επεκτείνεται και στον υπολογισμό των κρίσιμων παραμέτρων λειτουργίας του συστήματος, όπως η πιθανότητα απομάκρυνσης των πελατών λόγω αναμονής P_b και ο μέσος αριθμός πελατών στο σύστημα $E[N]$.

Ειδικότερα, η σύγκριση των τιμών της πιθανότητας P_b μεταξύ του αρχικού και του τροποποιημένου μοντέλου, για το σύνολο των 130 δοκιμών, κατέδειξε ότι η απόκλιση (εκφρασμένη μέσω των δεικτών ΜΑΔ και ΣΑΔ) κυμαίνεται σε εξαιρετικά χαμηλά επίπεδα. Σύμφωνα με τα αποτελέσματα των πειραματικών δεδομένων [[13]], η μέση τιμή της απόκλισης για την πιθανότητα P_b ανέρχεται στο 0,00196, γεγονός που υπογραμμίζει τη συνέπεια της προσέγγισης. Αυτό αποτυπώνεται οπτικά στο Σχήμα 3.6, όπου η προσέγγιση ακολουθεί με ακρίβεια την πορεία της ακριβούς επίλυσης (για μικρές τιμές του K) και της προσομοίωσης (για μεγάλες τιμές του K) υπό διαφορετικές συνθήκες φόρτου ρ .

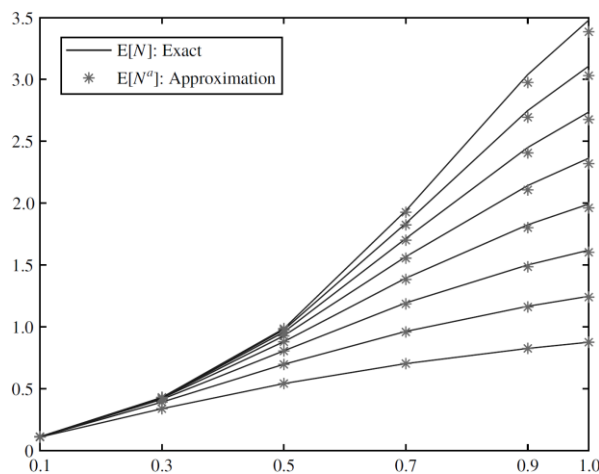


Σχήμα 3.6: Σύγκριση πιθανότητας απομάκρυνσης

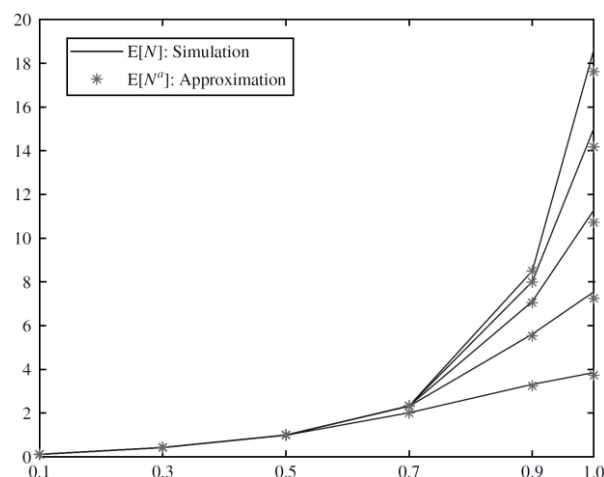
Επιπλέον, για την αξιολόγηση του μέσου πλήθους πελατών στο σύστημα, αξιοποιείται ο δείκτης Απόλυτου Ποσοστιαίου Σφάλματος – ΑΠΣ (Absolute Percentage Error - APE), ο οποίος ορίζεται ως:

$$\text{ΑΠΣ} = \frac{|E[N^a] - E[N]|}{E[N]} \times 100\% ,$$

όπου $E[N^a]$ είναι η προσεγγιστική τιμή και $E[N]$ η ακριβής. Όπως προκύπτει από την ίδια πειραματική μελέτη, το μέσο σφάλμα είναι 1,006%, ενώ ακόμα και η μέγιστη τιμή του δεν υπερβαίνει το 5,619%. Αυτό αποτυπώνεται οπτικά και στο Σχήμα 3.7, όπου η προσέγγιση διατηρεί την ευστάθειά της σε όλο το φάσμα των τιμών του ρυθμού κίνησης, επιβεβαιώνοντας ότι η μέθοδος είναι αξιόπιστη ακόμα και σε συνθήκες έντονου φόρτου.



(α) Αναμενόμενος αριθμός πελατών, συναρτήσει του ρυθμού κίνησης ρ , για $K = 2, \dots, 9$.



(β) Αναμενόμενος αριθμός πελατών, συναρτήσει του ρυθμού κίνησης ρ , για $K = 10, \dots, 50$.

Σχήμα 3.7: Σύγκριση αναμενόμενου αριθμού πελατών στο σύστημα

Επισημαίνεται ότι ο δείκτης ΑΠΣ, παρότι αποτελεί ένα κατατοπιστικό μέτρο σύγκρισης για αναμενόμενες τιμές, αποφεύχθηκε στην ανάλυση των πιθανοτήτων των κατανομών που προσεγγίζουν ταχύτατα το μηδέν. Σε τέτοιες περιπτώσεις, ο εξαιρετικά μικρός παρονομαστής μπορεί να οδηγήσει σε αριθμητική αστάθεια, καθιστώντας το αποτέλεσμα παραπλανητικό.

3.6.2 Ποιοτική Αξιολόγηση και Ιδιότητες των Βασικών Μέτρων Απόδοσης

Στη συνέχεια της ανάλυσης, πέρα από τον έλεγχο της αριθμητικής ακρίβειας, εξετάζονται ορισμένα πρόσθετα ευρήματα που προέκυψαν από την διερεύνηση της μελέτης προέλευσης [[13]], σχετικά με τη δομική σχέση μεταξύ των μεγεθών του αρχικού και του τροποποιημένου μοντέλου. Με βάση τα αποτελέσματα των αριθμητικών δοκιμών που πραγματοποιήθηκαν, στο πλαίσιο της εν λόγω έρευνας, διατυπώνεται η εικασία ότι το μέγεθος N^a αποτελεί ένα στοχαστικό κάτω φράγμα του πραγματικού αριθμού πελατών N στο σύστημα. Αυτό πρακτικά σημαίνει ότι η προσεγγιστική τιμή N^a παραμένει πάντα σε χαμηλότερα επίπεδα από την πραγματική τιμή N , υποεκτιμώντας ελαφρώς το πλήθος των πελατών.

Η υπόθεση αυτή εδράζεται στο γεγονός ότι για όλους τους συνδυασμούς παραμέτρων που εξετάστηκαν ($K = 1, 2, \dots, 9$ και $\rho = 0, 1, \dots, 1, 0$), η διαφορά $P(N \geq n) - P(N^a \geq n)$ των πιθανοτήτων παραμένει σταθερά μη αρνητική, με την απόκλιση να αυξάνεται όσο

μεγαλώνουν οι τιμές των K και ρ . Παρόλο που τα αριθμητικά δεδομένα υποστηρίζουν σθεναρά αυτή τη διαπίστωση, θα πρέπει να επισημανθεί ότι παραμένει σε επίπεδο υποθέσεως, καθώς δεν έχει διατυπωθεί ακόμα μία αυστηρή αναλυτική απόδειξη. Συνεπώς, η παρατήρηση αυτή αποτελεί μια ισχυρή πειραματική ένδειξη της στοχαστικής συμπεριφοράς του συστήματος.

Ένα ιδιαίτερο εύρημα που προκύπτει από τη μελέτη του τροποποιημένου μοντέλου είναι η ανάδειξη μίας σχεδόν γραμμικής σχέσης μεταξύ του αριθμού εισιτηρίου που κατέχει ένας πελάτης και της πραγματικής του θέσης στην ουρά αναμονής.

Στο προσεγγιστικό μοντέλο, η θέση εισιτηρίου καθορίζεται από το άθροισμα των πραγματικών πελατών L και των εικονικών θέσεων n . Στο χώρο καταστάσεων που απεικονίζεται στο Σχήμα 3.3, όλες οι καταστάσεις (L, n) που έχουν το ίδιο άθροισμα $L + n = d$ σχηματίζουν διαγώνιες διατάξεις. Για παράδειγμα, η θέση εισιτηρίου $d = 4$ περιλαμβάνει όλες τις πιθανές καταστάσεις όπου το συνολικό πλήθος πραγματικών και εικονικών πελατών ισούται με 4, δηλαδή τις καταστάσεις $(1, 3)$, $(2, 2)$, $(3, 1)$, $(4, 0)$.

Για να υπολογίσουμε την πιθανότητα το σύστημα να βρίσκεται σε μία τέτοια διαγώνιο – δηλαδή να ισχύει το ενδεχόμενο $\{D^a = d\}$ – αθροίζουμε τις πιθανότητες όλων των καταστάσεων (l, n) που ικανοποιούν τη συνθήκη $l + n = d$. Ουσιαστικά, χρησιμοποιούμε το άθροισμα των επιμέρους πιθανοτήτων για κάθε τιμή l , από 1 έως την ανώτατη επιτρεπτή τιμή του αθροίσματος, που μπορεί να πάρει το πλήθος των πραγματικών πελατών:

$$P(D^a = d) = \sum_{l=1}^{d \wedge K} p^a(l, d-l). \quad (3.41)$$

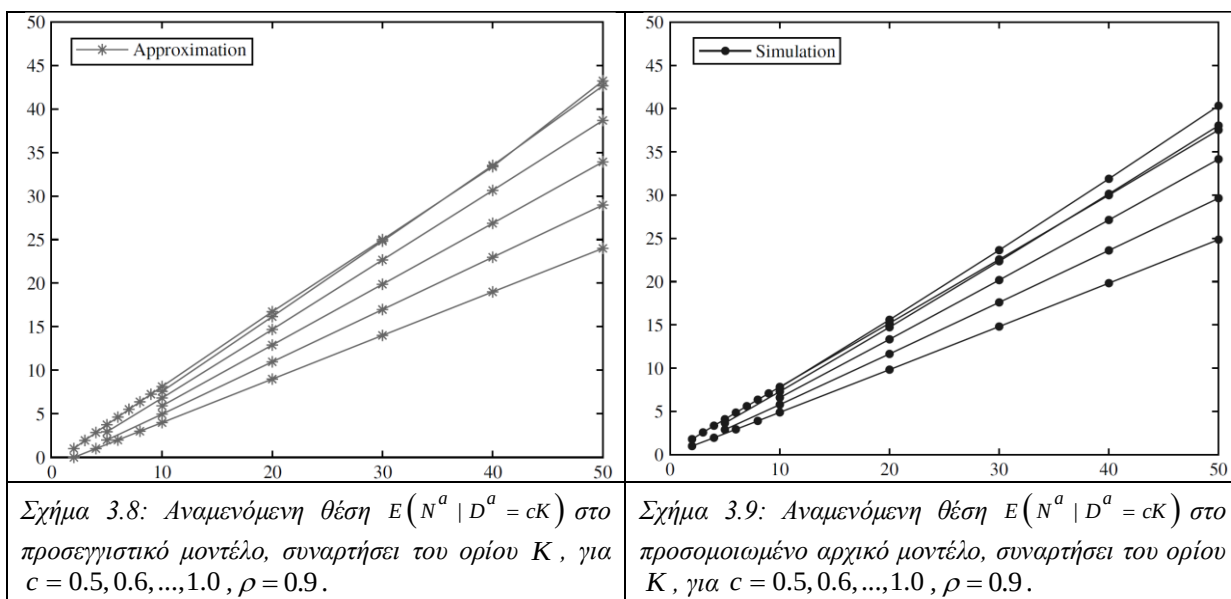
Έπειτα, η αναμενόμενη θέση του πραγματικού πελάτη στην ουρά, δεδομένου ότι η θέση του εισιτηρίου είναι η d , υπολογίζεται ως:

$$E(N^a | D^a = d) = \frac{\sum_{l=1}^{d \wedge K} l \cdot p^a(l, d-l)}{P(D^a = d)}. \quad (3.42)$$

Από την ανάλυση προκύπτει ότι αν διατηρήσουμε τη θέση του εισιτηρίου ως ένα σταθερό ποσοστό c του ορίου K (δηλαδή $d = c \cdot K$), τότε η αναμενόμενη θέση του πελάτη $E(N^a | D^a = cK)$ αυξάνεται με σταθερό γραμμικό ρυθμό, καθώς το K μεγαλώνει.

Η εγκυρότητα της προσεγγιστικά γραμμικής σχέσης επιβεβαιώνεται μέσω αριθμητικών υπολογισμών για ένα ευρύ φάσμα παραμέτρων, με το c να λαμβάνει τιμές από 0,5 έως 1,0 και για διάφορες τιμές των K και ρ . Ενδεικτικά, στο Σχήμα 3.8 παρουσιάζεται η γραμμική εξάρτηση της αναμενόμενης τιμής $E(N^a | D^a = cK)$ ως προς το όριο K , για ρυθμό $\rho = 0,9$.

Παρατηρείται ότι η γραμμικότητα αυτή διατηρείται σταθερή για όλες τις τιμές του c που εξετάστηκαν, εύρημα που επαληθεύεται και για τις υπόλοιπες τιμές του ρυθμού ρ στις αριθμητικές δοκιμές. Επιπλέον, προκειμένου να διαπιστωθεί αν η ιδιότητα αυτή αποτελεί χαρακτηριστικό του πραγματικού συστήματος και όχι κατασκευασμένο στοιχείο της προσέγγισης, πραγματοποιήθηκε αριθμητική πειραματική διαδικασία για τις ίδιες παραμέτρους και στο αρχικό μοντέλο. Τα αποτελέσματα αυτής, που αποτυπώνονται στο Σχήμα 3.9, επικυρώνουν πλήρως την ύπαρξη της γραμμικότητας και στο αρχικό σύστημα, με την αναμενόμενη τιμή $E(N^a | D^a = cK)$ να είναι γραμμική ως προς το όριο K , για ρυθμό $\rho = 0,9$.



Η παρατήρηση αυτή είναι ιδιαίτερα σημαντική, καθώς υποδηλώνει ότι η διαχείριση του συστήματος μπορεί να βασιστεί σε αυτή τη σχέση για τη διεξαγωγή συμπερασμάτων. Ειδικότερα, μπορεί να εκτιμηθεί πώς επηρεάζεται η αναμενόμενη θέση του πελάτη που βρίσκεται στο όριο της εισόδου (όταν $c=1$ το εισιτήριό του παίρνει την τιμή του ορίου K), καθώς μεταβάλλονται οι παράμετροι του συστήματος.

Επιπρόσθετα, η ανάλυση ανέδειξε ορισμένες αναμενόμενες αλλά ουσιαστικές συμπεριφορές του συστήματος, οι οποίες συνοψίζονται στα εξής:

- Για μικρές τιμές του c , η αναμενόμενη πραγματική θέση προσεγγίζει την τιμή του εισιτηρίου $E(N^a | D^a = cK) \approx cK$. Αυτό συμβαίνει επειδή, όταν ένας πελάτης βρίσκεται πολύ κοντά στην αρχή της ουράς, η πιθανότητα να έχουν αποχωρήσει οι ελάχιστοι προπορευόμενοι πελάτες είναι πολύ μικρή. Επομένως, στην πλειονότητα οι πελάτες είναι παρόντες και η εικονική θέση ταυτίζεται σχεδόν με την πραγματική.
- Παρατηρείται ότι η αναμενόμενη θέση $E(N^a | D^a = cK)$ τείνει να μειώνεται καθώς αυξάνεται η ένταση της κίνησης (traffic intensity) ρ . Αν και με μία γρήγορη σκέψη θα περίμενε κανείς μεγαλύτερη ουρά, λόγω του αυξημένου φόρτου, το αποτέλεσμα

αυτό εξηγείται από το γεγονός ότι η υψηλότερη ένταση κίνησης συνεπάγεται μεγαλύτερο ποσοστό αποχωρήσεων. Οι συχνότερες αποχωρήσεις δημιουργούν περισσότερες κενές θέσεις στην ουρά, μειώνοντας τελικά τον αριθμό των ατόμων που περιμένουν πραγματικά στην ουρά. Αυτό έχει ως αποτέλεσμα ένας πελάτης με συγκεκριμένο αριθμό εισιτηρίου να βρίσκεται πολύ πιο κοντά στην εξυπηρέτηση από το αναμενόμενο, λόγω των κενών θέσεων που άφησαν όσοι αποχώρησαν.

Η δυνατότητα εύκολου υπολογισμού των παραπάνω στοιχείων επιτρέπει στον διαχειριστή να αντλήσει πολύτιμα συμπεράσματα για τη συμπεριφορά του συστήματος. Η γνώση αυτή είναι καθοριστική και μπορεί να αξιοποιηθεί έμπρακτα αν παρασχεθεί ως πληροφορία στους εισερχόμενους πελάτες του συστήματος. Με τον τρόπο αυτό, οι πελάτες αποκτούν μία σαφέστερη εικόνα για τον αναμενόμενο χρόνο αναμονής τους, γεγονός που συμβάλλει δραστικά στη μείωση των ποσοστών πρόωρης εγκατάλειψης του συστήματος.

3.7 Βελτιστοποίηση της Απόδοσης των Υπηρεσιών του Συστήματος Εισιτηρίων

Όπως αναλύθηκε και στην πρώτη ενότητα του παρόντος κεφαλαίου, η χρήση του συστήματος εισιτηρίων προσφέρει σημαντικές διευκολύνσεις, αλλά παρουσιάζει το βασικό μειονέκτημα της έλλειψης πλήρους διαφάνειας ως προς την πραγματική κατάσταση του συστήματος. Στην πράξη, τόσο ο πάροχος της υπηρεσίας όσο και οι πελάτες δεν έχουν μία ξεκάθαρη εικόνα για το πόσα άτομα περιμένουν ενεργά στην ουρά. Καθώς ο αριθμός του εισιτηρίου δεν πληροφορεί για τις προηγούμενες αποχωρήσεις, δημιουργείται σημαντική αβεβαιότητα στη λήψη αποφάσεων, η οποία ενδέχεται να οδηγήσει σε λανθασμένες εκτιμήσεις σχετικά με τον χρόνο αναμονής. Το φαινόμενο αυτό και οι επιπτώσεις του στη λειτουργία της υπηρεσίας αποτελούν τον κεντρικό άξονα της ανάλυσης που ακολουθεί.

Στο σημείο αυτό γεννώνται ορισμένοι βασικοί προβληματισμοί που σχετίζονται με την αποτελεσματικότητα ενός τέτοιου συστήματος. Ειδικότερα, ανακύπτουν ερωτήματα ως προς το κάτω από ποιες συνθήκες αυτή η απώλεια πληροφορίας εμποδίζει έναν πελάτη από το να πάρει μία σωστή απόφαση, οδηγώντας τελικά σε πτώση της ποιότητας εξυπηρέτησης. Παράλληλα, αξίζει να εκτιμηθεί η σχέση μεταξύ του φόρτου εργασίας του συστήματος και της ανυπομονησίας των πελατών. Επιπλέον, έχει σημασία να εξεταστεί το μέγεθος αυτής της αλληλεπίδρασης, καθώς και σε τί βαθμό τελικά ωθεί τους πελάτες στο να αποχωρήσουν. Τέλος, θεωρείται ιδιαίτερα χρήσιμη η ανάδειξη συγκεκριμένων τρόπων ώστε η απόδοση της ουράς να βελτιωθεί, εξασφαλίζοντας οφέλη για τους πελάτες και, κατ' επέκταση, για την ομαλή λειτουργία των παρεχόμενων υπηρεσιών.

Για να διευκολυνθεί η μελέτη αυτών των ζητούμενων, υιοθετούμε την παραδοχή ότι οι πελάτες λαμβάνουν την απόφαση να εισέλθουν ή όχι στο σύστημα βασιζόμενοι αποκλειστικά στον κανόνα που ορίζει ότι παραμένουν μόνο εφόσον η θέση του εισιτηρίου τους D είναι μικρότερη από το όριο ανοχής τους K . Η προσέγγιση αυτή μας επιτρέπει τον παραλληλισμό της ουράς των εισιτηρίων με το μοντέλο της $M/M/1/K$ ουράς, στην οποία το K λειτουργεί ως το ανώτατο όριο

χωρητικότητας σε συνθήκες πλήρους πληροφόρησης. Αυτή η επιλογή ακολουθήθηκε προς όφελος της ανάλυσης, ώστε να καταστεί δυνατή η σύγκριση ανάμεσα στην πλήρη πληροφόρηση της φυσικής ουράς και στην περιορισμένη ενημέρωση του συστήματος εισιτηρίων. Έτσι, αναδεικνύεται πως η έλλειψη πληροφόρησης επηρεάζει τις αποφάσεις των πελατών και την τελική απόδοση της ουράς.

Παρόλο που αυτή η παραδοχή δεν λαμβάνει υπόψη κάθε πιθανή αντίδραση ενός πελάτη, εντούτοις, θεωρούμε ότι είναι ικανή να αποκαλύψει τη βασική στοχαστική συμπεριφορά του συστήματος. Μέσα από αυτή τη διαφορετική αναπλαισίωση της μελέτης μπορούν να προκύψουν χρήσιμα συμπεράσματα για τον πάροχο, ώστε να βελτιώσει την υπηρεσία του. Η ανάλυση αυτή οδηγεί σε μία αναβαθμισμένη μορφή του συστήματος εισιτηρίων που βοηθάει τον πελάτη να αποφασίσει σωστά και ταυτόχρονα περιορίζει στο ελάχιστο τη διαφορά στην απόδοση ανάμεσα στην ουρά με εισιτήριο και τη φυσική ουρά αναμονής.

3.7.1. Στοχαστική και Αριθμητική Συγκριτική Μελέτη: Ουρές με Εισιτήρια (Απλή και Τροποποιημένη) έναντι της Φυσικής Ουράς $M/M/1/K$

Προκειμένου να αναλυθεί διεξοδικά η επίδραση της έλλειψης πληροφόρησης στην απόδοση του συστήματος, κρίνεται σκόπιμη η εξέταση της ουράς εισιτηρίων μέσω της σύγκρισής της με το πρότυπο μοντέλο της φυσικής ουράς $M/M/1/K$. Στο πλαίσιο αυτό, εξετάζονται τρία διαφορετικά σενάρια λειτουργίας:

- Ουρά Εισιτηρίων (N, W, P_b) : Αφορά το βασικό μοντέλο όπου ο πελάτης λαμβάνει αποφάσεις γνωρίζοντας μόνο τη θέση που αναγράφεται στο εισιτήριό του.
- Τροποποιημένη Ουρά Εισιτηρίων (N^a, W^a, P_b^a) : Αφορά την αναβαθμισμένη εκδοχή της οποίας η εκτενής μελέτη έγινε στην υποενότητα 3.5.
- Φυσική Ουρά (N^p, W^p, P_b^p) : Αντιστοιχεί στο μοντέλο $M/M/1/K$, όπου ο πελάτης έχει πλήρη εικόνα του ακριβούς αριθμού των παρόντων ατόμων.

Σε κάθε περίπτωση, η παράμετρος N αντιπροσωπεύει τον αριθμό των πελατών που βρίσκονται στο σύστημα σε κατάσταση ισορροπίας, το W τους μέσους χρόνους αναμονής και το P_b την πιθανότητα αποχώρησης.

Η σύγκριση αυτή αναδεικνύει μια αξιοσημείωτη ιδιαιτερότητα στην λειτουργία των ουρών αναμονής με εισιτήρια σε σχέση με την φυσική ουρά, η οποία δημιουργεί ένα μη αναμενόμενο παράδοξο. Μολονότι τα συστήματα αυτά παρουσιάζουν χαμηλότερα επίπεδα συμφόρησης σε σχέση με την φυσική ουρά, εντούτοις, εμφανίζουν υψηλότερη πιθανότητα αποχώρησης. Αυτό συμβαίνει διότι οι πελάτες, λόγω της ελλιπούς ενημέρωσης για την πραγματική κατάσταση της αναμονής, αποθαρρύνονται από τον αριθμό του εισιτηρίου τους και εγκαταλείπουν το σύστημα. Με τον τρόπο αυτό, ο πάροχος στερείται της δυνατότητας

εξυπηρέτησης ατόμων που θα μπορούσαν να είχαν παραμείνει, αν γνώριζαν τον πραγματικό φόρτο εργασίας.

Η μαθηματική τεκμηρίωση αυτού του ισχυρισμού συνοψίζεται στην ακόλουθη Πρόταση 3.1, της οποίας η αναλυτική απόδειξη παραλείπεται. Μέσω αυτής της πρότασης, παρουσιάζεται η συγκριτική σχέση μεταξύ των παραμέτρων των διαφορετικών προς μελέτη μοντέλων.

Πρόταση 3.1: Θεωρούμε την ουρά εισιτηρίων, την τροποποιημένη ουρά εισιτηρίων και την φυσική ουρά, οι οποίες διέπονται από τον ίδιο ρυθμό αφίξεων λ , τον ίδιο ρυθμό εξυπηρέτησης μ και το ίδιο όριο αποχώρησης K . Τότε ισχύουν οι ακόλουθες σχέσεις:

- I. $P(N \leq n) \geq P(N^p \leq n)$ και $P(N^a \leq n) \geq P(N^p \leq n)$. Συνεπώς, συμπεραίνουμε ότι $E[N] \leq E[N^p]$ και $E[N^a] \leq E[N^p]$.
- II. $P_b \geq P_b^p$ και $P_b^a \geq P_b^p$.

Οι σχέσεις της παραπάνω πρότασης επιβεβαιώνουν και μαθηματικά την προαναφερθείσα ιδιαιτερότητα του συστήματος. Στο σκέλος I., οι ανισότητες των πιθανοτήτων υποδηλώνουν ότι το σύστημα της ουράς εισιτηρίων, τόσο στην απλή του μορφή, όσο και στην τροποποιημένη, παρουσιάζει μία συστηματική τάση να λειτουργεί με λιγότερους πελάτες σε σχέση με την φυσική ουρά. Συγκεκριμένα, η αυξημένη πιθανότητα να εντοπιστεί μικρός αριθμός πελατών στο σύστημα εισιτηρίων οδηγεί στο συμπέρασμα ότι ο μέσος αριθμός παρόντων πελατών θα είναι χαμηλότερος. Παράλληλα, το σκέλος II. καταδεικνύει ότι η πιθανότητα αποχώρησης παραμένει συστηματικά υψηλότερη και στις δύο μορφές του συστήματος εισιτηρίων, συγκριτικά, με την φυσική ουρά. Η συνδυαστική μελέτη αυτών των δύο σχέσεων τεκμηριώνει το γεγονός ότι η έλλειψη πληροφόρησης οδηγεί σε μία υπολειτουργία του συστήματος, όπου ο φόρτος είναι χαμηλότερος ακριβώς επειδή οι πελάτες αποθαρρύνονται ευκολότερα.

Τώρα, δεδομένου ότι ο χρόνος αναμονής W ορίζεται ως το άθροισμα των χρόνων εξυπηρέτησης T_n των παρόντων πελατών, όπου $T_n, n \geq 1$, είναι ανεξάρτητες και ισόνομες εκθετικές τυχαίες μεταβλητές, δηλαδή $W = \sum_{n=1}^N T_n$, $W^a = \sum_{n=1}^{N^a} T_n$ και $W^p = \sum_{n=1}^{N^p} T_n$, προκύπτουν περαιτέρω στοχαστικές ανισότητες, που αφορούν τους χρόνους αναμονής:

$$P(W \leq w) \geq P(W^p \leq w) \text{ και } P(W^a \leq w) \geq P(W^p \leq w), \quad (3.43)$$

$$E[W] \leq E[W^p] \text{ και } E[W^a] \leq E[W^p]. \quad (3.44)$$

Αξίζει να σημειωθεί ότι η Πρόταση 3.1 και οι σχέσεις που απορρέουν από αυτήν, αποτελούν κρίσιμα εργαλεία για την ανάλυση του συστήματος. Μας επιτρέπουν να εξάγουμε με απλό τρόπο ασφαλή φράγματα για τα βασικά μεγέθη απόδοσης, προσφέροντας μία αξιόπιστη προσέγγιση σε περιπτώσεις όπου ο ακριβής υπολογισμός τους καθίσταται δυσχερής ή και δυσεπίλυτος.

Επεκτείνοντας την ανάλυση στους χρόνους αναμονής, οι σχέσεις (3.43) και (3.44) επιβεβαιώνουν ότι η ίδια λογική ισχύει και για τον χρόνο που παραμένει ο πελάτης στο σύστημα. Δεδομένου ότι ο χρόνος αναμονής W εξαρτάται από τον αριθμό των παρόντων πελατών, αναδεικνύεται ότι τόσο η στοχαστική συμπεριφορά όσο και η μέση τιμή $E[W]$ της αναμονής είναι χαμηλότερες στα συστήματα εισιτηρίων σε σύγκριση με την φυσική ουρά. Το αποτέλεσμα αυτό συμπληρώνει την εικόνα του συστήματος, επαληθεύοντας ότι οι αυξημένες αποχωρήσεις οδηγούν σε μικρότερους χρόνους αναμονής, για όσους τελικά εξυπηρετούνται.

Για την περαιτέρω ποσοτικοποίηση της απόκλισης μεταξύ των δύο συστημάτων και την αξιολόγηση της ακρίβειας των παραπάνω σχέσεων, εισάγονται οι δείκτες της Απόλυτης Διαφοράς – ΑΔ (Absolute Difference – TAD) και του Ποσοστιαίου Σφάλματος ΠΣ (Percentage Error – APE). Στο σημείο αυτό, οι δείκτες αυτοί εστιάζουν ειδικά στην απόκλιση της πιθανότητας αποχώρησης της τροποποιημένης ουράς εισιτηρίων σε σχέση με εκείνη της φυσικής ουράς. Οι πιθανότητες αυτές αποτελούν τα κρίσιμα μέτρα απόδοσης του συστήματος και οι δείκτες ορίζονται ως εξής:

$$ΑΔ = P_b^a - P_b^p,$$

$$ΠΣ = \frac{P_b^a - P_b^p}{P_b^p} \times 100\%.$$

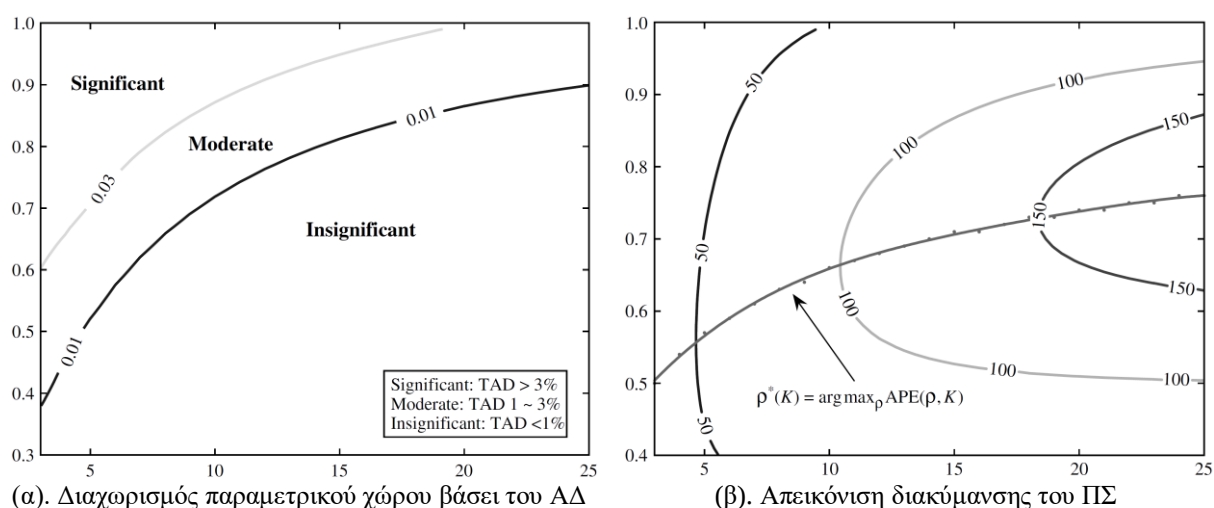
Οι εν λόγω δείκτες αξιοποιήθηκαν στη μελέτη [[13]] για την ανάλυση ενός ευρέος φάσματος παραμέτρων, που καλύπτει τις περιπτώσεις $2 \leq K \leq 50$ και $0 < \rho \leq 1$. Στα διαγράμματα του Σχήματος 3.10, τα οποία έχουν ληφθεί από την ίδια βιβλιογραφική πηγή, ο οριζόντιος άξονας αντιστοιχεί στο όριο αποχώρησης K και ο κατακόρυφος στον ρυθμό κίνησης ρ .

Συγκεκριμένα, στο Σχήμα 3.10 (α), βλέπουμε ότι με βάση τις τιμές ΑΔ, ο παραμετρικός χώρος διαχωρίζεται σε τρεις διακριτές περιοχές:

- **Περιοχή Σημαντικής Απόκλισης (Significant Region):** Περιλαμβάνει τις περιπτώσεις όπου ο ΑΔ υπερβαίνει το 3% (φτάνοντας έως 6%, σύμφωνα με τα αριθμητικά στοιχεία του Πίνακα 3.2, ο οποίος θα παρουσιαστεί παρακάτω). Όπως παρατηρείται στο Σχήμα 3.10 (α), η περιοχή αυτή εντοπίζεται ψηλά και προς τα αριστερά, για συνδυασμούς χαμηλού ορίου K και υψηλού ρυθμού κίνησης ρ .
- **Περιοχή Μέτριας Απόκλισης (Moderate Region):** Περιλαμβάνει τις περιπτώσεις όπου ο ΑΔ βρίσκεται ανάμεσα στο 1% και στο 3%, καταλαμβάνοντας τη ζώνη στην μέση του διαγράμματος.
- **Περιοχή Ασήμαντης Απόκλισης (Insignificant Region):** Περιλαμβάνει τις περιπτώσεις όπου ο ΑΔ είναι μικρότερος από το 1%. Όπως διακρίνεται στο Σχήμα 3.10 (α), η περιοχή αυτή βρίσκεται κάτω και δεξιά, καλύπτοντας το μεγαλύτερο φάσμα τιμών του παραμετρικού χώρου, ιδιαίτερα καθώς αυξάνεται το όριο K .

Κοιτάζοντας, λοιπόν, το γράφημα (α) του παρακάτω σχήματος, παρατηρούμε ότι τα όρια των περιοχών αυτών καθορίζονται από την δυναμική αλληλεπίδραση του K και του ρ . Όταν το σύστημα κινείται σε χαμηλές τιμές K , αλλά έχει υψηλό ρυθμό ρ , η απόκλιση ξεπερνά το 3%, που σημαίνει ότι υπάρχει μεγάλη ευαισθησία και τάση για απομάκρυνση. Από την άλλη, καθώς κινούμαστε προς τα δεξιά (όπου το όριο ανοχής K είναι μεγαλύτερο), το σύστημα μεταβαίνει σχεδόν αποκλειστικά στην ζώνη της ασήμαντης απόκλισης. Αυτό που καταλαβαίνουμε είναι ότι όταν οι πελάτες διαθέτουν μεγαλύτερη ανοχή στην αναμονή, δηλαδή είναι πιο υπομονετικοί, θα έχουμε περισσότερο κόσμο στην ουρά και έτσι για μεγάλο όριο K οι αποκλίσεις γίνονται αμελητέες, ανεξάρτητα από τις διακυμάνσεις του φόρτου. Η ανεξαρτησία αυτή οφείλεται στο γεγονός ότι στο πάνω και δεξί άκρο του γραφήματος δεν εντοπίζεται καθόλου περιοχή σημαντικής απόκλισης.

Το γενικό συμπέρασμα που λαμβάνουμε είναι ότι η υπομονή των πελατών καθορίζει τη δομή του παραμετρικού χώρου, δείχνοντας ότι όσο πιο μεγάλο όριο ανοχής έχουν οι πελάτες, τόσο σταθεροποιείται το σύστημα και ελαχιστοποιείται η Απόλυτη Διαφορά.



Σχήμα 3.10: Συγκριτική Απεικόνιση Δεικτών Απόκλισης

Το παραπάνω συμπέρασμα συμπληρώνεται οπτικά από το Σχήμα 3.10 (β), το οποίο δείχνει την μεταβολή του ποσοστιαίου σφάλματος στον παραμετρικό χώρο. Στο διάγραμμα αυτό, οι καμπύλες γραμμές ενώνουν τα σημεία εκείνα στα οποία το σφάλμα διατηρεί την ίδια τιμή (όπως 50%, 100%, 150%).

Ιδιαίτερο ενδιαφέρον παρουσιάζει η καμπύλη ρ^* , η οποία διατρέχει το γράφημα και ορίζεται από τη σχέση:

$$\rho^*(K) = \arg \max_{\rho} APE(\rho, K).$$

Ο όρος $\arg \max_{\rho}$ (argument of the maximum) υποδηλώνει ότι για κάθε συγκεκριμένη τιμή του K αναζητούμε εκείνη την τιμή του ρ η οποία μεγιστοποιεί το ΠΣ (APE). Συνεπώς, η καμπύλη $\rho^*(K)$, που όπως προκύπτει και από το όρισμα της εξαρτάται άμεσα από το όριο

K , αναπαριστά την διαδρομή των μεγίστων αποκλίσεων μέσα στον παραμετρικό χώρο, συνδέοντας τα σημεία όπου εμφανίζεται η μεγαλύτερη διαφορά μεταξύ των δύο μοντέλων.

Η αριθμητική ανάλυση των δεδομένων, όπως αποτυπώνεται στους Πίνακες 3.2 και 3.3, οι οποίοι ενσωματώνουν τα πειραματικά στοιχεία του συγγράμματος [[13]], αναδεικνύει τη σημασία της έλλειψης πληροφόρησης. Συγκεκριμένα, στις σημαντικές περιπτώσεις του Πίνακα 3.2, η μέση τιμή του ΑΔ ανέρχεται σε 0,044, ενώ το ΠΣ φτάνει το 38,45%. Αντιστοίχως, στις μέτριες περιπτώσεις του Πίνακα 3.3, ο μέσος ΑΔ είναι 0,020 με το ΠΣ να εκτοξεύεται στο 56,82%.

Κ	ρ	P_b^a	P_b^p	ΑΔ (TAD)	ΠΣ% (APE%)
2	0.7	0.258	0.224	0.034	15.36
2	0.8	0.3	0.262	0.038	14.43
2	0.9	0.339	0.299	0.04	13.46
2	1.	0.372	0.33	0.042	12.59
3	0.7	0.174	0.135	0.038	28.33
3	0.8	0.219	0.173	0.045	26.14
3	0.9	0.262	0.212	0.05	23.81
3	1.	0.3	0.246	0.054	21.74
4	0.7	0.121	0.087	0.034	39.84
4	0.8	0.166	0.122	0.044	36.22
4	0.9	0.212	0.16	0.052	32.23
4	1.	0.252	0.196	0.056	28.72
5	0.8	0.129	0.089	0.04	45.29
5	0.9	0.176	0.126	0.05	39.42
5	1.	0.218	0.163	0.056	34.27
6	0.8	0.102	0.066	0.036	53.71
6	0.9	0.149	0.102	0.047	45.78
6	1.	0.192	0.139	0.054	38.83
7	0.8	0.082	0.05	0.031	61.71
7	0.9	0.127	0.084	0.043	51.57
7	1.	0.172	0.121	0.051	42.67
8	0.9	0.11	0.07	0.04	56.94
8	1.	0.156	0.107	0.049	45.97
9	0.9	0.096	0.059	0.037	62.01
9	1.	0.142	0.096	0.047	48.85
10	0.9	0.085	0.051	0.034	66.86
10	1.	0.131	0.086	0.044	51.4
Μέσος Όρος		0.187	0.143	0.044	38.45

Πίνακας 3.2: Αποκλίσεις πιθανότητας αποχώρησης για τις Σημαντικές Περιπτώσεις (Significant Cases)

Κ	ρ	P_b^a	P_b^p	ΑΔ (TAD)	ΠΣ% (APE%)
2	0.3	0.075	0.065	0.01	15.59
2	0.4	0.12	0.103	0.017	16.61
2	0.5	0.167	0.143	0.024	16.67
2	0.6	0.213	0.184	0.03	16.16
3	0.4	0.051	0.039	0.012	29.68
3	0.5	0.087	0.067	0.02	30.72
3	0.6	0.129	0.099	0.03	30.03
4	0.5	0.046	0.032	0.014	42.62
4	0.6	0.08	0.056	0.024	42.33
5	0.6	0.05	0.033	0.017	53.43
5	0.7	0.086	0.057	0.029	50.43
6	0.6	0.031	0.019	0.012	63.51
6	0.7	0.062	0.038	0.023	60.37
7	0.7	0.045	0.026	0.018	69.8
8	0.7	0.032	0.018	0.014	78.78
8	0.8	0.066	0.039	0.027	69.42
9	0.7	0.023	0.012	0.011	87.36
9	0.8	0.053	0.03	0.023	76.91
10	0.8	0.043	0.023	0.02	84.25
20	0.9	0.029	0.014	0.015	110.63
20	1.	0.072	0.043	0.029	67.61
30	1.	0.049	0.028	0.021	76.96
40	1.	0.036	0.02	0.017	83.96
50	1.	0.029	0.015	0.014	89.87
Μέσος Όρος		0.07	0.05	0.02	56.82

Πίνακας 3.3: Αποκλίσεις πιθανότητας αποχώρησης για τις Μέτριες Περιπτώσεις (Moderate Cases)

Το γεγονός ότι το ποσοστιαίο σφάλμα παραμένει υψηλό ακόμη και όταν η απόλυτη διαφορά είναι μικρή, οφείλεται στο ότι οι πιθανότητες αποχώρησης στην φυσική ουρά είναι ήδη πολύ χαμηλές (λειτουργώντας ως μικρός παρονομαστής στο κλάσμα του ΠΣ). Συνεπώς, ακόμα και μία μικρή αριθμητική αύξηση στις αποχωρήσεις, αλλάζει σημαντικά την ποσοστιαία εικόνα του συστήματος. Αναδεικνύεται, λοιπόν, η σημασία της ακριβούς μελέτης της πληροφορίας, καθώς μικρές αποκλίσεις στην πιθανότητα αποχώρησης μπορούν να έχουν μεγάλο αντίκτυπο στην τελική αξιολόγηση της ουράς.

Συνοψίζοντας τα παραπάνω ποιοτικά και ποσοτικά ευρήματα, προκύπτουν αξιόλογα συμπεράσματα που αφορούν τον τρόπο λειτουργίας τέτοιων συστημάτων στην πράξη. Το

πρώτο σημείο εστιάζει στην ψευδαίσθηση της συμφόρησης που δημιουργείται στον παραμετρικό χώρο της ουράς εισιτηρίων. Όπως επισημάνθηκε και κατά την αρχική ανάλυση της Πρότασης 3.1, παρόλο που το σύστημα αυτό είναι στην ουσία λιγότερο φορτωμένο από τη φυσική ουρά, καθώς ο πραγματικός αριθμός των παρόντων πελατών είναι μικρότερος, στον πελάτη δίνει την εντύπωση ενός πολύ πιο πολυσχολου συστήματος. Αυτό συμβαίνει επειδή ο αριθμός που αναγράφεται στο εισιτήριο είναι άμεσα ορατός, ενώ ο αριθμός των ατόμων που περιμένουν ενεργά παραμένει κρυφός. Αυτή η απόκλιση προκαλεί μια συμπεριφορά των πελατών που αποδεικνύεται επιζήμια και για τους ίδιους αλλά και για την επιχείρηση. Συγκεκριμένα, η επιχείρηση έρχεται αντιμέτωπη με απώλεια πωλήσεων ακόμα και όταν υπάρχει διαθέσιμη χωρητικότητα για εξυπηρέτηση, ενώ οι πελάτες μένουν με την ανικανοποίητη ζήτηση παρόλο που θα μπορούσαν να είχαν εξυπηρετηθεί αν γνώριζαν την πραγματική κατάσταση. Προκειμένου, λοιπόν, να κατευθυνθεί σωστά η στάση του κοινού και να αναβαθμιστεί το επίπεδο εξυπηρέτησης, είναι απαραίτητο να βρεθούν τρόποι ώστε οι πελάτες να λαμβάνουν τα ποσοτικοποιημένα εκείνα στοιχεία που θα τους οδηγήσουν σε μια πιο ολοκληρωμένη και πιο αξιόπιστη εικόνα της αναμονής που έχουν να αντιμετωπίσουν.

Αν μας απασχολήσει τώρα το πώς οι βασικές παράμετροι του συστήματος επηρεάζουν το μέγεθος της απόκλισης, το Σχήμα 3.10 προσφέρει μια ξεκάθαρη εικόνα αυτής της ευαισθησίας. Συγκεκριμένα, παρατηρώντας το γράφημα (α), αν επιλέξουμε ένα σταθερό σημείο στον κατακόρυφο άξονα που αντιπροσωπεύει την ένταση της κίνησης ρ και κινηθούμε οριζόντια προς τα δεξιά αυξάνοντας την υπομονή του πελάτη K , βλέπουμε την απόκλιση TAD να μειώνεται, οδηγώντας μας από τη Σημαντική στην Ασήμαντη Περιοχή. Αυτό συμβαίνει διότι ένας πιο υπομονετικός πελάτης (που σημαίνει μεγαλύτερο K) επηρεάζεται λιγότερο από το ορατό νούμερο του εισιτηρίου, με αποτέλεσμα η συμπεριφορά του να προσεγγίζει εκείνη της φυσικής ουράς. Αντιθέτως, αν θεωρήσουμε την υπομονή K σταθερή στον οριζόντιο άξονα και κινηθούμε κατακόρυφα προς τα πάνω αυξάνοντας την κίνηση ρ , η απόκλιση μεγαλώνει αισθητά. Η λογική πίσω από αυτό είναι ότι όσο ο χώρος αναμονής γεμίζει, η έλλειψη πληροφόρησης για τον πραγματικό αριθμό των παρόντων γίνεται πιο καθοριστική, ωθώντας περισσότερους πελάτες σε αποχώρηση λόγω της εικόνας του υψηλού αριθμού στο εισιτήριο. Ανατρέχοντας, μάλιστα, στη στήλη $\Delta\Delta$ (TAD) του Πίνακα 3.2, γίνεται φανερό ότι η διαφορά στις πιθανότητες αποχώρησης, όπως προκύπτει από την επεξεργασία των ακραίων περιπτώσεων, παρουσιάζει αυξητική τάση υπό συνθήκες έντασης συμφόρησης, φτάνοντας ανώτατα να προσεγγίζει το 6%.

Παράλληλα με την απόλυτη διαφορά, ιδιαίτερη βαρύτητα έχει η ανάλυση του ποσοστιαίου σφάλματος ΠΣ (APE), το οποίο στις Μέτριες Περιπτώσεις του Πίνακα 3.3 ανέρχεται κατά μέσο όρο στο 56,82%, όπως έχουμε ήδη διαπιστώσει. Το εντυπωσιακό αυτό ποσοστό φανερώνει τον κίνδυνο υποεκτίμησης των αποχωρήσεων, καθώς δείχνει ότι η πραγματική πιθανότητα ένας πελάτης να φύγει από την ουρά εισιτηρίων είναι σταθερά μεγαλύτερη από αυτή της φυσικής ουράς. Η μελέτη των αριθμητικών δεδομένων αποκαλύπτει ότι το σύστημα παρουσιάζει μια ικανοποιητική προσέγγιση, κυρίως σε συνθήκες χαμηλής ανοχής, $K = 2$, όπου το ΠΣ διατηρείται σε χαμηλά επίπεδα κοντά στο 15,6%. Αντιθέτως, καθώς οι πελάτες γίνονται πιο υπομονετικοί, το ΠΣ τείνει να διογκώνεται, φτάνοντας ακόμα και το 110, 63%

για $K = 20$, παρόλο που η ΑΔ παραμένει σε χαμηλά επίπεδα. Αυτό συμβαίνει επειδή, όπως αναφέρθηκε και προγενέστερα, οι συνολικές αποχωρήσεις είναι ελάχιστες με αποτέλεσμα ακόμη και μία ανεπαίσθητη απόσταση στις πιθανότητες αποχώρησης να εμφανίζεται ως τεράστιο ποσοστό σφάλματος. Υπό αυτές τις συνθήκες, το μοντέλο των εισιτηρίων εξασφαλίζει υψηλές συνθήκες αναμονής για το κοινό, επιτυγχάνοντας αποδοτικότητα ανάλογη με εκείνη της φυσικής ουράς, μόνο στον βαθμό που η απόκλιση ΑΔ περιορίζεται σε αμελητέα επίπεδα. Πρακτικά, αυτό σημαίνει ότι η διαφορά στη συμπεριφορά δεν εξαλείφεται αυτόματα με την αύξηση της υπομονής, αλλά απαιτεί συγχρόνως χαμηλότερο φόρτο κίνησης, για να θεωρηθεί απόλυτα μηδαμινή.

Συμπληρωματικά, η μελέτη των δεδομένων επιβεβαιώνει ότι η απόδοση του συστήματος υποβαθμίζεται με επιταχυνόμενο ρυθμό όσο οι πελάτες γίνονται πιο ανυπόμονοι και ο χώρος εξυπηρέτησης γεμίζει. Συγκεκριμένα, η πιθανότητα αποχώρησης P_b^p στη φυσική ουρά αυξάνεται όσο εντείνεται η κίνηση ρ και μειώνεται όσο αυξάνεται η υπομονή K , επιβεβαιώνοντας την κοινή λογική ότι η έλλειψη ανοχής και η μεγάλη συμφόρηση οδηγούν σε περισσότερες αποχωρήσεις. Αυτό το μοτίβο δεν αφορά μόνο τη φυσική ουρά, αφού τα υπολογιστικά αποτελέσματα της έρευνας επιβεβαιώνουν ότι την ίδια ακριβώς τάση ακολουθεί τόσο η πιθανότητα αποχώρησης στα εισιτήρια P_b^a , όσο και η απόλυτη διαφορά ΑΔ μεταξύ των δύο συστημάτων. Στην πραγματικότητα, η χρήση εισιτηρίων λειτουργεί ως επιβαρυντικός παράγοντας που κάνει τη συνολική εικόνα της εξυπηρέτησης ακόμα χειρότερη από αυτή μιας παραδοσιακής ουράς.

Η ουσία της διαπίστωσης αυτής είναι ότι η κατάσταση στα εισιτήρια χειροτερεύει με επιταχυνόμενο ρυθμό. Αυτό σημαίνει ότι κάθε επιπλέον πελάτης που προσέρχεται ή κάθε βαθμός υπομονής που χάνεται προκαλεί όλο και μεγαλύτερη ζημιά στο σύστημα των εισιτηρίων σε σύγκριση με τη φυσική ουρά. Με άλλα λόγια, η ουρά εισιτηρίων τείνει να επιτείνει μια ήδη κακή κατάσταση, καθώς αν η φυσική ουρά δυσκολεύεται να συγκρατήσει τους πελάτες, τα εισιτήρια την οδηγούν σε ταχύτερη κατάρρευση.

Η παρατήρηση αυτή γίνεται αντιληπτή αν συγκρίνουμε τους μέσους όρους από τους πίνακες που παρουσιάστηκαν παραπάνω. Όπως φαίνεται στον Πίνακα 3.3, για τις Μέτριες Περιπτώσεις, ενώ σε μία φυσική ουρά αναμένουμε κατά μέσο όρο ένα 5% αποχωρήσεων, το ποσοστό αυτό ανεβαίνει στο 7%, όταν οι πελάτες χρησιμοποιούν εισιτήρια. Η επιδείνωση αυτή γίνεται ακόμα πιο δραματική στον Πίνακα 3.2, για τη Σημαντική Περιοχή, όπου η ήδη υψηλή πιθανότητα αποχώρησης 14,3% της φυσικής ουράς εκτοξεύεται στο 18,7% στα εισιτήρια, με την επιχείρηση να χάνει επιπλέον 4,4% της πελατείας της. Το εύρημα αυτό καθιστά σαφές ότι επειδή τα εισιτήρια αυξάνουν τις αποχωρήσεις όσο οι συνθήκες δυσκολεύουν, ο διαχειριστής οφείλει να επέμβει με στρατηγικές όπως η παροχή πληροφοριών για τον χρόνο αναμονής, ώστε να περιορίσει τις απώλειες και να βελτιώσει την παρεχόμενη υπηρεσία του.

Η ευρεία εφαρμογή της τεχνολογίας των εισιτηρίων σε συστήματα με υψηλό φόρτο κίνησης δεν είναι τυχαία, ωστόσο η αποτελεσματικότητά τους εξαρτάται άμεσα από το είδος της

υπηρεσίας και το προφίλ των πελατών. Αν θεωρήσουμε την ευαισθησία στον χρόνο ως μέτρο υπομονής του κοινού, γίνεται αντιληπτό ότι η παράμετρος K δεν αποτελεί έναν στατικό αριθμό, αλλά συνδέεται άμεσα με την υποκειμενική χρησιμότητα της παρεχόμενης υπηρεσίας. Για παράδειγμα, οι πελάτες σε τραπεζικά ιδρύματα τείνουν να είναι λιγότερο υπομονετικοί, καθώς η ευαισθησία τους στον χρόνο είναι υψηλή, με αποτέλεσμα το όριο αποχώρησης K να εξαντλείται ταχύτερα. Σε αντίθεση, σε υπηρεσίες μεταφορών ή δημόσιους οργανισμούς, η κρισιμότητα της διεκπεραίωσης αναγκάζει την ευαισθησία στον χρόνο να υποχωρήσει μπροστά στην ανάγκη, αυξάνοντας έτσι την ανοχή του πελάτη. Σε τέτοιες περιπτώσεις η διοίκηση κινδυνεύει να υποεκτιμήσει σημαντικά την πτώση του επιπέδου εξυπηρέτησης, εάν χρησιμοποιεί την απόδοση της φυσικής ουράς ως σημείο αναφοράς για το σύστημα των εισιτηρίων. Καταλήγοντας, επειδή στις περιπτώσεις όπου οι πιθανότητες αποχώρησης είναι εξαιρετικά χαμηλές και για τα δύο συστήματα, ο δείκτης ΠΣ τείνει να διογκώνει την ποσοστιαία διαφορά μεταξύ αυτών των μικρών τιμών, η ΑΔ κρίνεται ως το πλέον αξιόπιστο εργαλείο για την πρωτογενή αξιολόγηση. Καθώς αποτυπώνει την ουσιαστική αριθμητική κλίμακα της απόκλισης, η ΑΔ προσφέρει μία πιο σταθερή εικόνα, ενώ το ΠΣ λειτουργεί συμπληρωματικά για την ολοκληρωμένη ποσοτικοποίηση της στοχαστικής διαφοράς μεταξύ των δύο συστημάτων.

3.7.2 Η Ticket-Plus Queue και η Συγκριτική Ανάλυσή της με την M/M/1/K

Η ανάλυση που προηγήθηκε φανέρωσε σαφώς ότι η έλλειψη πληροφόρησης στα συστήματα εισιτηρίων λειτουργεί αποτρεπτικά για το κοινό, οδηγώντας σε αυξημένες αποχωρήσεις και κατ' επέκταση, σε σημαντική απώλεια δυνητικών εξυπηρετήσεων. Ωστόσο, η επιστημονική μελέτη δεν εξαντλείται στη διαπίστωση του κρίσιμου αυτού προβλήματος, αλλά καθιστά επιτακτική την ανάγκη διατύπωσης μίας συγκεκριμένης και εφαρμόσιμης λύσης.

Στην παρούσα υποενότητα, εξετάζεται μια πρόταση βελτίωσης του συστήματος, η οποία βασίζεται στην παροχή επιπλέον πληροφορίας προς τον πελάτη. Η κεντρική ιδέα έγκειται στην εκτύπωση του αναμενόμενου χρόνου αναμονής πάνω στο εισιτήριο, με στόχο την διόρθωση της μεροληπτικής εκτίμησης του πελάτη για την καθυστέρηση. Στο πλαίσιο αυτό, αναδεικνύεται ο καθοριστικός ρόλος της τεχνολογίας, η οποία δύναται να λειτουργήσει ως μέσο ενίσχυσης της συνολικής απόδοσης, μετατρέποντας το σύστημα αναμονής από σημείο συμφόρησης σε περιβάλλον ομαλής λειτουργίας.

Η εξελικτική αυτή προσέγγιση, που ονομάζεται Ticket-Plus Queue, αποσκοπεί –όπως θα δειχθεί στην πορεία της παρούσας υποενότητας– στο να επιτρέψει στη διοίκηση του συστήματος να εξισώσει την απόδοση των εισιτηρίων με αυτή της φυσικής ουράς M/M/1/K. Μέσω της στρατηγικής αυτής, επιδιώκεται ο επαναπροσδιορισμός της στοχαστικής συμπεριφοράς του πελάτη, ώστε ο πάροχος της υπηρεσίας να αναβαθμίσει την απόδοση στα επίπεδα μίας φυσικής ουράς, διασφαλίζοντας παράλληλα τα λειτουργικά οφέλη της οργανωμένης διαχείρισης αναμονής.

Η ανάγκη για την αναβάθμιση του συστήματος πηγάζει από το γεγονός ότι ο πελάτης, λαμβάνοντας ένα εισιτήριο με μία δεδομένη τιμή d (όπου το d αποτελεί μια συγκεκριμένη πραγμάτωση της μεταβλητής D), προβαίνει σε μια υποκειμενική εκτίμηση της αναμονής του. Η εκτίμηση αυτή ορίζεται από τον λόγο d/μ . Συγκεκριμένα, ο πελάτης υποθέτει λανθασμένα ότι οφείλει να περιμένει την εξυπηρέτηση και των $d-1$ προπορευόμενων ατόμων που έλαβαν αριθμό προτεραιότητας. Στην πραγματικότητα, η εκτίμηση αυτή είναι από τη φύση της υπερβολική, καθώς δεν συνυπολογίζει τις πιθανές αποχωρήσεις των προηγούμενων πελατών, οι οποίοι ενδέχεται να εγκαταλείψουν την ουρά λόγω της δικής τους αποθάρρυνσης. Συνεπώς, η πραγματική αναμονή προκύπτει από τον αριθμό των πραγματικά παρόντων πελατών και εκφράζεται από τον αναμενόμενο χρόνο $E(N|D=d)/\mu$, ο οποίος είναι βέβαια μικρότερος από το d/μ . Το σύστημα Ticket-Plus έρχεται να διορθώσει αυτή την πλάνη, ανακοινώνοντας τον ρεαλιστικό μειωμένο χρόνο αναμονής στους εισερχόμενους.

Η παροχή της ακριβούς πληροφορίας δεν αποσκοπεί μόνο στη μείωση των άσκοπων αποχωρήσεων, μέσω του ακριβέστερου υπολογισμού της αναμονής, αλλά δίνει μία επιπλέον ψυχολογική διάσταση στο ζήτημα. Η στρατηγική της πληροφόρησης υποστηρίζεται από βιβλιογραφικά δεδομένα, που της προσδίδουν ένα βαθύ ψυχολογικό υπόβαθρο. Συγκεκριμένα, η θεμελιώδης έρευνα του Larson ([7]), σελ. 900) υπογραμμίζει ότι οι πελάτες τείνουν να αισθάνονται καλύτερα όταν διαθέτουν δεδομένα που τους επιτρέπουν να εκτιμήσουν με ακρίβεια την αναμονή τους. Παράλληλα, στο πείραμά τους, οι Hui και Tse ([4]) εξέτασαν πώς διαφορετικά είδη πληροφοριών επηρεάζουν την ψυχολογία του προσερχόμενου κατά την αναμονή. Τα ευρήματά τους επιβεβαιώνουν ότι η γνώση του προβλεπόμενου χρόνου αναμονής συμβάλλει στη διαμόρφωση ενός καλύτερου κλίματος και ενισχύει τη θετική διάθεση του κοινού, αφού ο πελάτης αισθάνεται μεγαλύτερη βεβαιότητα και ικανοποίηση. Έτσι, αναβαθμίζεται η τελική εικόνα του παρόχου, έχοντας σαν συνέπεια να βελτιώνεται η συνολική αξιολόγηση της υπηρεσίας.

Η πρακτική εφαρμογή της Ticket-Plus Queue βασίζεται στην αξιοποίηση των υπολογιστικών δεδομένων που αναλύθηκαν στην ενότητα 3.6.2. Όπως παρατηρήθηκε, ο διαχειριστής του συστήματος είναι σε θέση να υπολογίσει την προσδοκώμενη πραγματική θέση του πελάτη $E(N|D=d)$, μέσω της σχέσης (3.42), εξαλείφοντας τη λανθασμένη εντύπωση που προκαλεί η απλή ανάγνωση του αριθμού προτεραιότητας d . Για την υλοποίηση αυτής της προσέγγισης, υποτίθεται ότι ο πάροχος γνωρίζει το όριο υπομονής K του κοινού. Ενώ στο απλό σύστημα εισιτηρίων ο πελάτης θα αποχωρούσε μόλις η θέση του d έφτανε ή ξεπερνούσε το K , εδώ η απόφαση αυτή προσδιορίζεται εκ νέου. Ο πελάτης πλέον ενημερώνεται για τον αντικειμενικό αναμενόμενο χρόνο αναμονής $E(N|D=d)/\mu$ και επιλέγει να παραμείνει μόνο εάν αυτή η τιμή δεν υπερβαίνει το προσωπικό του όριο. Στο πλαίσιο αυτό, αναζητείται το νέο κρίσιμο όριο εγκατάλειψης, το οποίο συμβολίζεται με $\tilde{K}$. Το όριο αυτό ορίζεται ως ο ελάχιστος ακέραιος για τον οποίο η αναμενόμενη πραγματική θέση ισούται ή υπερβαίνει το όριο K :

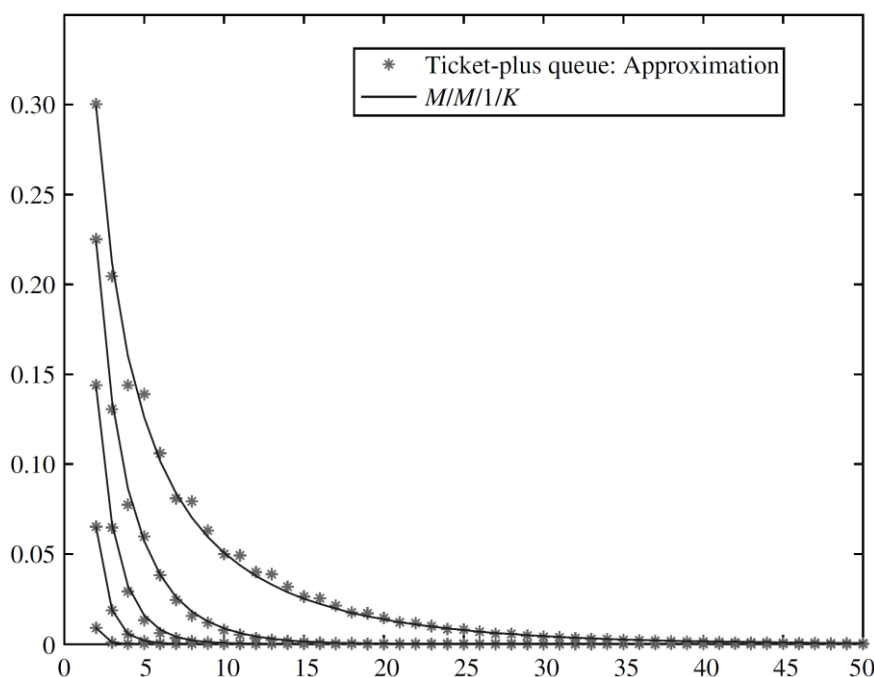
$$E(N|D=\tilde{K}) \geq K. \quad (3.45)$$

Λόγω της σχεδόν γραμμικής σχέσης μεταξύ της πραγματικής θέσης και του αριθμού του εισιτηρίου –ιδιότητα που διαπιστώθηκε στην ενότητα 3.6.2– το νέο όριο $\tilde{K}$ είναι μοναδικό, σαφώς ορισμένο και προσδιορίσιμο. Η εύρεσή του πραγματοποιείται μέσω μίας διαδικασίας ανοδικής αναζήτησης, ξεκινώντας από την τιμή K . Είναι αξιοσημείωτο ότι, επειδή η πραγματική μέση θέση $E[N]$ είναι πάντοτε μικρότερη ή ίση της ονομαστικής θέσης d (λόγω των ενδιάμεσων αποχωρήσεων), το νέο όριο αποθάρρυνσης θα είναι πάντοτε μεγαλύτερο από το αρχικό, δηλαδή $\tilde{K} > K$. Κατά συνέπεια, η παροχή της πληροφορίας επιτρέπει στην ουρά να λειτουργεί με διευρυμένη χωρητικότητα $\tilde{K}$, αυξάνοντας τον αριθμό των πελατών που περιμένουν προς εξυπηρέτηση.

Ως εκ τούτου, προτείνεται η αναγραφή του μέσου χρόνου αναμονής απευθείας πάνω στο εισιτήριο προτεραιότητας. Η πρακτική αυτή λειτουργεί ως ένας μηχανισμός παροχής μίας έγκυρης στατιστικής εκτίμησης για την κατάσταση της ουράς, καθοδηγώντας τους πελάτες να προσαρμόσουν τη συμπεριφορά τους με βάση ρεαλιστικά στοιχεία, απαλλαγμένοι από υποκειμενικές εντυπώσεις.

Η τελική αξιολόγηση της Ticket-Plus Queue πραγματοποιείται μέσα από τη σύγκρισή της με την $M/M/1/K$. Η διαδικασία αυτή λειτουργεί ως μία διπλή επιβεβαίωση –οπτική και αριθμητική– που ισχυροποιεί την αξιοπιστία της προσεγγιστικής μεθόδου.

Αρχικά, η οπτική τεκμηρίωση παρέχεται από το Σχήμα 3.11, όπου εξετάζεται η πιθανότητα αποχώρησης σε συνάρτηση με το όριο K , το οποίο διατρέχει τον οριζόντιο άξονα. Οι διαφορετικές καμπύλες του σχήματος αντιστοιχούν σε συγκεκριμένες τιμές της έντασης κίνησης $\rho = 0,1, \dots, 0,9$, επιτρέποντας την αξιολόγηση του συστήματος υπό ποικίλες συνθήκες φόρτου.



Σχήμα 3.11: Σύγκριση των P_b μεταξύ της Ticket-Plus Queue και της $M/M/1/K$

Όπως προκύπτει από τη γραφική παράσταση, οι καμπύλες των δύο μοντέλων συμπίπτουν σχεδόν πλήρως. Η οπτική αυτή ταύτιση υποδηλώνει ότι η μετάδοση της πληροφορίας του αναμενόμενου χρόνου επιτρέπει στο σύστημα εισιτηρίων να αναπαράγει με ακρίβεια τη συμπεριφορά μίας φυσικής ουράς αναμονής.

Πέρα από την οπτική σύγκριση, η ανάλυση συμπληρώνεται από τα αριθμητικά δεδομένα του Πίνακα 3.4, στον οποίο παρουσιάζονται οι 10 περιπτώσεις με τη μεγαλύτερη απόκλιση, ανάμεσα στα 130 σενάρια που εξετάστηκαν στο [[13]]. Ο πίνακας συγκρίνει τη δεσμευμένη αναμενόμενη τιμή του χρόνου αναμονής για τους πελάτες που αποφάσισαν να εισέλθουν στο σύστημα. Λεπτομερέστερα, εξετάστηκε η σύγκλιση μεταξύ του χρόνου αναμονής W^+ στην ουρά Ticket-Plus, υπό τη δέσμευση ότι η θέση του εισιτηρίου D^+ είναι μικρότερη του ορίου $\tilde{K}$, και του αντίστοιχου χρόνου W^P στην ουρά M/M/1/K, υπό τη δέσμευση ότι η θέση του D^P είναι μικρότερη του ορίου K .

K	ρ	$E[W^+ D^+ < \tilde{K}]$	$E[W^P D^P < K]$	ΠΣ% (APE%)
2	1.	0.96	0.99	2.99
3	1.	1.43	1.49	3.79
4	1.	1.9	1.98	4.
5	1.	2.37	2.47	4.
6	1.	2.84	2.96	3.92
7	1.	3.32	3.45	3.8
8	1.	3.79	3.93	3.68
9	1.	4.26	4.42	3.55
10	1.	4.73	4.9	3.42
20	0.9	6.6	6.42	2.81
Μέγιστη τιμή των 130 δοκιμών				4.
Μέσος όρος των 130 δοκιμών				0.78

Πίνακας 3.4: Σύγκριση αναμενόμενων χρόνων αναμονής – 10 δυσμενέστερες περιπτώσεις

Από τα αριθμητικά δεδομένα προκύπτει ότι οι αναμενόμενοι χρόνοι αναμονής παραμένουν σχεδόν ταυτόσημοι, με το ποσοστιαίο σφάλμα να κυμαίνεται σε πολύ χαμηλά επίπεδα και τη μέση τιμή του να μην ξεπερνά το 0.78%. Μάλιστα, ακόμα και στις δυσμενέστερες συνθήκες έντονης κίνησης ($\rho = 1,0$) το σφάλμα δεν υπερβαίνει το 4%. Η σύγκριση αυτή επιβεβαιώνει αριθμητικά ότι $E(N^+ | D^+ < \tilde{K}) \approx E(N^P | D^P < K)$, γεγονός που αποδεικνύει ότι η αναβαθμισμένη προσέγγιση ανταποκρίνεται με υψηλή ακρίβεια στην πραγματικότητα.

Παρά την εξαιρετική σύγκλιση των αναμενόμενων χρόνων αναμονής, παρατηρείται μία ποιοτική διαφορά στη φύση των δύο συστημάτων. Στο σύστημα M/M/1/K, η θέση του πελάτη είναι προκαθορισμένη από τη δομή της ουράς. Αντίθετα, στο σύστημα Ticket-Plus, η

ενημέρωση του πελάτη για τον αναμενόμενο χρόνο αναμονής εισάγει μία δυναμική διαδικασία στη λήψη απόφασης, η οποία εξαρτάται από τις αποφάσεις εγκατάλειψης (balking) των λοιπών πελατών που προηγούνται.

Συμπερασματικά, η συνδυαστική μελέτη του σχήματος και του πίνακα φανερώνει ότι η χορήγηση ακριβούς πληροφόρησης επιτρέπει στη στρατηγική Ticket-Plus να επιτυγχάνει απόδοση σχεδόν ταυτόσημη με εκείνη της ουράς $M/M/1/K$. Με τον τρόπο αυτό, η διοίκηση του συστήματος μπορεί να διατηρήσει τα πρακτικά πλεονεκτήματα των εισιτηρίων, καθοδηγώντας τη συμπεριφορά των πελατών και περιορίζοντας τις άσκοπες αποχωρήσεις. Η προσέγγιση αυτή μετατρέπει την ουρά από έναν μηχανισμό αβέβαιης αναμονής σε ένα διαφανές σύστημα εξυπηρέτησης, συμβάλλοντας στην ικανοποίηση του κοινού και τη μέγιστη αξιοποίηση των πόρων του παρόχου.

Κεφάλαιο 4^ο: Ανάλυση Συμπεριφοράς Πελατών και Στρατηγικές Κατωφλιού σε Συστήματα Εξυπηρέτησης με Εισιτήρια

Στο κεφάλαιο αυτό, μετακινούμαστε από την κλασική μελέτη των συστημάτων ουράς προς μία προσέγγιση που εστιάζει στις επιλογές των πελατών. Αντί να εξετάσουμε απλώς τους δείκτες απόδοσης του συστήματος, στρέφουμε το ενδιαφέρον μας στο πώς ενεργούν οι πελάτες όταν φτάνουν σε μία ουρά και η εικόνα που έχουν για την κατάσταση του συστήματος είναι περιορισμένη. Η ανάλυση βασίζεται στη μελέτη των Yoav Kerner, Eliran Sherzer και Mor Ann Yanco με τίτλο *On non-equilibria threshold strategies in ticket queues*. Με τη χρήση Μαρκοβιανών Διεργασιών εντοπίζονται τα σημεία όπου οι απλές στρατηγικές κατωφλιού αποτυγχάνουν. Με τον τρόπο αυτό επιδιώκεται η κατανόηση της καλύτερης δυνατής αντίδρασης από την πλευρά του πελάτη και αναζητούνται μέθοδοι για να επιτευχθεί μια στρατηγική ισορροπίας σε συνθήκες αβεβαιότητας.

4.1 Εισαγωγή στην Στρατηγική Συμπεριφορά των Πελατών

Όπως αναλύθηκε στο Κεφάλαιο 3, το σύστημα αναμονής με εισιτήρια βασίζεται σε έναν μηχανισμό ψηφιακής αναγγελίας όπου οι πελάτες λαμβάνουν αποφάσεις υπό συνθήκες αβεβαιότητας σχετικά με το πόσοι από τους προπορευόμενους αριθμούς εισιτηρίων παραμένουν ενεργοί. Οι αποφάσεις των πελατών, για την παραμονή τους στο σύστημα, λαμβάνονται βάσει ενός δεδομένου σταθερού ορίου ανοχής K .

Ωστόσο, στο παρόν κεφάλαιο εισάγεται η διάσταση της στρατηγικής αλληλεπίδρασης, εξετάζοντας κατά πόσο το όριο αυτό αποτελεί όντως τη βέλτιστη επιλογή για τον πελάτη στο δεδομένο περιβάλλον.

Στο πλαίσιο αυτό, η έννοια της ισορροπίας Nash αποτελεί το κεντρικό εργαλείο ανάλυσης. Αυτή περιγράφει μια κατάσταση σταθερότητας, όπου κανένας πελάτης δεν έχει κίνητρο να αλλάξει τη στρατηγική του μεμονωμένα, δεδομένων των στρατηγικών που ακολουθούν οι υπόλοιποι. Δηλαδή, οι αποφάσεις όλων των πελατών είναι τέτοιες, ώστε κανένας δεν έχει λόγο να εγκαταλείψει την τακτική του, εφόσον οι υπόλοιποι επιμένουν στη δική τους. Πρόκειται για μία κατάσταση όπου κάθε πελάτης έχει επιλέξει την ευνοϊκότερη δυνατή κίνηση, συνυπολογίζοντας τις αντίστοιχες στρατηγικές των υπολοίπων.

Αναδεικνύεται, στην πορεία του παρόντος κεφαλαίου, ότι σε συστήματα που απουσιάζει η πλήρης πληροφόρηση, οι παραδοσιακές στρατηγικές που βασίζονται σε ένα μεμονωμένο όριο ανοχής, που στο εξής καλείται «κατώφλι», δεν καταφέρνουν να επιτύχουν αυτή την ισορροπία.

Βέβαια, εάν το σύνολο των πελατών υιοθετήσει μια στρατηγική που επιτρέπει την παραμονή στο σύστημα μόνο όταν η παρατηρούμενη διαφορά αριθμών στο εισιτήριο είναι χαμηλή, ο κάθε πελάτης ξεχωριστά διαπιστώνει ότι τον συμφέρει να διαφοροποιηθεί και να εισέλθει ακόμη και όταν η ουρά

εμφανίζεται μεγαλύτερη από το όριο που είχε τεθεί ως αρχικό κριτήριο παραμονής. Η κίνηση αυτή δικαιολογείται από το γεγονός ότι οι υψηλές τιμές στον πίνακα προτεραιότητας ενδέχεται να υποδηλώνουν μαζικές αποχωρήσεις προηγούμενων πελατών, καθιστώντας την ουρά στην πράξη λιγότερο επιβαρυνμένη. Ως εκ τούτου, η βέλτιστη απόκριση του πελάτη απέναντι στις επιλογές των άλλων οδηγεί τελικά στην υιοθέτηση ενός πιο σύνθετου σχήματος, γνωστού ως στρατηγική διπλού κατωφλιού.

Για να φανεί στην πράξη πώς προκύπτει αυτό, στη συνέχεια θα παρουσιαστεί το μαθηματικό μοντέλο του συστήματος. Εκεί θα αναλυθεί ο τρόπος με τον οποίο κάθε πελάτης εκτιμά την κατάσταση του συστήματος για να λάβει την απόφασή του. Οι πελάτες θα διαχωριστούν σε «πραγματικούς» και «εικονικούς», ανάλογα με το αν θα παραμείνουν ή θα εγκαταλείψουν το σύστημα.

4.2 Μοντέλο και Αναπαράσταση του Χώρου Καταστάσεων

Για τις ανάγκες της μελέτης της παρούσας ενότητας, παρουσιάζεται εκ νέου το μαθηματικό πλαίσιο του συστήματος, υιοθετώντας έναν ελαφρώς αναδιαμορφωμένο συμβολισμό, ώστε να ικανοποιούνται οι στόχοι του κεφαλαίου. Αν και εννοιολογικά ο χώρος ταυτίζεται με το μοντέλο που εξετάστηκε στο προηγούμενο κεφάλαιο, η παρούσα αναπαράσταση είναι απαραίτητη για την ακριβή περιγραφή και την εις βάθος μελέτη των Μαρκοβιανών Διεργασιών, οι οποίες αποτελούν την βάση για την περαιτέρω διερεύνηση.

Στο πλαίσιο της παρούσας μελέτης, το σύστημα ticket queue μοντελοποιείται ως μια Μαρκοβιανή Αλυσίδα Συνεχούς Χρόνου. Εξετάζεται ένα σύστημα ουράς M/M/1 με ρυθμό αφίξεων λ και ρυθμό εξυπηρέτησεων μ , που διέπεται από την προτεραιότητα FCFS (First Come, First Served). Όταν ένας πελάτης φτάνει και βρίσκει τον διακομιστή ελεύθερο, η εξυπηρέτησή του ξεκινά αμέσως. Αν ο διακομιστής είναι απασχολημένος, ο πελάτης καλείται να λάβει την απόφαση για το αν θα ενταχθεί στην ουρά ή θα αποχωρήσει. Στην περίπτωση της αποχώρησης, η απουσία του πελάτη γίνεται αντιληπτή μόνο κατά τη στιγμή που θα κληθεί για εξυπηρέτηση, οπότε και θεωρείται ότι εξυπηρέτηθηκε σε μηδενικό χρόνο. Οι επόμενοι πελάτες δεν διαθέτουν πληροφόρηση σχετικά με τις προηγούμενες αποχωρήσεις και η απόφαση ένταξης βασίζεται αποκλειστικά στο παρατηρούμενο μήκος της ουράς, δηλαδή στην απόσταση του εκδοθέντος εισιτηρίου από την ένδειξη του τελευταίου αριθμού που έχει καλέσει το σύστημα.

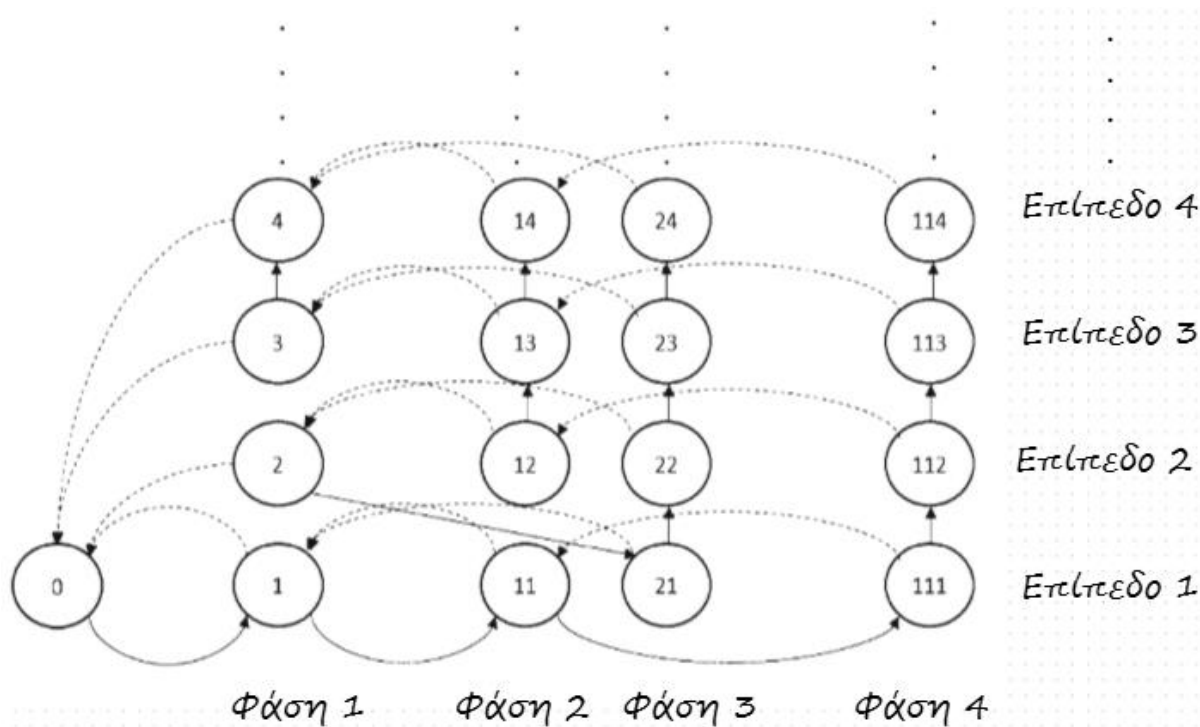
Υποθέτουμε ότι οι πελάτες ακολουθούν μία στρατηγική κατωφλιού K , εφόσον εισέρχονται στην εξυπηρέτηση μόνο όταν το συνολικό πλήθος πελατών D παραμένει μικρότερο του K ($D < K$). Θεωρούμε την κατάσταση 0 ως το κενό σύστημα, ενώ, για κάθε άλλη περίπτωση η κατάσταση έχει τη μορφή $\mathbf{n} = (n_1, \dots, n_i, \dots, n_y)$, με το n_i να αντιστοιχεί στον αριθμό των εισιτηρίων που εκδόθηκαν μεταξύ του i -οστού και του $i+1$ -οστού πελάτη που εισέρχεται. Το y , $1 \leq y \leq K$, αντιπροσωπεύει το πλήθος των παρόντων πελατών του συστήματος.

Έτσι, ο χώρος καταστάσεων του συστήματος διαμορφώνεται ως:

$$A = \{0\} \bigcup_{y=1}^K \left\{ \mathbf{n} \in \mathbf{N}^y : n_i \geq 1, i=1, \dots, y, y=1, \dots, K \text{ και } \sum_{i=1}^{y-1} n_i < K \right\}. \quad (4.1)$$

Μπορούμε να φανταστούμε το σύστημα σαν μία ουρά όπου έχουμε πραγματικούς πελάτες (αυτούς που περιμένουν να εξυπηρετηθούν) και εικονικούς πελάτες (αυτούς που εκδίδουν εισιτήριο, αλλά αποχωρούν λόγω του ορίου K). Το διάνυσμα κατάστασης $\mathbf{n} = (n_1, \dots, n_i, \dots, n_y)$ διαθέτει όλη την πληροφορία του συστήματος, όπου το y είναι ο αριθμός πραγματικών πελατών που έχουν ενεργή θέση στο σύστημα και το n_i είναι πόσοι εικονικοί πελάτες μεσολαβούν από την είσοδο του i -οστού πραγματικού πελάτη και του αμέσως επόμενου. Η λειτουργία του συστήματος απαιτεί την ταξινόμηση σε Επίπεδα (Levels) και Φάσεις (Phases). Τα επίπεδα λαμβάνουν αριθμηση σύμφωνα με το τελευταίο στοιχείο της κατάστασης, δηλαδή το κάθε επίπεδο αντιπροσωπεύει τον αριθμό των εισιτηρίων που εκδόθηκαν από τον τελευταίο πραγματικό πελάτη και μετά, ενώ οι φάσεις καθορίζονται από τη διαφορετική εσωτερική διάταξη των πελατών.

Για να γίνει απόλυτα κατανοητός ο τρόπος λειτουργίας του συστήματος παρατίθεται στο Σχήμα 4.1 το διάγραμμα μεταβάσεων για την ειδική περίπτωση που το κατώφλι αποχώρησης λαμβάνει την τιμή $K=3$. Αν και έχει ήδη παρουσιαστεί και μελετηθεί μία ανάλογη περίπτωση στο προηγούμενο κεφάλαιο, για $K=4$, κρίνεται χρήσιμο να επαναληφθεί η σχηματική απεικόνιση και εδώ. Ο λόγος είναι ότι η ανάλυση στο παρόν κεφάλαιο ακολουθεί μια διαφορετική πορεία και η οπτικοποίηση και σε αυτή την περίπτωση στοχεύει να δοθεί μια αυτονομία στην παρουσίαση του υλικού του παρόντος κεφαλαίου.



Σχήμα 4.1: Διάγραμμα μεταβάσεων για $K=3$

Όπως αποτυπώνεται στο Σχήμα 4.1, η αναπαράσταση του συστήματος αποτελείται από την κενή κατάσταση 0, που περιγράφει ένα αδρανές σύστημα, και από μία δυαδική αναπαράσταση, που δομείται σε δύο διαστάσεις: το Επίπεδο (Level) και τη Φάση (Phase). Το πλήθος των Επιπέδων είναι άπειρο και το κάθε ένα από αυτά αντιστοιχεί στον αριθμό n_L των πελατών που βρίσκονται μετά από τον τελευταίο πελάτη, προσανυζόμενο κατά μία ακόμα θέση. Η Φάση αποτελεί μία πεπερασμένη διάσταση, η οποία προκύπτει από την απεικόνιση των υπόλοιπων στοιχείων της κατάστασης, που προηγούνται του n_L και η ταξινόμησή τους γίνεται σύμφωνα με λεξικογραφική αναπαράσταση, κατά αντιστοιχία με μία, ένα προς ένα, απεικόνιση των στοιχείων του $(n_1, \dots, n_l, \dots, n_L)$ στο $\{1, 2, \dots, 2^{K-1}\}$. Γι' αυτό και στο παραπάνω σχήμα βλέπουμε ότι στη φάση 1 συμπεριλαμβάνονται όλα τα διανύσματα με μήκος 1, στη φάση 2 όλα εκείνα με μήκος 2 και 1^ο στοιχείο το 1, στη φάση 3 όλα εκείνα με μήκος 2 και 1^ο στοιχείο το 2 και, τέλος, στη φάση 4 όλα εκείνα με μήκος 3.

Ας δούμε τώρα την μετακίνηση μεταξύ των καταστάσεων του συστήματος. Οι μεταβάσεις μεταξύ των επιπέδων οφείλονται σε άφιξη ενός πελάτη, είτε που εγκατέλειψε το σύστημα, είτε που παρέμεινε για εξυπηρέτηση, υπό την προϋπόθεση ότι υπάρχουν εικονικοί πελάτες στο σύστημα και δεν υπάρχει υπέρβαση της τιμής K . Οι μεταβάσεις προς την κενή κατάσταση συμβαίνουν, αποκλειστικά, όταν υπάρχει μονάχα ένας πελάτης στο σύστημα (περίπτωση καταστάσεων $\{1\}, \{2\}, \{3\}, \{4\}$) και του οποίου η ολοκλήρωση εξυπηρέτησης αδειάζει το σύστημα στιγμιαία, ακόμα, και από τους εικονικούς πελάτες, που ενδεχομένως υπάρχουν. Οι μεταβάσεις μεταξύ των φάσεων μαρτυρούν μεταβολή του πλήθους των πραγματικών πελατών του συστήματος, οδηγώντας σε κατάσταση με μεγαλύτερο μήκος, στην περίπτωση που ο επόμενος νεοεισαχθείς παραμένει στο σύστημα ή οδηγώντας σε κατάσταση με μικρότερο μήκος, στην περίπτωση εξυπηρέτησης του πρώτου πελάτη της ουράς.

Έτσι, εάν φανταστούμε το διάγραμμα σαν ένα πλέγμα, τότε μπορούν να εμφανιστούν σε αυτό είτε κάθετες μεταβάσεις, είτε οριζόντιες, είτε ταυτόχρονες. Η ταξινόμηση των μεταβάσεων αυτών γίνεται ως εξής:

- Κάθετες Μεταβάσεις (Αλλαγή Επιπέδου)
 - Προς τα πάνω: με την είσοδο ενός εικονικού πελάτη.
 - Προς το 0: με την εξυπηρέτηση του μοναδικού πραγματικού πελάτη.
- Οριζόντιες Μεταβάσεις (Αλλαγή Φάσης)
 - Προς τα δεξιά: με την είσοδο ενός πραγματικού πελάτη, εφόσον δεν υπάρχουν εικονικοί πελάτες στο σύστημα και δεν υπάρχει υπέρβαση του κατωφλίου.
 - Προς τα αριστερά: με την εξυπηρέτηση ενός πραγματικού πελάτη.
- Ταυτόχρονες Μεταβάσεις (Αλλαγή Επιπέδου & Φάσης)
 - Προς τα κάτω και δεξιά: με την είσοδο ενός πραγματικού πελάτη, εφόσον υπάρχουν εικονικοί πελάτες στο σύστημα και δεν υπάρχει υπέρβαση του κατωφλίου.

4.3 Ανάλυση Φάσεων, Επιτρεπτές Μεταβάσεις και Προσδιορισμός της Στάσιμης Κατανομής

Η παρούσα ενότητα εστιάζει στον προσδιορισμό της στάσιμης κατανομής της Διδιάστατης Μαρκοβιανής Αλυσίδας. Οι αναλυτικές εκφράσεις των πιθανοτήτων στο οριακό τμήμα δεν θα δοθούν εκ νέου, καθώς δεν εξυπηρετούν τους σκοπούς της μελέτης του παρόντος κεφαλαίου. Ωστόσο, ο υπολογισμός αυτών είναι εφικτός και βασισμένος σε ό,τι έχει αναπτυχθεί στο Κεφάλαιο 3 της εργασίας. Η στάσιμη κατανομή παρουσιάζεται σε δύο μορφές: την Πινακογεωμετρική Μορφή (Matrix-Geometric Form) και τη Βαθμωτή Μορφή (Scalar Form). Για τον καθορισμό της δεσμευμένης κατανομής απαιτούνται οι από κοινού πιθανότητες. Ως εκ τούτου, κρίνεται ως πιο λειτουργική η βαθμωτή μορφή, καθώς διευκολύνει την παρουσίαση των πιθανοτήτων αυτών.

Στο εξής, ορίζουμε ως $L = n_y$ το Επίπεδο (Level) και ως PH τη Φάση (Phase), θέτοντας τα θεμέλια για το συμβολιστικό πλαίσιο του μοντέλου. Θεωρούμε π_m το διάνυσμα των στάσιμων πιθανοτήτων διάστασης 2^{K-1} , έτσι ώστε το j -οστό στοιχείο του να είναι $(\pi_m)_j = P(L = m, PH = j)$. Από την Πινακογεωμετρική προσέγγιση, για κάθε $m \geq K$, ισχύει η σχέση $\pi_m = \pi_K R^{m-K}$.

Ανάλογα με τη διαδικασία και τις τεχνικές που χρησιμοποιήθηκαν στο προηγούμενο κεφάλαιο της εργασίας, ο πίνακας ρυθμού R υπολογίζεται από τον τύπο $R = -\lambda B^{-1}$, με τον πίνακα B να αποτελεί το block-διαγώνιο στοιχείο του γεννήτορα, το οποίο περιλαμβάνει τους ρυθμούς μετάβασης που διατηρούν το επίπεδο σταθερό. Έτσι, στο εξεταζόμενο μοντέλο, όλα τα στοιχεία της διαγωνίου του B είναι ίσα με $-(\lambda + \mu)$, καθώς ως διαγώνια στοιχεία του γεννήτορα πίνακα εξασφαλίζουν ότι οι γραμμές του αθροίζουν στο μηδέν, ενώ, στα στοιχεία που βρίσκονται κάτω από τη διαγώνιο του B εμφανίζεται ο ρυθμός μ , ο οποίος αντιστοιχεί στην ολοκλήρωση μίας εξυπηρέτησης, δηλαδή σε μία μετάβαση από την κατάσταση $(n_1, n_2, \dots, n_y)$ στην κατάσταση $(n_2, \dots, n_y)$.

Στο υπό μελέτη μοντέλο η αύξηση του επιπέδου καθορίζεται από τους πίνακες που βρίσκονται πάνω από τη διαγώνιο του γεννήτορα, οι οποίοι λαμβάνουν τη μορφή λI , υποδηλώνοντας μία διαδικασία αφίξεων ρυθμού λ .

Πιο συγκεκριμένα, για την περίπτωση που μελετήθηκε στην υποενότητα 4.2 και αφορά το κατώφλι $K = 3$, ο γεννήτορας πίνακας Q είναι:

$$Q = \begin{pmatrix} A_{0,0} & A_{0,1} & 0 & 0 & 0 & 0 \\ C & A_{1,1} & A_{1,2} & 0 & 0 & 0 \\ C & A_{2,1} & B & A_{2,3} & 0 & 0 \\ C & 0 & 0 & B & \lambda I_4 & 0 \\ C & 0 & 0 & 0 & B & \lambda I_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix},$$

ενώ, οι υποπίνακες του γεννήτορα είναι:

$$B = \begin{pmatrix} -(\lambda + \mu) & 0 & 0 & 0 \\ \mu & -(\lambda + \mu) & 0 & 0 \\ \mu & 0 & -(\lambda + \mu) & 0 \\ 0 & \mu & 0 & -(\lambda + \mu) \end{pmatrix}, \quad C = \begin{bmatrix} \mu \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad A_{0,0} = (-\lambda), \quad A_{0,1} = (\lambda \quad 0 \quad 0 \quad 0)$$

και $A_{1,1} = B + \lambda(E_{1,2} + E_{2,4})$, $A_{1,2} = \lambda I_4 - \lambda(E_{1,1} + E_{2,2})$, $A_{2,1} = \lambda E_{1,3}$, $A_{2,3} = \lambda I_4 - \lambda E_{1,1}$, με $E_{m \times n}$ να είναι ένας 4×4 πίνακας με μονάδα στην αντίστοιχη θέση $m \times n$ και 0 οπουδήποτε αλλού.

Ακολουθεί η αναζήτηση ενός τύπου για τον πίνακα ρυθμού R υψωμένο στη δύναμη $m-K$. Ο υπολογισμός αυτός είναι απαραίτητος για τον προσδιορισμό της στάσιμης κατανομής, καθώς η πιθανότητα το σύστημα να βρίσκεται σε ένα επίπεδο $m \geq K$ εξαρτάται άμεσα από τον πίνακα αυτόν. Στην πορεία, θα προσπαθήσουμε να βρούμε έναν τρόπο υπολογισμού των στοιχείων του, καθώς η απευθείας ύψωση του R σε δυνάμεις δεν είναι εύκολη διαδικασία. Πριν προχωρήσουμε, να αναφέρουμε, στο σημείο αυτό, ότι ο R είναι κάτω τριγωνικός, συνεπώς, το ίδιο ισχύει και για τις δυνάμεις του.

Στη συνέχεια, θα αναλύσουμε τη σχέση μεταξύ του αριθμού των πραγματικών πελατών και των φάσεων του συστήματος, δημιουργώντας το απαραίτητο υπόβαθρο για την παρουσίαση και την απόδειξη της πρότασης που ακολουθεί. Έτσι, για να συνεχίσουμε, χρειάζεται να συσχετίσουμε τις φάσεις του συστήματος με τον αριθμό των παρόντων πελατών, που τις απαρτίζουν.

Ορίζουμε, λοιπόν, ως $n(j)$ τον αριθμό των πραγματικών πελατών που αντιστοιχούν σε μία συγκεκριμένη φάση j , με $1 \leq j \leq 2^{K-1}$. Είναι σημαντικό να επισημάνουμε ότι δεν έχουν όλες οι φάσεις τον ίδιο αριθμό πελατών, οπότε η τιμή $n(j)$ λειτουργεί ως ένας δείκτης ταξινόμησης, ο οποίος μας επιτρέπει να οργανώσουμε τις φάσεις με βάση το πλήθος των πελατών τους. Ο υπολογισμός του $n(j)$ δίνεται από τη σχέση:

$$n(j) = \min \left\{ n : \sum_{l=0}^{n-1} \binom{K-1}{l} \geq j \right\}.$$

Για να γίνει απολύτως κατανοητός ο αλγεβρικός τρόπος υπολογισμού, εξετάζουμε την περίπτωση του μοντέλου με κατώφλι $K=3$, όπου έχουμε $2^{3-1} = 4$ φάσεις. Εφαρμόζοντας τον παραπάνω τύπο, παρατηρούμε τα εξής:

- Για τη φάση $j=1$: Το άθροισμα για $n=1$ δίνει $\binom{2}{0} = 1$. Εφόσον $1 \geq 1$, η φάση αντιστοιχεί στο $n(1)=1$.

- Για τη φάση $j=2$: Το άθροισμα για $n=1$ δίνει $1 < 2$. Προχωρώντας στο $n=2$, το άθροισμα γίνεται $\binom{2}{0} + \binom{2}{1} = 1 + 2 = 3$. Εφόσον $3 \geq 2$, η φάση αντιστοιχεί στο $n(2)=2$.
- Για τη φάση $j=3$: Ακολουθώντας την ίδια λογική, εφόσον το άθροισμα για $n=2$ είναι 3, το οποίο ικανοποιεί τη συνθήκη $3 \geq 3$, η φάση αυτή αντιστοιχεί επίσης στο $n(3)=2$.
- Για τη φάση $j=4$: Το άθροισμα για $n=2$ είναι $3 < 4$. Προχωρώντας στο $n=3$, το άθροισμα γίνεται $1+2+1=4$. Εφόσον $4 \geq 4$, η φάση αντιστοιχεί στο $n(4)=3$.

Αφού ορίστηκαν οι φάσεις, το επόμενο βήμα είναι να προσδιοριστούν ποιες μεταβάσεις μεταξύ τους είναι εφικτές. Θεωρούμε, λοιπόν, ως $A(j)$ το σύνολο όλων των φάσεων i , τέτοιων ώστε μία μετάβαση μεταξύ των φάσεων i και j -συμπεριλαμβανομένων πιθανών αλλαγών επιπέδου- να μπορεί να πραγματοποιηθεί, χωρίς το σύστημα να οδηγηθεί σε κατάσταση μηδέν (δηλαδή να αδειάσει τελείως η ουρά). Οπότε, θεωρώντας ως $a_j = (n_{1,j}, n_{2,j}, \dots, n_{n(j),j})$ το διάνυσμα μήκους $n(j)$ που αντιπροσωπεύει την φάση j , $1 \leq j \leq 2^{K-1}$, για οποιοδήποτε δεδομένο επίπεδο $n_{n(j)}$, το σύνολο αυτό $A(j)$ ορίζεται ως:

$$A(j) = \{i : a_i = (n_{1,i}, n_{2,i}, \dots, n_{n(i)-n(j)}, a_j)\}.$$

Με άλλα λόγια, το $A(j)$ περιλαμβάνει όλες εκείνες τις φάσεις i που περιέχουν την φάση j ως υποακολουθία στο τέλος της δομής τους. Για να διευκρινιστεί αυτό ακριβώς το σημείο, η φάση j μπορεί να παρουσιαστεί ως η φάση i , εφόσον απομακρυνθούν οι πρώτες $w = w(i, j)$ συνιστώσες της, όπου $w \geq 0$ και $w = n(i) - n(j)$. Επιπλέον, για να ανήκει μία φάση i στο σύνολο $A(j)$, δηλαδή $i \in A(j)$, θα πρέπει το διάνυσμα a_i της φάσης i να έχει μήκος $n(i)$ μεγαλύτερο ή ίσο από το μήκος $n(j)$ του διανύσματος a_j , $n(i) \geq n(j)$. Αυτό σημαίνει αλγεβρικά ότι αν από την αρχή του διανύσματος a_i απομακρυνθούν ακριβώς αυτές οι πρώτες w συνιστώσες του και το υπόλοιπο τμήμα που απομένει συμπίπτει με το a_j .

Για ακόμα μία φορά, για να γίνει πιο ξεκάθαρη η εικόνα του πώς αυτό αποτυπώνεται στην πράξη και πώς δομούνται εσωτερικά τα στοιχεία των διανυσμάτων, θα συνεχιστεί η μελέτη του μοντέλου με κατώφλι $K=3$. Σε αυτή την περίπτωση, όπου υπάρχουν συνολικά $2^{3-1} = 4$ φάσεις, το σύνολο των επιτρεπτών μεταβάσεων, για κάθε φάση, διαμορφώνεται ως εξής:

- Για τη φάση $j=1$: Έχουμε $n(1)=1$ και $a_1 = (n_{1,1})$. Έτσι, επειδή όλες οι φάσεις του συστήματος διαθέτουν τουλάχιστον έναν πραγματικό πελάτη $n(i) \geq 1$, το a_1 περιέχεται ως κατάληξη σε κάθε a_i . Άρα, ισχύει $A(1) = \{1, 2, 3, 4\}$.

- Για τη φάση $j = 2$: Έχουμε $n(2) = 2$ και $\alpha_2 = (n_{1,2}, n_{2,2})$. Αποκλείεται η επιλογή της φάσης 1, διότι $n(1) = 1 < 2 = n(2)$ και η επιλογή της φάσης 3, διότι παρόλο που $n(2) = n(3) = 2$, εντούτοις, έχουμε $\alpha_2 \neq \alpha_3$. Αντίθετα, περιέχεται η φάση 2, για προφανείς λόγους, αφού ταυτίζεται με τον εαυτό της, αλλά και η φάση 4, αφού $w(4, 2) = n(4) - n(2) = 3 - 2 = 1$, το οποίο σημαίνει η αφαίρεση της συνιστώσας από το α_4 αφήνει ένα τμήμα που συμπίπτει με το α_2 . Άρα, ισχύει $A(2) = \{2, 4\}$.
- Για τη φάση $j = 3$: Έχουμε $n(3) = 2$ και $\alpha_3 = (n_{1,3}, n_{2,3})$. Κατ' ανάλογο τρόπο με την προηγούμενη περίπτωση, προκύπτει ότι $A(3) = \{3, 4\}$.
- Για τη φάση $j = 4$: Έχουμε $n(4) = 3$ και $\alpha_4 = (n_{1,4}, n_{2,4}, n_{3,4})$. Αποκλείεται η επιλογή των φάσεων 1, 2, 3, καθώς έχουν μήκος διανύσματος μικρότερο από αυτό της φάσης 4, $n(1) = 1 < 3, n(2) = n(3) = 2 < 3$. Περιλαμβάνεται μόνο η φάση 4, αφού για $w(4, 4) = 0$, το διάνυσμα ταυτίζεται με τον εαυτό του. Άρα, ισχύει $A(4) = \{4\}$.

Με βάση τα παραπάνω, οδηγούμαστε στην ακόλουθη πρόταση.

Πρόταση 4.1: Έστω δύο φάσεις i και j , με $i \in A(j)$. Έχουμε:

$$(R^{m-K})_{ij} = \binom{m-K+w-1}{w} \left(\frac{\mu}{\mu+\lambda} \right)^w \left(\frac{\lambda}{\mu+\lambda} \right)^{m-K},$$

όπου $w = w(i, j)$. Για $i \notin A(j)$, ισχύει $(R^{m-K})_{ij} = 0$, για όλα τα j .

Απόδειξη: Υπάρχουν δύο προσεγγίσεις για την απόδειξη της πρότασης. Η πρώτη είναι τεχνική και περιλαμβάνει τις ακριβείς λεπτομέρειες του πολλαπλασιασμού πινάκων. Επιλέγεται να παραληφθούν οι λεπτομέρειες αυτές και να ακολουθηθεί η δεύτερη εκδοχή της πιθανοσυνδυαστικής προσέγγισης. Σε αυτή την προσέγγιση, παρατηρείται η διαδικασία σε διακριτά στάδια εστιάζοντας

σε μία ισοδύναμη ΜΑΔΧ με πίνακα μετάβασης $P = I + \frac{1}{\lambda + \mu} Q$.

Στον χώρο καταστάσεων του συστήματος, οι φάσεις αντιπροσωπεύουν τη διάταξη των πραγματικών και των εικονικών πελατών. Όταν το επίπεδο αυξάνεται εξαιτίας της αφίξης νέου πελάτη που αποφασίζει να παραμείνει, το σύστημα μεταβαίνει σε υψηλότερο επίπεδο, ενώ η εσωτερική φάση καθορίζεται από την δομή των πελατών.

Σε επίπεδα μεγαλύτερα ή ίσα του ορίου K , παρατηρείται ότι οι αλλαγές στις φάσεις μπορούν να γίνουν μόνο προς μία κατεύθυνση μείωσης. Αυτό συμβαίνει διότι οι μεταβάσεις που επηρεάζουν την εσωτερική φάση πραγματοποιούνται αποκλειστικά και μόνο κατόπιν ολοκλήρωσης μίας εξυπηρέτησης, η οποία μειώνει την φάση. Αντίθετα, η αλλαγή επιπέδου οφείλεται αποκλειστικά στις αφίξεις.

Αν προβάλλουμε την διαδικασία αποκλειστικά στον άξονα των φάσεων, διαπιστώνουμε ότι υπάρχει μία συγκεκριμένη πορεία μέχρι να φτάσει η διαδικασία στο όριο. Ειδικότερα, οι μεταβάσεις χωρίζονται σε δύο διακριτές κατηγορίες, στις αφίξεις με ρυθμό λ , οι οποίες αυξάνουν το επίπεδο κατά μία μονάδα χωρίς να μεταβάλλουν τη φάση και στις ολοκληρώσεις εξυπηρέτησης με ρυθμό μ , οι οποίες μειώνουν τη φάση. Επειδή η αλλαγή επιπέδου γίνεται αποκλειστικά λόγω αφίξεων και η αλλαγή φάσης αποκλειστικά λόγω εξυπηρέτησεων, για να μεταβούμε από το βασικό επίπεδο K στο επιθυμητό επίπεδο m κάνοντας συνολικά w βήματα μείωσης φάσης, η διαδικασία ακολουθεί συγκεκριμένες διαδρομές.

Γενικά, ο δείκτης $w = w(i, j)$ εκφράζει τη συνολική μείωση των πραγματικών πελατών και μετράει πόσες φορές η εσωτερική διάταξη μετατοπίστηκε προς λιγότερους πελάτες κατά τη διάρκεια ανόδου του επιπέδου.

Στην ειδική περίπτωση, όπου $i = j$, (δηλαδή η αρχική και η τελική φάση συμπίπτουν $w = 0$), το επίπεδο αυξάνεται κατά $m - K$ χωρίς αλλαγή φάσης, πράγμα που σημαίνει ότι υπήρξαν ακριβώς $m - K$ αφίξεις χωρίς καμία ενδιάμεση ολοκλήρωση εξυπηρέτησης.

Για να κατανοήσουμε πλήρως τον μηχανισμό των πιθανοτήτων, εξετάζουμε αναλυτικά τα επιμέρους τμήματα του γινομένου του προς απόδειξη τύπου.

Αρχικά, ο λόγος $\frac{\mu}{\mu + \lambda}$ εκφράζει την πιθανότητα να συμβεί μία ολοκλήρωση εξυπηρέτησης σε σχέση με οποιοδήποτε γεγονός στο σύστημα, ενώ η ύψωσή του στη δύναμη w , δηλαδή ο όρος $\left(\frac{\mu}{\mu + \lambda}\right)^w$ αντιπροσωπεύει τη συνολική συνεισφορά των w διαδοχικών ολοκληρώσεων που προκάλεσαν τις ισάριθμες μειώσεις φάσης.

Αντίστοιχα, ο λόγος $\frac{\lambda}{\mu + \lambda}$ εκφράζει την πιθανότητα να συμβεί μία άφιξη πελάτη και η ύψωσή του στην δύναμη $m - K$, δηλαδή ο όρος $\left(\frac{\lambda}{\mu + \lambda}\right)^{m-K}$ αποτυπώνει τη αλληλουχία των $m - K$ αφίξεων που ανέβασαν το επίπεδο του συστήματος από το K στο m .

Για να βρούμε με πόσους τρόπους μπορούν να συνδυαστούν αυτές οι αφίξεις και οι εξυπηρέτησεις, αξιοποιούμε τη συνδυαστική ανάλυση, όπου ο συνολικός αριθμός των διαφορετικών τροχιών που αλλάζουν το επίπεδο κατά $m - K$, ταυτόχρονα με την πραγματοποίηση ολοκληρώσεων εξυπηρέτησης, δίνεται από τον διωνυμικό συντελεστή $\binom{m-K+w-1}{w}$.

Ο πολλαπλασιασμός του συνολικού αριθμού των δυνατών τροχιών επί τις πιθανότητες των επιμέρους γεγονότων προκύπτει από την Αρχή Απαρίθμησης και τον Πολλαπλασιαστικό Κανόνα των Πιθανοτήτων για ανεξάρτητα γεγονότα. Έτσι, δεδομένου ότι κάθε συγκεκριμένη τροχιά

αποτελεί μία ανεξάρτητη ακολουθία από $m-K$ αφίξεις και w εξυπηρετήσεις με σταθερές πιθανότητες εμφάνισης, καταλήγουμε στον τύπο:

$$(R^{m-K})_{ij} = \binom{m-K+w-1}{w} \left(\frac{\mu}{\mu+\lambda} \right)^w \left(\frac{\lambda}{\mu+\lambda} \right)^{m-K}.$$

□

Υπενθυμίζοντας ότι το σύμβολο $(\pi_m)_i$ εκφράζει την από κοινού πιθανότητα του επιπέδου m και της φάσης i , η Πρόταση 4.2, που ακολουθεί, αξιοποιεί απευθείας αυτό το αποτέλεσμα για να προσδιορίσει τις πιθανότητες της στάσιμης κατανομής σε οποιοδήποτε επίπεδο m , πέρα από τιμές του ορίου K .

Πρόταση 4.2: Για $1 \leq i \leq 2^{K-1}$ μπορούμε να γράψουμε:

$$(\pi_m)_j = \sum_{i \in A(j)} (\pi_K)_i \binom{m-K+w-1}{w} \left(\frac{\mu}{\mu+\lambda} \right)^w \left(\frac{\lambda}{\mu+\lambda} \right)^{m-K}, \quad m \geq K+1,$$

όπου $(\pi_K)_i$ μπορεί να προκύψει μέσω ενός συστήματος γραμμικών εξισώσεων, που βρίσκεται από την οριακή συνθήκη.

Απόδειξη: Η απόδειξη προκύπτει άμεσα από πινακογεωμετρική δομή της στοχαστικής διαδικασίας σε συνδυασμό με το αποτέλεσμα της Πρότασης 4.1.

Από την θεωρία του μοντέλου, η στάσιμη κατανομή για τα επίπεδα που ξεπερνούν το όριο εισόδου, εκφράζεται συναρτήσει του πίνακα ρυθμού R ως $\pi_m = \pi_K R^{m-K}$, όπου π_m είναι το διάνυσμα πιθανοτήτων του επιπέδου.

Απομονώνοντας το j -οστό στοιχείο αυτού του διανύσματος, ο πολλαπλασιασμός πινάκων δίνει το άθροισμα $(\pi_m)_j = \sum_i (\pi_K)_i (R^{m-K})_{ij}$. Επειδή οι μεταβάσεις προς j είναι επιτρεπτές όταν $i \in A(j)$, τότε για κάθε άλλη περίπτωση ο αντίστοιχος όρος μηδενίζεται, περιορίζοντας το άθροισμα σε:

$$(\pi_m)_j = \sum_{i \in A(j)} (\pi_K)_i (R^{m-K})_{ij}.$$

Τέλος, αντικαθιστώντας τον όρο $(R^{m-K})_{ij}$ με τον τύπο της Πρότασης 4.1, προκύπτει η προς απόδειξη σχέση:

$$(\pi_m)_j = \sum_{i \in A(j)} (\pi_K)_i \binom{m-K+w-1}{w} \left(\frac{\mu}{\mu+\lambda} \right)^w \left(\frac{\lambda}{\mu+\lambda} \right)^{m-K}, \quad m \geq K+1,$$

όπου $w = w(i, j)$ παριστάνει την διαφορά των πραγματικών πελατών, μεταξύ των φάσεων i και j .

□

Ολοκληρώνοντας την ανάλυση της συμπεριφοράς του συστήματος και έχοντας προσδιορίσει πλήρως τις εκφράσεις της στάσιμης κατανομής, μέσω των προηγούμενων προτάσεων, το σύστημα διαθέτει πλέον την απαιτούμενη περιγραφή για την εξέλιξη των καταστάσεων. Η γνώση αυτής της συμπεριφοράς μας προετοιμάζει ώστε, στο επόμενο τμήμα, να προχωρήσουμε στην εξαγωγή της έκφρασης για τον δεσμευμένο χρόνο αναμονής, δεδομένου του συνολικού αριθμού πελατών, που παρατηρεί κατά την άφιξη του ένας νεοεισρχθείς πελάτης.

4.4 Δεσμευμένη Κατανομή των Πραγματικών Πελατών και Ανάδειξη της Αναγκαιότητας του Διπλού Κατωφλιού

Σε συνέχεια της προηγούμενης ενότητας, προχωράμε τώρα, με στόχο τον καθορισμό της κατανομής του χρόνου αναμονής, αλλά και την εύρεση της καλύτερης δυνατής αντιμετώπισης από την πλευρά του πελάτη που εισέρχεται στο σύστημα, με δεδομένες τις πληροφορίες που του παρέχονται κατά την άφιξή του.

Για να το πετύχουμε αυτό, θυμίζουμε ότι έχουμε ορίσει ως N το πλήθος των πραγματικών πελατών του συστήματος, οι οποίοι κατά την είσοδό τους έλαβαν την απόφαση να μην εγκαταλείψουν την ουρά, αλλά και ως D το συνολικό πλήθος όλων των πελατών (πραγματικών και εικονικών μαζί).

Ανατρέχοντας στη διδιάστατη ΜΑΣΧ που αναλύθηκε στην προηγούμενη υποενότητα, κάθε ζεύγος (m, j) αντιστοιχεί μοναδικά σε μία συγκεκριμένη κατάσταση του συστήματος, όπου το m αντιπροσωπεύει το επίπεδο (level) και λαμβάνει τιμές πέραν του ορίου K ($m \geq K+1$) και το j τη φάση (phase), η οποία κυμαίνεται από 1 έως 2^{K-1} ($1 \leq j \leq 2^{K-1}$). Το διάνυσμα $(n_1, n_2, \dots, n_{n(j)})$ περιγράφεται από το ζεύγος (m, j) και προκύπτει από τον τύπο (4.1), που παρουσιάστηκε στην υποενότητα 4.2 του παρόντος κεφαλαίου, λαμβάνοντας την παράμετρο y ως $n(j)$.

Επομένως, βάσει αυτής της δομής, ο συνολικός αριθμός πελατών ισούται με το άθροισμα των επιμέρους στοιχείων του διανύσματος, δηλαδή $D = \sum_{i=1}^{n(j)} n_i$, ενώ ο αριθμός των πραγματικών πελατών N αντιστοιχεί ακριβώς στη διάσταση $n(j)$.

Έπειτα, θεωρούμε $X_i, 1 \leq i \leq n(j)$, ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, κάθε μία από τις οποίες εκφράζει τον χρόνο εξυπηρέτησης κάποιου πελάτη και ακολουθεί εκθετική κατανομή παραμέτρου μ . Έτσι, αφού ο χρόνος αναμονής W για έναν αφιχθέντα πελάτη του συστήματος περιλαμβάνει τους χρόνους εξυπηρέτησης όλων των πραγματικών πελατών που βρίσκει κατά την άφιξή του, τότε $W = \sum_{i=1}^{n(j)} X_i$.

Επομένως, στο εξής, επικεντρώνεται η προσοχή μας στη δεσμευμένη κατανομή του πλήθους των πραγματικών πελατών N , δεδομένου του συνολικού πλήθους D , το οποίο θα αποτελέσει και το κύριο εργαλείο στην πορεία του κεφαλαίου. Για να οδηγηθούμε στα βασικά θεωρήματα που ακολουθούν και να είμαστε σε θέση να τα αποδείξουμε, απαιτείται πρώτα να παρουσιαστούν τρία λήμματα και κάποιοι ορισμοί.

Ακολουθεί η παρουσίαση των σημαντικών αυτών βοηθητικών εργαλείων και εννοιών για την επίτευξη της διατύπωσης αλλά και απόδειξης των θεωρημάτων που βρίσκονται παρακάτω στο παρόν υποκεφάλαιο.

Αρχικά, θεωρούμε S_d το σύνολο που περιλαμβάνει όλες εκείνες τις καταστάσεις στις οποίες το συνολικό πλήθος των πελατών D ισούται με d , για τιμές $d \geq K$. Το σύνολο αυτό εκφράζεται ως:

$$S_d = \bigcup_{y=1}^K \left\{ \mathbf{n} \in \mathbf{N}^y : n_i \geq 1, \sum_{i=1}^y n_i = d, \quad i=1, \dots, y \right\}, \quad d \geq K.$$

Επίσης, θεωρούμε A_d το πλήθος των πραγματικών πελατών του συστήματος, κατά την χρονική εκείνη στιγμή όπου για πρώτη φορά μέσα σε μία τυχαία περίοδο απασχόλησης (busy period), το σύστημα φτάνει σε μία κατάσταση που ανήκει στο παραπάνω σύνολο S_d .

Ευθύς αμέσως, διατυπώνονται τα τρία απαραίτητα βοηθητικά λήμματα.

Λήμμα 4.1: Για οποιαδήποτε τιμή του συνολικού πλήθους πελατών d , όπου $d \geq K$, το σύνολο καταστάσεων S_d περιέχει ακριβώς 2^{K-1} διαφορετικές καταστάσεις. Κάθε μία από αυτές τις καταστάσεις αντιστοιχεί σε μία διαφορετική φάση. Επίσης, η κατάσταση που αντιστοιχεί στην τελευταία φάση 2^{K-1} είναι εκείνη που διαθέτει την μικρότερη τιμή για το επίπεδο n_y και ισούται ακριβώς με $d - (K - 1)$.

Απόδειξη: Η αποδεικτική διαδικασία διακρίνεται σε δύο σκέλη. Στο πρώτο σκέλος, αποδεικνύουμε γιατί το σύνολο των καταστάσεων περιέχει ακριβώς 2^{K-1} καταστάσεις. Αρχικά, λαμβάνουμε υπόψη μας την δομή του χώρου καταστάσεων και παρατηρούμε πώς μεταβάλλονται οι τιμές, όταν κινούμαστε μέσα στην ίδια φάση. Εντός μίας συγκεκριμένης φάσης, οι τιμές που μπορούν να πάρουν οι επιμέρους καταστάσεις διαφέρουν αποκλειστικά και μόνο ως το τελευταίο επίπεδο n_y και καθώς μεταβαίνουμε από το ένα επίπεδο στο αμέσως επόμενο, το συνολικό πλήθος των πελατών d , αυξάνεται σταδιακά κατά μία μονάδα, με την χαμηλότερη τιμή να βρίσκεται στο επίπεδο 1. Επίσης, ανάλογα με τη φάση, το d του επιπέδου 1 μπορεί να πάρει αέριαιες τιμές μεταξύ των αριθμών 1 και K . Οπότε, με δεδομένο ότι κάθε διαδοχική αύξηση επιπέδου προσθέτει ακριβώς μία μονάδα στο d , προκύπτει ότι για οποιαδήποτε επιλεγμένη τιμή του d , με $d \geq K$, έχουμε ακριβώς μία αντίστοιχη κατάσταση σε κάθε μία από τις συνολικά 2^{K-1} διακριτές φάσεις του συστήματος. Συνεπώς, το S_d περιλαμβάνει ακριβώς 2^{K-1} καταστάσεις, αντιστοιχίζοντας κάθε μία από αυτές σε μία μοναδική κατάσταση της φάσης.

Στο δεύτερο σκέλος, στόχος μας είναι να αποδειχθεί ότι η μικρότερη τιμή του n_y ισούται με $d - (K - 1)$ και βρίσκεται στην κατάσταση με φάση 2^{K-1} . Εστιάζουμε, λοιπόν, στην τελευταία φάση του συστήματος, που είναι η 2^{K-1} , και στην οποία η εσωτερική δομή των προηγούμενων στοιχείων, πλην του τελευταίου, είναι τέτοια ώστε το άθροισμα τους να φτάνει το μέγιστο επιτρεπόμενο όριο που επιβάλλει ο χώρος καταστάσεων. Αυτό σημαίνει ότι, το διάνυσμα κατάστασης αποτελείται από $K-1$ μονάδες πριν το τελευταίο στοιχείο του n_y , ώστε το σύστημα να φτάσει στην μέγιστη χωρητικότητα του K . Τώρα, για να έχουμε συνολικό άθροισμα d , θα πρέπει το n_y να πάρει την τιμή $d - \sum_{i=1}^{y-1} n_i$. Η τιμή της διαφοράς αυτής γίνεται ελάχιστη όταν το άθροισμα μεγιστοποιείται, δηλαδή όταν $\sum_{i=1}^{y-1} n_i = K - 1$. Συνεπώς, η κατάσταση με τη φάση 2^{K-1} έχει τη χαμηλότερη τιμή του n_y , που είναι $d - (K - 1)$.

□

Λήμμα 4.2: Για $d \geq 2K - 1$, κατά τη διάρκεια μίας περιόδου απασχόλησης, ο αριθμός των πραγματικών πελατών στο σύστημα μειώνεται.

Απόδειξη: Ο συνολικός αριθμός των πελατών, που αποτελεί το συνολικό άθροισμα των στοιχείων του διανύσματος, ισούται με: $d = \sum_{i=1}^y n_i \Leftrightarrow d = \sum_{i=1}^{y-1} n_i + n_y$. Άρα, ισχύει: $d - n_y = \sum_{i=1}^{y-1} n_i$.

Από την κατασκευή του χώρου καταστάσεων, που αποτυπώνεται στην σχέση (4.1), γνωρίζουμε ότι υπάρχει είσοδος νέου πελάτη μόνο εάν η ένδειξη είναι μικρότερη από την τιμή K , το οποίο ισοδυναμεί με: $\sum_{i=1}^{y-1} n_i \leq K - 1$.

Έτσι, οι δύο παραπάνω σχέσεις μας δίνουν: $d - n_y \leq K - 1$. Άρα, $d \leq n_y + K - 1$.

Από υπόθεση, θεωρούμε ότι $d \geq 2K - 1$ και συνδυάζοντας τις δύο ανισοτικές σχέσεις, παίρνουμε $2K - 1 \leq d \leq n_y + K - 1$.

Οπότε, ισχύει: $2K - 1 \leq n_y + K - 1$. Άρα, $n_y \geq K$, το οποίο σημαίνει ότι οι εικονικές θέσεις μετά τον τελευταίο πραγματικό πελάτη είναι τουλάχιστον K . Επομένως, οποιοσδήποτε νέος πελάτης φτάσει θα αποθαρρυνθεί και θα αποχωρήσει.

Ως εκ τούτου, η συνθήκη $n_y \geq K$ είναι αναγκαία και ικανή ώστε κανένας νέος πελάτης να μην μπορεί να εισέλθει στην ουρά μέχρι το σύστημα να αδειάσει, αφού $d \geq K$ για το υπόλοιπο της περιόδου απασχόλησης. Επειδή, όμως δεν μπορούν να συμβούν αφίξεις, αφού το σύστημα δεν δέχεται καμία προσθήκη, τότε η μόνη επιλογή που υπάρχει είναι να συμβούν αποχωρήσεις, λόγω ολοκληρώσεων εξυπηρέτησης των παρευρισκομένων πελατών. Έτσι, καταλήγουμε στο ότι ο αριθμός των πραγματικών πελατών μπορεί μονάχα να μειώνεται.

□

Πριν προχωρήσουμε στην διατύπωση του επόμενου λήμματος, υπενθυμίζουμε ότι με A_d συμβολίζεται η τυχαία μεταβλητή που αντιπροσωπεύει τον αριθμό των πραγματικών πελατών στο σύστημα κατά την χρονική στιγμή όπου, για πρώτη φορά σε μία περίοδο απασχόλησης το σύστημα φτάνει σε κατάσταση με συνολικό πλήθος d , ενώ το σύμβολο $\leq_{st}$ εκφράζει μία στοχαστική ανισότητα μεταξύ των δύο εκατέρωθεν μεγεθών.

Λήμμα 4.3: Για $d \geq 2K - 1$, ισχύει $A_{d+1} \leq_{st} A_d$.

Απόδειξη: Όπως είδαμε στο Λήμμα 4.2, για $d \geq 2K - 1$, εντός μίας περιόδου απασχόλησης, δεν πραγματοποιούνται αφίξεις, παρά μόνο το σύστημα χάνει σταδιακά τον αριθμό των πελατών του. Λόγω αυτής της φθίνουσας πορείας του συστήματος, για να υπάρχουν τουλάχιστον n πελάτες, όταν βρισκόμαστε στο S_{d+1} , υποχρεωτικά θα πρέπει να υπήρχαν προηγουμένως τουλάχιστον n πελάτες, όταν βρισκόμασταν στο S_d . Που σημαίνει ότι το γεγονός $\{A_{d+1} \geq n\}$ είναι υποσύνολο του $\{A_d \geq n\}$, αφού όποτε συμβαίνει το πρώτο, συμβαίνει υποχρεωτικά και το δεύτερο. Άρα, η τομή των δύο αυτών συμπίπτει με το ίδιο το $\{A_{d+1} \geq n\}$.

Έπειτα, από την θεωρία των δεσμευμένων πιθανοτήτων, γνωρίζουμε ότι ισχύει το εξής:

$$P(A_{d+1} \geq n | A_d \geq n) = \frac{P(\{A_{d+1} \geq n\} \cap \{A_d \geq n\})}{P(A_d \geq n)}.$$

Όμως, λόγω της έννοιας του υποσυνόλου, που αναφέρθηκε παραπάνω, η οποία μας δίνει την ισότητα $P(\{A_{d+1} \geq n\} \cap \{A_d \geq n\}) = P(A_{d+1} \geq n)$, οδηγούμαστε στην σχέση:

$$P(A_{d+1} \geq n) = P(A_d \geq n) \cdot P(A_{d+1} \geq n | A_d \geq n).$$

Τέλος, επειδή η δεσμευμένη πιθανότητα του δεξιού μέλους δεν παύει να είναι μία πιθανότητα, τότε η τιμή της εξ' ορισμού είναι μικρότερη ή ίση με 1, δηλαδή $P(A_{d+1} \geq n | A_d \geq n) \leq 1$.

Οπότε, ισχύει: $P(A_d \geq n) \cdot P(A_{d+1} \geq n | A_d \geq n) \leq P(A_d \geq n) \cdot 1$.

Έτσι, φτάσαμε σε: $P(A_{d+1} \geq n) \leq P(A_d \geq n)$, το οποίο σημαίνει ότι η πιθανότητα να έχουμε τουλάχιστον n πελάτες είναι πάντα μικρότερη ή ίση στην επόμενη κατάσταση $d + 1$, σε σύγκριση με την προηγούμενη κατάσταση d . Οπότε, η A_{d+1} τείνει να παίρνει στοχαστικά μικρότερες τιμές από την A_d . Συνεπώς, $A_{d+1} \leq_{st} A_d$.

□

Πλέον, έχοντας παρουσιάσει τα τρία απαραίτητα λήμματα, μπορούμε να προχωρήσουμε στην διατύπωση και απόδειξη του θεωρήματος, που σχετίζεται με την δεσμευμένη κατανομή των πραγματικών πελατών, με δεδομένο το πλήθος των συνολικών πελατών του συστήματος.

Θεώρημα 4.1: Για $d \geq 2K - 1$, η δεσμευμένη κατανομή $N | D = d$ είναι στοχαστικά φθίνουσα ως προς d . Επιπλέον, ισχύει $\lim_{d \rightarrow \infty} P(N = 1 | D = d) = 1$.

Απόδειξη: Η διαδικασία χωρίζεται σε δύο διακριτά αποδεικτικά μέρη, ένα που αφορά την μονοτονία της $N | D = d$ και ένα που αφορά το όριο $\lim_{d \rightarrow \infty} P(N = 1 | D = d)$.

Όσον αφορά το σκέλος της μονοτονίας, θα λέγαμε ότι βασίζεται στη σύγκριση της εξέλιξης του συστήματος για διαφορετικές τιμές του πλήθους των πελατών. Επισημαίνουμε ότι οι μελλοντικές μεταβάσεις και οι μηχανισμοί αλλαγής καταστάσεων της στοχαστικής διαδικασίας δεν εξαρτώνται από την τρέχουσα κατάσταση του συστήματος. Επειδή ο αριθμός των πραγματικών πελατών, την εκάστοτε δεδομένη στιγμή, εξαρτάται μόνο από τους προηγούμενους πραγματικούς πελάτες και όχι από τις αφίξεις εικονικών πελατών, συμπεραίνουμε ότι με δεδομένο το πλήθος των παρόντων ενεργών πελατών, η γνώση του πλήθους είτε των εικονικών, είτε των συνολικών πελατών, που βρίσκονται στο σύστημα, δεν επηρεάζει την μελλοντική συμπεριφορά του συστήματος. Η ιδέα αυτή αποτυπώνεται με την ισοδυναμία των κατανομών $N | D = d + 1, A_{d+1} = a$ και $N | D = d, A_d = a$, όπου a οποιοσδήποτε ακέραιος στο $\{1, 2, \dots, K\}$. Συνεπώς, η όποια διαφοροποίηση προκύπτει μεταξύ των περιπτώσεων $d - 1$ και d απορρέει απευθείας από σχέση των μεταβλητών A_{d-1} και A_d . Τώρα, λαμβάνουμε υπόψη το Λήμμα 4.3, το οποίο μέσω του τύπου $A_{d+1} \leq_{st} A_d$ αποδεικνύει ότι καθώς ο συνολικός αριθμός d μεγαλώνει, ο αριθμός των πραγματικών πελατών τείνει να είναι μικρότερος. Αυτό σημαίνει ότι μεγαλύτερο συνολικό πλήθος οδηγεί σε λιγότερους πραγματικούς πελάτες. Ωστόσο, κατά τη σύγκριση των $N | D = d$, η τιμή των συνολικών πελατών θεωρείται δεδομένη και εξετάζεται η δεσμευμένη κατανομή των πραγματικών πελατών. Συνεπώς, δεδομένου του $D = d$, η αντίστροφη αυτή σχέση συνεπάγεται ότι για να προκύψει μεγαλύτερη τιμή στην κατανομή του N , απαιτείται μικρότερο συνολικό πλήθος. Ως εκ τούτου, προκύπτει άμεσα ότι η κατανομή $N | D = d$ είναι στοχαστικά μεγαλύτερη από την $N | D = d - 1$. Έτσι, αποδείχθηκε ότι, για $d \geq 2K - 1$, η δεσμευμένη κατανομή $N | D = d$, είναι στοχαστικά φθίνουσα ως προς d .

Αφού ολοκληρώθηκε το πρώτο σκέλος της απόδειξης, τώρα θα εξετάσουμε την συμπεριφορά της δεσμευμένης κατανομής του N , καθώς ο παρατηρούμενος αριθμός πελατών μεγαλώνει απεριόριστα, τείνοντας στο άπειρο.

Με χρήση της Πρότασης 4.2, για $d \geq 2K - 1$, η από κοινού πιθανότητα $P(N = 1, D = d)$ μπορεί να πάρει την ακόλουθη πολυωνυμική μορφή ως προς d :

$$P(N = n, D = d) = \sum_{j: n(j)=n} \sum_{i \in A(j)} (\pi_K)_i \binom{m-K+w-1}{w} \left(\frac{\mu}{\mu+\lambda} \right)^w \left(\frac{\lambda}{\mu+\lambda} \right)^{m-K}, \quad (4.2)$$

όπου $m = d - \sum_{l=1}^{n(j)-1} n_l$.

Αναπτύσσοντας τον διωνυμικό παράγοντα και αντικαθιστώντας την τιμή του m αποκλειστικά στον εκθέτη του τελευταίου παράγοντα, η σχέση (4.2) παίρνει την μορφή:

$$P(N = n, D = d) = \sum_{j: n(j)=n} \sum_{i \in A(j)} (\pi_K)_i \frac{(m-K) \cdots (m-K+w-1)}{w!} \left(\frac{\mu}{\mu+\lambda} \right)^w \left(\frac{\lambda}{\mu+\lambda} \right)^{d - \sum_{l=1}^{n(j)-1} n_l - K}.$$

Τώρα, κοιτώντας τον αριθμητή του κλάσματος, αυτός αποτελείται από γινόμενο παραγόντων, το πλήθος των οποίων καθορίζει και τον βαθμό του ως προς m και κατ' επέκταση ως προς d , μιας και ο m ισούται με τον d μειούμενο κατά έναν σταθερό όρο. Το πλήθος των παραγόντων στον αριθμητή είναι $(m-K+w-1) - (m-K) + 1 = w$ άρα αυτό που προέκυψε από τον διωνυμικό όρο είναι ένα πολυώνυμο βαθμού w ως προς d .

Έπειτα, γνωρίζουμε ότι ο μέγιστος αριθμός πραγματικών πελατών είναι K , δηλαδή $n(i) \leq K$, ενώ, από την δομή των φάσεων, όπου κάθε φάση j αντιστοιχεί σε έναν συγκεκριμένο αριθμό πραγματικών πελατών $n(j)$, εξαιτίας του περιορισμού που θέτει το εξωτερικό άθροισμα του τύπου (4.2), λαμβάνουμε $n(j) = n$. Άρα, ισχύει $w = n(i) - n(j) \leq K - n$.

Επιπλέον, αξιοποιώντας μία από τις ιδιότητες των δυνάμεων στον τελευταίο όρο $\left(\frac{\lambda}{\mu+\lambda} \right)^{d - \sum_{l=1}^{n(j)-1} n_l - K}$, μπορούμε να απομονώσουμε τον $\left(\frac{\lambda}{\mu+\lambda} \right)^d$, αφήνοντας κατά μέρος τον σταθερό όρο $\left(\frac{\lambda}{\mu+\lambda} \right)^{-\sum_{l=1}^{n(j)-1} n_l - K}$. Έτσι, συμπεραίνουμε ότι, το άθροισμα είναι ένα πολυώνυμο βαθμού $K - n$ ως προς d , πολλαπλασιασμένο με τον γεωμετρικό όρο $\left(\frac{\lambda}{\mu+\lambda} \right)^d$.

Συνεχίζοντας, από τη θεωρία πιθανοτήτων, έχουμε την δεσμευμένη πιθανότητα:

$$P(N \geq n | D = d) = \frac{P(N \geq n, D = d)}{P(D = d)}, \quad n \geq 1.$$

Σε αντιστοιχία με αυτό που παραπάνω διατυπώθηκε για τον βαθμό του αθροίσματος στην από κοινού πιθανότητα, έχουμε ότι η απλή πιθανότητα $P(D = d)$ αποτελεί ένα πολυώνυμο βαθμού

$K - 1$ ως προς d , πολλαπλασιασμένο με τον γεωμετρικό όρο $\left(\frac{\lambda}{\mu+\lambda} \right)^d$.

Ως εκ τούτου, αφού απλοποιηθεί ο γεωμετρικός παράγοντας και από τους δύο όρους του κλάσματος, αυτό που απομένει είναι το κλάσμα των πολωνύμων. Η ρητή αυτή τελική μορφή έχει ένα πολυώνυμο βαθμού $K - n$ στον αριθμητή και ένα πολυώνυμο βαθμού $K - 1$ στον παρονομαστή.

Στην περίπτωση που $n > 1$, υπερಿಸχύει ο βαθμός του αριθμητή, οδηγώντας το όριο $\lim_{d \rightarrow \infty} P(N \geq n | D = d)$ στο 0. Ενώ, στην περίπτωση που $n = 1$, οι βαθμοί είναι ίδιοι και λόγω απλοποίησης το όριο δίνει 1. Οπότε:

$$\lim_{d \rightarrow \infty} P(N \geq n | D = d) = \begin{cases} 1, & n = 1, \\ 0, & n > 1. \end{cases}$$

Τώρα, με δεδομένο ότι η πιθανότητα να έχει το σύστημα ακριβώς n πραγματικούς πελάτες, υπό τη συνθήκη $D = d$, δίνεται από τον τύπο $P(N = n | D = d) = P(N \geq n | D = d) - P(N \geq n+1 | D = d)$, θα αναζητήσουμε την τιμή του ορίου $\lim_{d \rightarrow \infty} P(N = n | D = d)$ για κάθε περίπτωση χωριστά.

- Για $n = 1$, έχουμε: $P(N = 1 | D = d) = P(N \geq 1 | D = d) - P(N \geq 2 | D = d) = 1 - 0 = 1$.
- Ενώ, για $n > 1$, έχουμε: $P(N = n | D = d) = P(N \geq n | D = d) - P(N \geq n+1 | D = d) = 0$, επειδή για $n > 1$ και τα δύο όρια είναι ίσα με 0.

Συνεπώς, αποδείχθηκε ότι $\lim_{d \rightarrow \infty} P(N = 1 | D = d) = 1$.

□

Σε αυτό το σημείο, αξίζει να εξετασθούν τα συμπεράσματα του Θεωρήματος 1 για τη συμπεριφορά του συστήματος. Η βασική εικόνα που αποτυπώνεται είναι ότι, όσο μεγαλύτερο συνολικό μήκος ουράς παρατηρεί ένας νέος πελάτης κατά την άφιξή του, τόσο περισσότερος χρόνος εκτιμάται ότι έχει παρέλθει από την τελευταία είσοδο πραγματικού πελάτη. Δίνεται η εντύπωση, λοιπόν, ότι όταν το d αυξάνεται, ο αριθμός των πελατών που δεν έχουν εγκαταλείψει το σύστημα μειώνεται, με συνέπεια οι πελάτες που προηγούνται να παρουσιάζουν αυξημένη πιθανότητα να είναι εικονικοί.

Έτσι, είδαμε ότι για μεγάλες τιμές του συνολικού πλήθους πελατών ($d \geq 2K - 1$), η κατανομή του αριθμού των πραγματικών πελατών μειώνεται σταδιακά, περιορίζοντας το πλήθος αυτών στο σύστημα. Η συμπεριφορά αυτή ενισχύεται περαιτέρω καθώς το d τείνει στο άπειρο, όπου η πιθανότητα να υπάρχει μονάχα ένας πελάτης φτάνει στην μονάδα. Δηλαδή, εάν ο παρατηρούμενος αριθμός πελατών γίνει τεράστιος, με βεβαιότητα παραμένει ενεργός αποκλειστικά ο πελάτης που εξυπηρετείται τη δεδομένη στιγμή, ενώ όλοι οι υπόλοιποι έχουν αποχωρήσει.

Πριν προχωρήσουμε στην διατύπωση και απόδειξη του επόμενου θεωρήματος, κρίνεται αναγκαίο να οριστούν ορισμένα μεγέθη. Έστω V το όφελος από την εξυπηρέτηση. Έστω $C(t)$ η συνάρτηση κόστους αναμονής (waiting cost function), όπου t είναι ο χρόνος αναμονής του πελάτη. Υποθέτουμε ότι η $C(t)$ είναι μία γνησίως αύξουσα συνάρτηση του t , που σημαίνει ότι όσο περισσότερο περιμένει κάποιος, τόσο μεγαλύτερο είναι το κόστος για αυτόν. Έπειτα, ορίζουμε $U(t)$ την συνάρτηση χρησιμότητας (utility function) για έναν πελάτη που παραμένει χρόνο t στην ουρά πριν εξυπηρετηθεί και η οποία υπολογίζεται από τον τύπο $U(t) = V - C(t)$. Αν τώρα θεωρήσουμε με $E[W | D = d]$ τον αναμενόμενο χρόνο αναμονής του πελάτη, δεδομένου ότι ο συνολικός αριθμός

που παρατηρεί κατά την άφιξή του είναι d , τότε η αναμενόμενη χρησιμότητα (expected utility) για αυτόν τον πελάτη δίνεται από τον τύπο $E[U] = V - C(E[W | D = d])$.

Θεώρημα 4.2: Θεωρούμε σύστημα ουράς εισιτηρίων M/M/1 με ομοιογενείς πελάτες ως προς την αξία εξυπηρέτησης και το κόστος αναμονής. Υποθέτουμε ότι η αξία εξυπηρέτησης είναι αρκετά μεγάλη, ώστε για χαμηλές τιμές του d να ισχύει $V - C(E[W | D = d]) \geq 0$. Κάτω από αυτές τις συνθήκες, για οποιαδήποτε συνάρτηση κόστους αναμονής $C(t)$, ισχύει ότι μία στρατηγική μονού κατωφλιού δεν μπορεί να αποτελεί στρατηγική ισορροπίας Nash. Επιπλέον, εάν όλοι οι πελάτες υιοθετήσουν την ίδια στρατηγική κατωφλιού, τότε η προτιμότερη αντίδραση για έναν μεμονωμένο πελάτη είναι να ακολουθήσει μία στρατηγική διπλού κατωφλιού (K_1, K_2) , με $K_1 < K_2$, όπου ο κανόνας συμμετοχής ορίζει ότι κάθε πελάτης επιλέγει να εισέλθει στην ουρά αν και μόνον αν $D < K_1$ ή $D > K_2$.

Απόδειξη: Υποθέτουμε ότι όλοι οι πελάτες υιοθετούν κάποιο κατώφλι K και εξετάζουμε τι θα συμβεί στις διαφορετικές περιπτώσεις τιμών της παρατηρούμενης ποσότητας D . Αν ο πελάτης φτάνοντας δει μία σχετικά μικρή ουρά αναμονής, ο χρόνος αναμονής είναι μικρός, άρα η αναμενόμενη χρησιμότητα θα είναι θετική ($V - C(E[W | D = d]) \geq 0$) και κατά συνέπεια τον συμφέρει να μείνει για εξυπηρέτηση. Για ενδιάμεσες τιμές της παρατηρούμενης ουράς, όπου το D έχει ξεπεράσει το αρχικό κατώφλι, αλλά δεν βρίσκεται ακόμα σε ακραία επίπεδα, ο αναμενόμενος χρόνος αναμονής αυξάνεται τόσο ώστε η χρησιμότητα να γίνεται αρνητική, αναγκάζοντας τον πελάτη να αποχωρήσει. Από την άλλη πλευρά, εάν ο πελάτης φτάνοντας δει εξαιρετικά μεγάλη ουρά αναμονής, τότε σύμφωνα με το Θεώρημα 1, εκτιμά ότι οι περισσότεροι από τους πελάτες έχουν αποχωρήσει. Επομένως, η πραγματική ουρά είναι τελικά μικρή, ο αναμενόμενος χρόνος αναμονής όχι απαγορευτικά μεγάλος και η αναμενόμενη χρησιμότητα παίρνει θετικές τιμές. Οπότε και σε αυτή την περίπτωση ο πελάτης είναι προτιμότερο να μείνει στην ουρά. Συνεπώς, ολοκληρώνοντας την απόδειξη, καταλήξαμε ότι η βέλτιστη ατομική απόφαση δεν οδηγεί σε ένα απλό κατώφλι, αλλά στην καθιέρωση μίας στρατηγικής διπλού κατωφλιού.

□

4.5 Ανασκόπηση, Συμπεράσματα και Προοπτικές Γενίκευσης

Στο παρόν κεφάλαιο μελετήθηκε η συμπεριφορά ενός συστήματος ουράς εισιτηρίων M/M/1, προχωρώντας στον υπολογισμό της κατανομής του χρόνου αναμονής και της βέλτιστης απόκρισης ενός εισερχόμενου πελάτη. Βασικός στόχος ήταν να γίνει κατανοητό πώς η πληροφορία που παρατηρεί ο αφιχθείς πελάτης επηρεάζει την εκτίμησή του για την πραγματική κατάσταση της ουράς. Για τον λόγο αυτό, ορίστηκε το σύνολο S_d , όλων εκείνων των καταστάσεων όπου το

συνολικό πλήθος ισούται με d ($d \geq K$), καθώς και ο αριθμός των πραγματικών πελατών A_d , κατά τη διάρκεια μίας περιόδου, όταν το σύστημα βρίσκεται σε μία κατάσταση του S_d .

Προκειμένου να αναδειχθεί η συμπεριφορά του συστήματος, παρουσιάστηκαν τρία διαδοχικά λήμματα, τα οποία λειτούργησαν ως απαραίτητη μαθηματική προετοιμασία για την απόδειξη δύο βασικών θεωρημάτων.

Αρχικά, μέσω του Λήμματος 1, καθορίστηκε το πλήθος των καταστάσεων του S_d , το οποίο περιέχει ακριβώς 2^{K-1} καταστάσεις, καθεμία από τις οποίες αντιστοιχεί σε διαφορετική φάση. Παρότι το λήμμα αυτό δεν αξιοποιήθηκε ως άμεσος υπολογιστικός τύπος στις αλγεβρικές σχέσεις των επόμενων βημάτων, εντούτοις, λειτούργησε ως περιγραφικός ορισμός για την δομή του χώρου καταστάσεων, εντοπίζοντας τη μικρότερη τιμή του πλήθους και τη φάση στην οποία αντιστοιχεί.

Εν συνεχεία, το Λήμμα 2 απέδειξε ότι κατά τη διάρκεια μίας περιόδου απασχόλησης, ο αριθμός των πραγματικών πελατών στο σύστημα μειώνεται σταδιακά. Το συμπέρασμα αυτό αξιοποιήθηκε άμεσα στην απόδειξη του επόμενου λήμματος, δικαιολογώντας ότι το σύστημα δεν δέχεται νέους ενεργούς πελάτες, εφόσον έχει ξεπεραστεί το καθορισμένο όριο και πραγματοποιούνται αποκλειστικά ολοκληρώσεις εξυπηρετήσεων.

Αμέσως μετά, το Λήμμα 3 αξιοποίησε τα συμπεράσματα του προηγούμενου σταδίου για να αποδείξει μία σχέση στοχαστικής διάταξης ως προς το πλήθος των πραγματικών πελατών. Έτσι, μέσω της ανισοτικής σχέσης $A_{d+1} \leq_{st} A_d$, είδαμε ότι καθώς το d αυξάνεται, ο αριθμός των πραγματικών πελατών έχει την τάση να μειώνεται. Το εν λόγω λήμμα αποτέλεσε βασικό εργαλείο για την απόδειξη του Θεωρήματος 1, τεκμηριώνοντας τη φθίνουσα συμπεριφορά της ουράς.

Έπειτα, πατώντας σε όλα τα παραπάνω εξαγόμενα, το Θεώρημα 1 απέδειξε ότι η δεσμευμένη κατανομή του πλήθους των πραγματικών πελατών N , δεδομένου του ορατού μήκους $D = d$, είναι στοχαστικά φθίνουσα ως προς το d . Παράλληλα, αποδείχθηκε ότι καθώς το d τείνει στο άπειρο, η πιθανότητα ύπαρξης ακριβώς ενός πραγματικού πελάτη στην ουρά τείνει στη μονάδα, οδηγώντας στο συμπέρασμα ότι μια τεράστια ουρά αποτελείται ουσιαστικά από άτομα που έχουν εγκαταλείψει το σύστημα.

Τέλος, το Θεώρημα 2 αξιοποίησε όλη την προηγούμενη αλυσίδα λημμάτων μέσα από τις πληροφορίες που παρείχε το Θεώρημα 1 και συνέθεσε το κεντρικό ζήτημα του παρόντος κεφαλαίου. Κατέληξε στο συμπέρασμα ότι το αναμενόμενο όφελος ενός πελάτη γίνεται θετικό, όταν αντικρίζει μία πολύ μεγάλη ουρά, διότι εκτιμά ότι ο χρόνος αναμονής θα είναι σχετικά μικρός, εξαιτίας των πολλών αποχωρήσεων. Επί της ουσίας, έδειξε ότι εάν σε μία ουρά εισιτηρίων όλοι υιοθετήσουν την στρατηγική μονού κατωφλιού και παραμείνουν στην αναμονή για εξυπηρέτηση μονάχα όταν η θέση τους δεν ξεπερνά ένα προκαθορισμένο όριο, τότε η βέλτιστη απόκριση ενός ατόμου είναι να ακολουθήσει τη στρατηγική διπλού κατωφλιού και να εισέλθει στην περίπτωση που η θέση του ξεπερνά έναν άλλο προκαθορισμένο δείκτη ορίου, ικανοποιητικά ανώτερο από αυτόν του μονού κατωφλιού.

Θέλοντας να διευρύνουμε τα τελικά συμπεράσματα του παρόντος κεφαλαίου, θα λέγαμε ότι παρόλο που στην πράξη οι πελάτες υιοθετούν απλές στρατηγικές κατωφλιού, αποδεικνύεται ότι τέτοιες στρατηγικές δεν μπορούν να αποτελέσουν στρατηγικές ισορροπίας Nash. Κατά την αναζήτηση μίας πιο αποτελεσματικής προσέγγισης, διαπιστώνεται ότι πιο αποδοτικό για ένα μεμονωμένο άτομο - όταν όλοι οι υπόλοιποι ακολουθούν στρατηγική κατωφλιού- είναι να επιλέξει να αποκλίνει, ακολουθώντας μια στρατηγική διπλού κατωφλιού (K_1, K_2) . Δηλαδή, συμφέρει τον πελάτη να παραμείνει στην ουρά είτε όταν το παρατηρούμενο πλήθος είναι χαμηλό ($D < K_1$), είτε όταν είναι πάρα πολύ μεγάλο ($D > K_2$), γνωρίζοντας ότι στην τελευταία περίπτωση οι περισσότεροι προηγούμενοι πελάτες έχουν αποχωρήσει και η πραγματική αναμονή θα είναι μικρή.

Ωστόσο, ούτε η στρατηγική διπλού κατωφλιού μπορεί να θεμελιωθεί ως ισορροπία Nash, αφού εάν όλοι υιοθετήσουν μία στρατηγική διπλού κατωφλιού, τότε κάθε πελάτης που παρατηρεί μία εξαιρετικά μεγάλη ουρά εκτιμά ότι αυτή αποτελείται από ενεργούς πελάτες, με συνέπεια ο χρόνος αναμονής του να είναι τελικά πολύ μεγάλος και, έτσι, να προτιμά να αποχωρήσει.

Προχωρώντας πέρα από το βασικό μοντέλο, εξετάζεται τι συμβαίνει αν αλλάξουν ορισμένες από τις υποθέσεις. Το πρώτο ζήτημα αφορά τη συμπεριφορά των πελατών κατά την άφιξή τους. Στο μοντέλο που μελετήθηκε, έγινε η υπόθεση ότι η επιλογή της εισόδου ενός πελάτη είναι δυνατή μονάχα κατά την άφιξή του. Όμως, υπάρχει ενδιαφέρον να μελετηθεί η περίπτωση όπου οι πελάτες έχουν τη δυνατότητα να εισέλθουν στο σύστημα και να επανεξετάσουν την παραμονή τους σε αυτό σε μεταγενέστερο χρόνο. Έτσι, μπορούν να ενταχθούν στην ουρά και ενδεχομένως να διαμορφώσουν μια άλλη άποψη, εάν το επιθυμούν, εγκαταλείποντας τελικά το σύστημα. Βέβαια, ακόμα και εάν προστεθεί αυτή η δυνατότητα, το βασικό συμπέρασμα δεν αλλάζει, αφού και σε αυτή την περίπτωση μία απλή στρατηγική κατωφλιού εξακολουθεί να μην αποτελεί ισορροπία Nash. Παράλληλα, τίθεται το ερώτημα κατά πόσο η στρατηγική διπλού κατωφλιού αποτελεί την αποδοτικότερη επιλογή από την πλευρά του πελάτη. Η απάντηση εξαρτάται από το πώς διαμορφώνεται το κόστος αναμονής με την πάροδο του χρόνου. Εξαιτίας του μαρκοβιανού χαρακτήρα του συστήματος, όσο περισσότερο περιμένει κάποιος, τόσο λιγότερος χρόνος τού απομένει για να εξυπηρετηθεί. Οπότε, εάν το κόστος αναμονής είναι κοίλο, όποιος μπει στην ουρά δεν φεύγει ποτέ, αφού όσο περνάει η ώρα η αναμονή φαντάζει όλο και μικρότερη, κάνοντας τον πελάτη να πιστεύει ότι πλησιάζει η στιγμή της εξυπηρέτησής του. Όμως, εάν το κόστος είναι κυρτό ως προς τον χρόνο, η καθυστέρηση γίνεται όλο και πιο ανυπόφορη, κάνοντας τον πελάτη να θέλει κάποια στιγμή να αποχωρήσει. Συνεπώς, προκύπτει ότι η αποδοτικότητα και η συμπεριφορά των πελατών εξαρτώνται άμεσα από τη δομή του κόστους αναμονής.

Στο πλαίσιο γενίκευσης του μοντέλου, πέρα από το βασικό σύστημα με έναν μόνο εξυπηρετητή, έχει σημασία να εξετασθεί πώς συμπεριφέρεται το σύστημα όταν υπάρχουν περισσότεροι από ένας διαθέσιμοι εξυπηρετητές, δηλαδή σε ένα σύστημα τύπου M/M/c. Σε ένα τέτοιο περιβάλλον, η επιλογή του πελάτη επηρεάζεται από τη συνολική ικανότητα εξυπηρέτησης. Και σε αυτό το σενάριο τα συμπεράσματα της μελέτης εξακολουθούν να ισχύουν, διότι αν υπάρχει τουλάχιστον ένας ελεύθερος εξυπηρετητής, οι πελάτες επιλέγουν να εισέλθουν για πραγματοποιήσιμες τιμές των παραμέτρων. Αντίθετα, αν όλοι οι εξυπηρετητές είναι απασχολημένοι, η συνολική συμπεριφορά

επηρεάζεται από τον αριθμό των εξυπηρετήσεων αποκλειστικά μέσω του συνολικού ρυθμού εξυπηρέτησης.

Μία άλλη επέκταση, πολύ συνηθισμένη στα πραγματικά συστήματα, προκύπτει όταν η επιβεβαίωση της αποχώρησης ενός πελάτη, που αποφασίζει να μην μείνει στην ουρά, απαιτεί ένα θετικό, έστω και μικρό, χρονικό διάστημα. Στη μοντελοποίηση, αυτό εκφράζεται θεωρώντας ότι ο πελάτης που αποχωρεί αντιμετωπίζεται σαν να εξυπηρετείται πάρα πολύ γρήγορα, με έναν ιδιαίτερα υψηλό ρυθμό θ , γεγονός που σημαίνει ότι ο χρόνος που χρειάζεται για να ολοκληρωθεί η αποχώρησή του είναι ελάχιστος. Συγκεκριμένα, σε αυτό το γενικευμένο πλαίσιο περικλείεται, ως ειδική περίπτωση, το μοντέλο που μελετήθηκε στο παρόν κεφάλαιο, όπου αντιστοιχούσε σε οριακή περίπτωση ρυθμού $\theta = \infty$, κάνοντας τον χρόνο καθυστέρησης μηδενικό. Όμως, όταν το θ είναι πεπερασμένο, δηλαδή υπάρχει αυτή η μικρή καθυστέρηση στην ενημέρωση του συστήματος για την αποχώρηση του πελάτη, τα προηγούμενα συμπεράσματα παύουν να ισχύουν. Αυτό συμβαίνει διότι αν ένας πελάτης αντικρίσει μία πολύ μεγάλη ουρά, θα χρειαστεί να περιμένει σημαντικό χρόνο για να εξυπηρετηθεί, αν αποφασίσει να παραμείνει στο σύστημα. Συνεπώς, κάτω από αυτές τις συνθήκες, δεν αποκλείεται μία στρατηγική κατωφλιού να αποτελεί συμμετρική στρατηγική ισορροπία Nash. Επομένως, η ακριβής δομή της ισορροπίας Nash σε αυτό το μοντέλο παραμένει ένα ανοιχτό ζήτημα προς μελέτη.

Μία τελευταία προοπτική γενίκευσης είναι να ξεφύγουμε από την παραδοχή των εκθετικών χρόνων εξυπηρέτησης. Η εξαγωγή ακριβών ποσοτικών αποτελεσμάτων γίνεται εξαιρετικά δύσκολη, αν όχι ακατόρθωτη, καθώς εξαρτάται άμεσα από τη μορφή που έχει η κατανομή του χρόνου εξυπηρέτησης. Ως προς τα ποιοτικά συμπεράσματα, εκτιμάται ότι η κατανομή έχει την ιδιότητα να ευνοεί την ολοκλήρωση της εξυπηρέτησης όσο περνάει ο χρόνος. Τα βασικά μας συμπεράσματα θα παραμείνουν έγκυρα, επειδή μια μεγάλη ουρά εξακολουθεί να δείχνει σύντομο χρόνο αναμονής. Γενικότερα, τα συστήματα ουράς εισιτηρίων με γενικούς χρόνους εξυπηρέτησης είναι πολύ δύσκολο να μελετηθούν, τουλάχιστον με τρόπο που να δίνει ποιοτικά αποτελέσματα.

Βιβλιογραφία/Παραπομπές

- [[1]] Adan, I., van Leeuwen, J., & Selen, J. (2017). *Analysis of structured Markov processes*, arXiv preprint arXiv:1709.09060.
- [[2]] Harchol-Balter, M. (2013). *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*. Cambridge University Press.
- [[3]] Hassin, R., & Haviv, M. (2006). *To Queue Or Not To Queue: Equilibrium Behavior In Queueing Systems*. Kluwer Academic Publishers.
- [[4]] Hui, M. K., & Tse, D. K. (1996). *What to Tell Customers in Waits of Different Lengths: An Integrative Model of Service Evaluation*. *Journal of Marketing*, 60(2), 81-90.
- [[5]] Kerner, Y., Sherzer, E., & Yanco, M. A. (2017). *On non-equilibria threshold strategies in ticket queues*. *Queueing Systems*, 86(4), 419-431.
- [[6]] Kulkarni, V. G. (2010). *Introduction to Modeling and Analysis of Stochastic Systems*. Springer.
- [[7]] Larson, R.C. (1987). *Perspectives on Queues: Social Justice and the Psychology of Queueing*. *Operations Research*, 35(6), 895-905.
- [[8]] Λουλάκης, Μ.(2019). *Στοχαστικές Διαδικασίες*. Εκδόσεις Κάλλιπος, Αθήνα.
- [[9]] Neuts, M. F. (1994). *Matrix Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Dover Publications, Mineola, NY.
- [[10]] Osborne, M. J. (2000). *An Introduction to Game Theory*. Oxford University Press.
- [[11]] Ramaswami, V., & Latouche, G. (1999). *Introduction to Matrix Analytic Methods in Stochastic Modeling*. SIAM.
- Ramaswami, V., & Latouche, G. (1986). *A general class of Markov processes with explicit matrix-geometric solutions*. *Operations-Research-Spectrum*, 8(4), 209-218.
- [[12]] Ross, S. M. (2014). *Introduction to Probability Models*. Academic Press.
- [[13]] Xu, S. H., Gao, L., & Ou, J. (2007). *Service Performance Analysis and Improvement for a Ticket Queue with Balking Customers*. *Management Science*, 53(6), 971-990.

Παράρτημα Α: Κώδικας Αριθμητικής Εφαρμογής της MGM στο M/Cox/1

Το Παράρτημα αυτό περιλαμβάνει τον πλήρη κώδικα που χρησιμοποιήθηκε στο μοντέλο M/Cox/1 της ενότητας 1.4.3.

```
In[1]:= ClearAll["Global`*"]

(* Ορισμός Παραμέτρων *)
λ = 1.0;
μ1 = 2.0;
μ2 = 3.0;
ρ = 0.4;
q = 1.0 - ρ;

(* Ορισμός Πινάκων *)
B = {{q * μ1, 0}, {μ2, 0}};
L = {{-(λ + μ1), ρ * μ1}, {0, -(λ + μ2)}};
F = {{λ, 0}, {0, λ}};
B0 = {{q * μ1}, {μ2}};
L0 = {{-λ}};
F0 = {{λ, 0}};
Linverse = Inverse[L];

(* Υπολογισμός του Rate Matrix R *)
Rcurrent = {{0.0, 0.0}, {0.0, 0.0}}; (*Θέτουμε Αρχική Εκτίμηση*)
Do[Rnext = -(F + B . Rcurrent . Rcurrent) . Linverse;
  Rcurrent = Rnext;
  If[Max[Abs[Rnext - Rcurrent]] < 10^-7, Break[]];, {i, 1, 100}];
R = Rcurrent;

(* Ορίζουμε Βοηθητικούς Πίνακες *)
Matrix1 = L + R . B;
Matrix2 = IdentityMatrix[2] - R;
Matrix3 = Inverse[Matrix2];
```

```
(* Εξίσωση 1: Επίπεδο 0 *)
Eq1 = (x0 * L0[[1, 1]] + ({x1, x2} . B0)[[1, 1]]) == 0;

(* Εξίσωση 2: Επίπεδο 1 (μία εκ των δύο εξισώσεων) *)
Eq2 = (x0 * F0[[1, 1]] + ({x1, x2} . Matrix1)[[1, 1]]) == 0;

(* Εξίσωση 3: Συνθήκη Κανονικοποίησης *)
e = {1}, {1};
Eq3 = x0 + ({x1, x2} . Matrix3 . e)[[1, 1]] == 1;

(* Εύρεση Οριακών Πιθανοτήτων *)
Solution = Solve[{Eq1, Eq2, Eq3}, {x0, x1, x2}];
π0 = N[x0 /. Solution[[1]];
π1 = N[{x1, x2} /. Solution[[1]];

(* Παρουσίαση Αποτελεσμάτων *)
Print[MatrixForm[R]];
Print["π0 = ", π0];
Print["π1 = ", π1];

R = ( 0.333333  0.0666667
      0.      0.25 )
π0 = 0.511848
π1 = {0.236967, 0.0758294}
```

Κατάλογος και Επεξήγηση Όρων

Ξενόγλωσσος Όρος	Ελληνικός Όρος	Επεξήγηση
Absolute Percentage Error-APE	Απόλυτο Ποσοστιαίο Σφάλμα	Δείκτης που ποσοτικοποιεί τη σχετική απόκλιση επί τοις εκατό μεταξύ της προσεγγιστικής και της ακριβούς τιμής.
Avoid the Crowd-ATC	Αποφυγή του Πλήθους	Περίπτωση όπου η βέλτιστη απόκριση είναι φθίνουσα.
Balk	Αποχώρηση	Ενέργεια κατά την οποία ο πελάτης εγκαταλείπει το σύστημα λόγω υπερβολικού φόρτου που παρατηρεί κατά την άφιξή του.
Balking Customers	Πελάτες που Αποχωρούν	Πελάτες οι οποίοι φτάνουν στο σύστημα αναμονής, αλλά αποφασίζουν να μην παραμείνουν σε αυτό επειδή η ουρά είναι πολύ μεγάλη, δηλαδή ο χρόνος αναμονής κρίνεται υπερβολικός για τους ίδιους.
Best Response	Βέλτιστη Απόκριση	Η στρατηγική ενός παίκτη (πελάτη) που μεγιστοποιεί την απόδοσή του, δεδομένων των στρατηγικών που έχουν επιλέξει οι υπόλοιποι.
Birth-Death Processes	Διαδικασίες Γεννήσεων-Θανάτων	Μαρκοβιανές Αλυσίδες Συνεχούς Χρόνου (ΜΑΣΧ) με μεταβάσεις μόνο μεταξύ γειτονικών καταστάσεων.
Boundary Portion (Level 0)	Επίπεδο Ορίου	Το αρχικό τμήμα μίας QBD διαδικασίας που μπορεί να έχει διαφορετική διάσταση από τα επόμενα επίπεδα.
Continuous Time Markov Chain (CTMC)	Μαρκοβιανή Αλυσίδα Συνεχούς Χρόνου (ΜΑΣΧ)	Στοχαστική διαδικασία με διακριτό χώρο καταστάσεων και συνεχή χρόνο.
Discrete Time Markov Chain (DTMC)	Μαρκοβιανή Αλυσίδα Διακριτού Χρόνου (ΜΑΔΧ)	Στοχαστική διαδικασία με διακριτό χώρο καταστάσεων και διακριτό χρόνο.
Drift	Μετατόπιση	Η τάση μίας στοχαστικής διαδικασίας να κινείται μακροπρόθεσμα είτε προς τα δεξιά (mean drift to the right), είτε προς τα αριστερά (mean drift to the left).
Drift Off	Ολίσθηση στο άπειρο	Η κατάσταση όπου η στοχαστική διαδικασία απομακρύνεται απεριόριστα προς τα δεξιά, οδηγώντας σε αστάθεια το σύστημα.
Expected Utility	Αναμενόμενη Χρησιμότητα	Το μέγεθος που προκύπτει ως διαφορά μεταξύ του οφέλους εξυπηρέτησης και του κόστους αναμονής.
Follow the Crowd-FTC	Προσχώρηση στο Πλήθος	Περίπτωση όπου η βέλτιστη απόκριση είναι αύξουσα.
GI/M/1 Queueing System	Σύστημα Ουράς GI/M/1	Μοντέλο ουράς με γενική κατανομή αφίξεων, εκθετικούς χρόνους εξυπηρέτησης και έναν εξυπηρετητή.
Infinitesimal Generator	Απειροστός Γεννήτορας (Q)	Πίνακας ρυθμών μετάβασης μιας ΜΑΣΧ που καθορίζει την συμπεριφορά της.
Join	Συμμετοχή	Ενέργεια κατά την οποία ο πελάτης αποφασίζει να εισέλθει/παραμείνει στο σύστημα για εξυπηρέτηση.

Level	Επίπεδο	Η κύρια μεταβλητή μιας QBD διαδικασίας, η οποία εκφράζει το συνολικό πλήθος των πελατών στο σύστημα και μεταβάλλεται μόνο κατά ένα βήμα τη φορά.
Matrix Geometric Method	Πινακογεωμετρική Μέθοδος	Η μέθοδος υπολογισμού της στάσιμης κατανομής σε διαδικασίες QBD.
Mixed Strategy	Μικτή Στρατηγική	Ένας συνδυασμός καθαρών στρατηγικών, όπου ο πελάτης επιλέγει ποια θα ακολουθήσει βάσει μίας κατανομής πιθανότητας (με ορισμένα ποσοστά).
Maximum Absolute Difference - MAD	Μέγιστη Απόλυτη Διαφορά	Στατιστικός δείκτης που μετρά τη μέγιστη απόλυτη απόκλιση μεταξύ των αποτελεσμάτων ενός μεγέθους δύο συγκρινόμενων μοντέλων.
Modified Ticket Queue	Τροποποιημένη Ουρά Εισιτηρίων	Αναδιάρθρωση του συστήματος ουράς σε δύο διακριτές υποομάδες πελατών (πραγματικών, εικονικών), με στόχο την αντιμετώπιση του προβλήματος της υπολογιστικής πολυπλοκότητας.
Nash Equilibrium	Ισορροπία Nash	Κατάσταση όπου κανένας πελάτης δεν έχει κίνητρο να μεταβάλλει την στρατηγική του, με δεδομένο ότι όλοι οι άλλοι κρατάνε την δική τους σταθερή.
Net Drift	Καθαρή Μετατόπιση	Η συνολική τάση κίνησης της διαδικασίας QBD, η οποία αποτελεί καθοριστικό παράγοντα για την αξιολόγηση της ευστάθειας.
Neuts' Mean Drift Condition	Συνθήκη Μέσης Ολίσθησης του Neuts	Συνθήκη ευστάθειας για τις QBD διαδικασίες, η οποία απαιτεί η τάση μετατόπισης προς αριστερά να υπερτερεί της αντίστοιχης προς δεξιά.
Non-degenerate	Μη εκφυλισμένα	Περίπτωση όπου τα σημεία άλματος της βέλτιστης απόκρισης δεν συμπίπτουν με ακέραιες τιμές.
Normalization Condition	Συνθήκη Κανονικοποίησης	Συνθήκη όπου το άθροισμα όλων των πιθανοτήτων ισούται ακριβώς με 1, επιβάλλοντας στα στοιχεία κάθε γραμμής του πίνακα μεταβάσεων να αθροίζουν στην μονάδα.
Phase	Φάση	Η δευτερεύουσα μεταβλητή μιας QBD διαδικασίας, η οποία εκφράζει μία εσωτερική λειτουργία του συστήματος, π.χ. αν ο εξυπηρετητής είναι διαθέσιμος ή απασχολημένος.
Pure Strategy	Καθαρή Στρατηγική	Μια στρατηγική όπου ο πελάτης επιλέγει με βεβαιότητα μία συγκεκριμένη ενέργεια (π.χ. να παραμείνει ή να εγκαταλείψει την ουρά).
Quasi Birth Death Processes	QBD Διαδικασίες	Μια Διδιάστατη Μαρκοβιανή Αλυσίδα (με επίπεδο και φάση) για τη μελέτη ουρών αναμονής.
Queueing Position	Θέση στην Ουρά	Η ακριβής θέση που καταλαμβάνει ένας πελάτης σε ένα σύστημα αναμονής, η οποία καθορίζεται από το συνολικό πλήθος των πελατών..
Rate Matrix	Πίνακας Ρυθμού	Ο πίνακας που καθορίζει τους ρυθμούς μετάβασης μεταξύ διαφορετικών καταστάσεων σε μία στοχαστική διαδικασία, στο πλαίσιο της πινακογεωμετρικής μεθόδου

Rational Choice Theory	Θεωρία Ορθολογικής Επιλογής	Το πλαίσιο σύμφωνα με το οποίο το άτομο επιλέγει πάντα την καλύτερη δυνατή ενέργεια βάσει των προτιμήσεών του.
Rationality	Ορθολογικότητα	Η ιδιότητα ενός ατόμου να είναι συνεπής στις αποφάσεις του, όταν αλλάζουν κάθε φορά οι διαθέσιμες επιλογές του.
Repetitive Portion	Επαναλαμβανόμενο Τμήμα	Τα λοιπά διαδοχικά επίπεδα, πέραν του επιπέδου 0, σε μία διαδικασία QBD, στα οποία η δομή των μεταβάσεων παραμένει η ίδια, επιτρέποντας τη επίλυση του συστήματος με πινακογεωμετρική μέθοδο.
Stability Condition	Συνθήκη Ευστάθειας	Η προϋπόθεση που πρέπει να πληρείται ώστε ένα σύστημα να μην οδηγείται σε υπερβολική συσσώρευση και να παραμένει σε ισορροπία σε βάθος χρόνου.
Threshold Strategies	Στρατηγικές Κατωφλιού	Στρατηγικές όπου η απόφαση (π.χ. παραμονής ή μη ενός πελάτη) βασίζεται στη σύγκριση της κατάστασης με ένα καθορισμένο όριο (κατώφλι).
Ticket Plus Queue	Τροποποιημένη Ουρά Εισιτηρίων	Επέκταση του απλού συστήματος εξυπηρέτησης, όπου τυπώνεται ο αναμενόμενος χρόνος αναμονής πάνω στο εισιτήριο, με στόχο τη διόρθωση της μεροληπτικής εκτίμησης του πελάτη.
Ticket Position	Θέση Εισιτηρίου	Η διαφορά μεταξύ του αριθμού εισιτηρίου ενός πελάτη και της τρέχουσας ένδειξης του πίνακα εξυπηρέτησης.
Ticket Queue	Ουρά Εισιτηρίων	Σύστημα όπου εκδίδονται εισιτήρια και οι πελάτες εξυπηρετούνται σύμφωνα με μια καθορισμένη προτεραιότητα.
Total Absolute Difference - TAD	Συνολική Απόλυτη Διαφορά	Στατιστικός δείκτης που εκφράζει το άθροισμα των απόλυτων διαφορών σε όλο το φάσμα των εξεταζόμενων τιμών.
Traffic Intensity	Ένταση Κίνησης	Ο συνολικός φόρτος εργασίας που φτάνει στο σύστημα, πριν συνυπολογιστούν οι αποχωρήσεις πελατών.
Utility Function	Συνάρτηση Χρησιμότητας	Συνάρτηση η οποία εκφράζει το όφελος ενός πελάτη από μία επιλογή.
Utilization Factor	Συντελεστής Αξιοποίησης	Ο πραγματικός φόρτος εργασίας του εξυπηρετητή, λαμβάνοντας υπόψη μόνο τους πελάτες που τελικά παραμένουν στην ουρά εξυπηρέτησης.
Waiting Cost Function	Συνάρτηση Κόστους Αναμονής	Συνάρτηση που εκφράζει το κόστος για τον πελάτη όσο βρίσκεται στην αναμονή.

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν.1599/1986, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης.