



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής

Εξισορρόπηση Φορτίου σε Δορυφορικά Δίκτυα με Χρήση Θεωρίας Παιγνίων και Μετανοητικής Μάθησης

Διπλωματική Εργασία

του

ΧΡΗΣΤΟΥ ΑΝΑΣΤΑΣΙΟΥ ΜΠΟΒΟΛΗ

Επιβλέπων: Συμεών Παπαβασιλείου
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2025



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής

Εξισορρόπηση Φορτίου σε Δορυφορικά Δίκτυα με Χρήση Θεωρίας Παιγνίων και Μετανοητικής Μάθησης

Διπλωματική Εργασία

του

ΧΡΗΣΤΟΥ ΑΝΑΣΤΑΣΙΟΥ ΜΠΟΒΟΛΗ

Επιβλέπων: Συμεών Παπαβασιλείου
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 30η Ιουνίου 2025.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....

Συμεών Παπαβασιλείου
Καθηγητής Ε.Μ.Π.

.....

Ιωάννα Ρουσάκη
Αναπληρώτρια Καθηγήτρια Ε.Μ.Π.

.....

Ελένη Στάη
Επίκουρη Καθηγήτρια Ε.Μ.Π.

Αθήνα, Ιούνιος 2025



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής

(Υπογραφή)

.....
Χρήστος Α. Μποβολής

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Χρήστος Α. Μποβολής, 2025.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Η παρούσα διπλωματική εργασία διερευνά το πρόβλημα της εξισορρόπησης φορτίου σε Δορυφορικές Υποδομές Υπολογιστικής Παρυφής (Satellite Edge Computing - SEC). Στο πλαίσιο αυτό, επίγειοι χρήστες του Διαδικτύου των Πραγμάτων (Internet of Things - IoT) αναθέτουν υπολογιστικές διεργασίες σε δορυφορικούς εξυπηρετητές για καταναεμημένη επεξεργασία. Η ετερογένεια των διεργασιών αυτών και η στοχαστικότητα της άφιξής τους οδηγεί σε ανισοκατανομή του φορτίου μεταξύ των δορυφόρων. Στόχος είναι η ανάπτυξη μηχανισμών βέλτιστης εξισορρόπησης, με σκοπό τη χρονική και ενεργειακή αποδοτικότητα, αλλά και τη δικαιοσύνη (fairness).

Η προτεινόμενη στρατηγική ακολουθεί μια επαναληπτική διαδικασία που περιλαμβάνει: (i) επιλογή κατάλληλου δορυφόρου-στόχου για κάθε δορυφόρο-πηγή σε κάθε γύρο, και (ii) προσδιορισμό της βέλτιστης ποσότητας φορτίου προς εκφόρτωση. Η τελευταία βασίζεται σε ένα ντετερμινιστικό κανόνα, ενώ η επιλογή στόχου μοντελοποιείται με παιγνιο-θεωρητικούς μηχανισμούς που λαμβάνουν υπόψη πραγματικούς τηλεπικοινωνιακούς περιορισμούς. Οι μηχανισμοί αυτοί διακρίνονται σε στοχαστικά παίγνια (stochastic games), καθώς και θορυβώδη παίγνια δυναμικού (potential games) σε μη ανατροφοδοτούμενα (bandit) ή ημιανατροφοδοτούμενα (semi-bandit) περιβάλλοντα.

Για την επίλυση αυτών των παιγνίων αξιοποιούνται τεχνικές της Μετανοητικής Μάθησης (No-Regret Learning), με κύριο εργαλείο την παραλλαγή ϵ -Hedge του Αλγορίθμου Εκθετικών Βαρών (Exponential Weights Algorithm - EWA), η οποία προσφέρει εγγυήσεις σύγκλισης σε δ-ισορροπίες και ισορροπίες Nash για τα μοντέλα που βασίζονται σε παίγνια δυναμικού. Παρουσιάζονται επίσης ενεργειακά αποδοτικές επεκτάσεις των διατυπώσεων των παιγνίων, που επιτρέπουν το συμβιβασμό μεταξύ εξισορρόπησης φορτίου και ενεργειακής κατανάλωσης.

Τα αποτελέσματα προσομοιώσεων καταδεικνύουν αποτελεσματική εξισορρόπηση φορτίου, με το μη ανατροφοδοτούμενο παίγνιο δυναμικού να ξεχωρίζει λόγω της υψηλής του απόδοσης, των εγγυήσεων σύγκλισης και της πλήρως αποκεντρωμένης, επεκτάσιμης φύσης του. Περαιτέρω αξιολογήσεις επιβεβαιώνουν την ανθεκτικότητα της προσέγγισης αυτής υπό διαφορετικά επίπεδα φόρτου συστήματος, τοπολογίες δικτύων και παραμέτρους εκμάθησης.

Λέξεις Κλειδιά

Εξισορρόπηση Φορτίου, Εκφόρτωση Υπολογισμών, Δορυφορικές Υποδομές Υπολογιστικής Παρυφής, Στοχαστικά Παίγνια, Παίγνια Δυναμικού, Μετανοητική Μάθηση

Abstract

This thesis investigates the challenge of load balancing in Satellite Edge Computing (SEC) environments. In this context, ground-based Internet of Things (IoT) users offload computational tasks to satellite servers for distributed processing. The heterogeneity of these tasks and the stochastic nature of their arrival lead to an uneven workload distribution among the satellites. The objective is to develop optimal load-balancing mechanisms aimed at achieving temporal and energy efficiency, as well as fairness.

The proposed strategy follows an iterative two-fold process: (i) selecting an appropriate target satellite for each source satellite in every round, and (ii) determining the optimal load amount to offload. While a deterministic approach governs load amount selection, the target selection leverages game-theoretic mechanisms that account for communication-specific constraints. These mechanisms include a stochastic game formulation, along with noisy potential games modeled in both bandit and semi-bandit frameworks.

The solution of these games utilizes No-Regret Learning techniques, particularly through implementation of the ϵ -Hedge variant of the Exponential Weights Algorithm (EWA), which offers convergence guarantees to δ -equilibria and Nash equilibria for the potential game-based models. Energy-aware extensions of the game formulations are also proposed, allowing for a trade-off between load balancing and energy efficiency.

Simulation results demonstrate effective load balancing, with the potential game in the bandit setting standing out due to its strong performance, convergence guarantees, and fully decentralized, scalable nature. Further evaluations confirm the robustness of this approach under varying system loads, network topologies, and learning parameters.

Keywords

Load Balancing, Computation Offloading, Satellite Edge Computing, Stochastic Games, Potential Games, No-Regret Learning

Ευχαριστίες

Θα ήθελα κατ'αρχάς να ευχαριστήσω τον κ. Καθηγητή Συμεών Παπαβασιλείου, καθώς είμαι ευγνώμων για την επίβλεψη, την εμπιστοσύνη και τη συνεχή του υποστήριξη στα πλαίσια αυτής της διπλωματικής αλλά και ευρύτερα. Φυσικά, ευχαριστώ επίσης τη μεταδιδαστορική ερευνήτρια Μαρία Διαμαντή και τον υποψήφιο διδάκτορα Παναγιώτη Χαρατσάρη για την πολύτιμη καθοδήγηση και συνεργασία.

Η πενταετής διαδρομή των σπουδών είναι μια διαδικασία εξερεύνησης και συνεχών αλλαγών, αλλά οι σταθερές σε αυτή τη διαδρομή -η οικογένεια και οι φίλοι- θα είναι πάντα ο παράγοντας που την καθιστά πλήρη. Το ευχαριστώ δε μπορεί παρά να απευθύνεται και σε αυτούς.

Περιεχόμενα

Περίληψη	1
Abstract	3
Ευχαριστίες	5
Περιεχόμενα	8
Κατάλογος Σχημάτων	9
1 Εισαγωγή	11
1.1 Πρόλογος	11
1.2 Βιβλιογραφική Επισκόπηση	12
1.3 Κίνητρο και Συνεισφορά	13
1.4 Διάρθρωση Διπλωματικής Εργασίας	13
2 Θεωρητικό Υπόβαθρο	15
2.1 Βασικές Αρχές Θεωρίας Παιγνίων	15
2.1.1 Γενικοί Ορισμοί	15
2.1.2 Παίγνια σε Δυναμικά Περιβάλλοντα	17
2.1.3 Παίγνια Δυναμικού (Potential Games)	18
2.1.4 Στοχαστικά Παίγνια (Stochastic Games)	18
2.2 Μετανοητική Μάθηση (No-Regret Learning)	19
2.2.1 Εξωτερική Μετάνοια (External Regret)	19
2.2.2 Μέθοδος "Ακολουθήσε-τον-Ηγέτη" (Follow the Leader - FTL)	20
2.2.3 Μέθοδος "Ακολουθήσε-τον-Κανονικοποιημένο-Ηγέτη" (Follow the Regularized Leader - FTRL)	20
2.2.4 Αλγόριθμος Εκθετικών Βαρών (Exponential Weights Algorithm)	22

2.2.5	Σύγκλιση για παίγνια δυναμικού	24
3	Μοντελοποίηση Συστήματος	27
3.1	Δορυφορικό Δίκτυο	27
3.2	Υπολογιστικό Σύστημα	28
3.3	Στρατηγική εκφόρτωσης υπολογισμών	29
3.3.1	Στρατηγική ποσού εκφόρτωσης	31
3.3.2	Στρατηγική στόχου εκφόρτωσης	33
4	Παιγνιο-θεωρητική μοντελοποίηση	35
4.1	Γενική μοντελοποίηση του προβλήματος απόφασης	35
4.2	Παίγνιο G_1 (Στοχαστικό Παίγνιο)	37
4.3	Παίγνιο G_2/G_2^{EXP} (Παίγνιο Δυναμικού)	37
4.4	Ενεργειακά αποδοτικά παίγνια	45
4.5	Σύνοψη μοντελοποιήσεων	46
5	Αξιολόγηση και Αριθμητικά Αποτελέσματα	47
5.1	Παραμετροποίηση	47
5.2	Αποτελέσματα	47
5.2.1	Σύγκριση μεταξύ των τριών παιγνιο-θεωρητικών μοντέλων	47
5.2.2	Συμπεριφορά του συστήματος υπό μεταβλητό αριθμό δορυφόρων	50
5.2.3	Συμπεριφορά του συστήματος υπό μεταβλητό μέσο ρυθμό παραγωγής εργασιών	51
5.2.4	Συμπεριφορά του συστήματος υπό μεταβλητό παράγοντα εξερεύνησης	52
5.2.5	Συμπεριφορά του ενεργειακά αποδοτικού παιγνίου	52
6	Σύνοψη και Συμπεράσματα	55
	Βιβλιογραφία	57
	Γλωσσάριο	61

Κατάλογος σχημάτων

3.1	Δορυφορική Υπολογιστική Υποδομή Παρυφής	28
5.1	Πείραμα 1 - Παιγνίο G_1 : Σύγκλιση πιθανοτήτων επιλογής δορυφόρου-στόχου για το δορυφόρο (πομπό) νο.8.	48
5.2	Πείραμα 1 - Παιγνίο G_2 (Μη ανατροφοδοτούμενη περίπτωση): Σύγκλιση πιθανοτήτων επιλογής δορυφόρου-στόχου για το δορυφόρο (πομπό) νο.8.	49
5.3	Πείραμα 1 - Παιγνίο G_2 (Ημιανατροφοδοτούμενη περίπτωση): Σύγκλιση πιθανοτήτων επιλογής δορυφόρου-στόχου για το δορυφόρο (πομπό) νο.8.	49
5.4	Πείραμα 1: Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (σύγκριση των τριών παιγνίων)	50
5.5	Πείραμα 2 - Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (μεταβολή του πλήθους των δορυφόρων).	51
5.6	Πείραμα 3 - Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (μεταβολή του εύρους των τιμών λ_i).	52
5.7	Πείραμα 4 - Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (μεταβολή του παράγοντα εξερεύνησης).	53
5.8	Πείραμα 5 - Συνολική ενέργεια που καταναλώνεται σε κάθε γύρο λόγω των εκφορτώσεων (μεταβολή του συντελεστή ενεργειακής απόδοσης).	54
5.9	Πείραμα 5 - Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (μεταβολή του συντελεστή ενεργειακής απόδοσης).	54

1.1 Πρόλογος

Η ραγδαία εξέλιξη των δικτύων επόμενης γενιάς—με χαρακτηριστικά όπως η καθολική συνδεσιμότητα, η πυκνότητα των συσκευών και η ανάγκη για απόκριση σε πραγματικό χρόνο—έχει καταστήσει το Διαδίκτυο των Πραγμάτων (IoT) θεμελιώδη παράγοντα των σύγχρονων ψηφιακών υποδομών. Καθώς οι IoT συσκευές εξαπλώνονται σε αστικές, αγροτικές αλλά και απομακρυσμένες περιοχές, οι παραδοσιακές επίγειες υποδομές συχνά αδυνατούν να προσφέρουν την απαιτούμενη ποιότητα εξυπηρέτησης. Στο πλαίσιο αυτό, οι Δορυφορικές Υπολογιστικές Υποδομές Παρυφής (Satellite Edge Computing) αναδεικνύονται ως μια υποσχόμενη αρχιτεκτονική, επιτρέποντας στους δορυφόρους να λειτουργούν όχι μόνο ως κόμβοι επικοινωνίας, αλλά και ως αποκεντρωμένοι υπολογιστικοί πόροι που επεξεργάζονται άμεσα δεδομένα από IoT συσκευές. Ωστόσο, η πλήρης αξιοποίηση αυτής της αρχιτεκτονικής φέρνει στο προσκήνιο νέες προκλήσεις, με κυριότερη την εξισορρόπηση φορτίου (load balancing).

Η ανάγκη για εξισορρόπηση φορτίου προκύπτει από το γεγονός ότι ορισμένοι δορυφόροι δέχονται, κατά μέσο όρο, υπερβολικά μεγάλο φορτίο προς επεξεργασία—είτε λόγω του πλήθους των διεργασιών, είτε λόγω της υπολογιστικής πολυπλοκότητας κάθε διεργασίας— ενώ ταυτόχρονα διαθέτουν περιορισμένες υπολογιστικές δυνατότητες. Αυτό έχει ως αποτέλεσμα το γρήγορο κορεσμό των ουρών επεξεργασίας τους, οι οποίες καθίστανται ανίκανες να δεχθούν νέες διεργασίες. Αντιθέτως, η αξιοποίηση της συλλογικής υπολογιστικής ικανότητας όλων των δορυφόρων μέσω κατάλληλης εξισορρόπησης θα μπορούσε να διατηρήσει σταθερές ουρές και να αποτρέψει τον κορεσμό του συστήματος.

Το πρόβλημα αυτό μπορεί να διατυπωθεί ως ένα πρόβλημα απόφασης, στο οποίο κάθε δορυφόρος καλείται να επιλέξει πόσες διεργασίες θα εκφορτώσει σε άλλους δορυφόρους, και ποιοι θα είναι οι δορυφόροι αυτοί. Είναι λοιπόν κατάλληλο για μοντελοποίηση μέσω θεωρίας παιγνίων, καθώς οι δορυφόροι λειτουργούν ως ορθολογικοί παίκτες που στοχεύουν στη βελτιστοποίηση της απόδοσής τους.

Συνεπώς, το επίκεντρο αυτής της διπλωματικής εργασίας είναι η μελέτη κατάλληλων παιγνιοθεωρητικών μηχανισμών που επιτρέπουν την επίτευξη αποδοτικών στρατηγικών εξισορρόπησης φορτίου σε Δορυφορικές Υπολογιστικές Υποδομές Παρυφής.

1.2 Βιβλιογραφική Επισκόπηση

Η καταναμεμένη εξισορρόπηση φορτίου με παιγνιο-θεωρητικούς μηχανισμούς έχει μελετηθεί εκτενώς σε διάφορα υπολογιστικά περιβάλλοντα, όπως σε συστήματα επεξεργαστών, σε Υπολογιστικές Υποδομές Παρυφής (Mobile Edge computing - MEC) και σε ενσωματωμένα επίγεια/δορυφορικά δίκτυα (Terrestrial Networks / Non-Terrestrial Networks). Οι προσεγγίσεις ποικίλλουν ανάλογα με το αν η λήψη αποφάσεων καθοδηγείται από τον χρήστη που εκφορτώνει διεργασίες στα συστήματα αυτά, ή από τις ίδιες τις υπολογιστικές μονάδες, με τη διαθεσιμότητα καθολικών πληροφοριών για το σύστημα, αλλά και με την επιλογή του κατάλληλου παιγνιο-θεωρητικού πλαισίου.

Στις πρώτες μελέτες, η εξισορρόπηση φορτίου σε καταναμεμένους υπολογιστές διατυπώθηκε ως συνεργατικό παίγνιο μεταξύ διακομιστών. Για παράδειγμα, στο [GCL02], το πρόβλημα της στατικής εξισορρόπησης φορτίου μοντελοποιήθηκε ως συνεργατικό παίγνιο και χρησιμοποιήθηκε η Λύση Διαπραγματεύσεως Nash (Nash Bargaining Solution - NBS) για να εξασφαλιστεί η κατά Pareto βελτιστότητα και η δικαιοσύνη μεταξύ των διεργασιών, υποθέτοντας πλήρη γνώση των δυνατοτήτων επεξεργασίας των διακομιστών και των ρυθμών άφιξης διεργασιών. Αντίθετα, τα [GC05] και [FE25] ανέπτυξαν μη συνεργατικά παίγνια εξισορρόπησης φορτίου από την οπτική γωνία του χρήστη, στα οποία οι χρήστες κατανέμουν τις διεργασίες τους σε ετερογενείς διακομιστές με κατάλληλη στρατηγική ανάθεσης ώστε να βελτιστοποιήσουν την ατομική τους απόδοση.

Η τελευταία προσέγγιση, κατά την οποία η στοχευμένη εκφόρτωση από πλευράς χρηστών επιτυγχάνει έμμεσα την εξισορρόπηση του φορτίου των διακομιστών, έχει υιοθετηθεί ευρέως στην εξισορρόπηση φορτίου των Υπολογιστικών Υποδομών Παρυφής. Στο [Hu+20], οι συγγραφείς διατύπωσαν ένα μη συνεργατικό παίγνιο μεταξύ κινητών συσκευών με ελλιπείς πληροφορίες για το συνολικό σύστημα, και απέδειξαν την ύπαρξη Ισορροπιών Nash (Nash Equilibria). Άλλοι ερευνητές χρησιμοποίησαν τη Θεωρία Παιγνίων Μέσου Πεδίου (Mean-field Game Theory) για να μοντελοποιήσουν μεγάλους πληθυσμούς χρηστών που αλληλεπιδρούν με διακομιστές Υπολογιστικών Υποδομών Παρυφής ([Wu+24], [YCM24]).

Στα δορυφορικά δίκτυα, ιδίως σε εκείνα που περιλαμβάνουν Δορυφόρους Χαμηλής Τροχιάς (LEO), η έμφαση δίνεται και πάλι στην έμμεση εξισορρόπηση του φορτίου των εξυπηρετητών μέσω της στοχευμένης εκφόρτωσης υπολογισμών από τους χρήστες, και όχι στην άμεση εξισορρόπηση του φορτίου μέσω της εκφόρτωσης μεταξύ των εξυπηρετητών. Αρκετές εργασίες ([Jia+22], [Gao+24b], [Gao+24a], [Wan+20]) διατύπωσαν την εκφόρτωση πολλαπλών χρηστών και τον έλεγχο της ισχύος μετάδοσης ως προβλήματα θεωρίας παιγνίων, αποδεικνύοντας την ύπαρξη ισορροπιών Nash. Αυτές οι μελέτες επικεντρώνονται στην ελαχιστοποίηση της παρεμβολής στο κανάλι, στη βελτιστοποίηση της κατανάλωσης ενέργειας ή στη μείωση των χρόνων απόκρισης διεργασιών σε ενσωματωμένα περιβάλλοντα επικοινωνίας δορυφόρου-εδάφους.

Η συνεργατική εκφόρτωση μεταξύ εξυπηρετητών για την εξισορρόπηση φορτίου δεν έχει διερευνηθεί σημαντικά μέσω προσεγγίσεων θεωρίας παιγνίων. Ορισμένες μελέτες ([Zha+21], [Lin+25]) πρότειναν την εξισορρόπηση φορτίου με παιγνιο-θεωρητικούς μηχανισμούς μεταξύ Διασυνδεδεμένων Εξυπηρετητών Παρυφής (Intelligent Edge Servers -IES), οι οποίοι επικοινωνώ-

νούν μέσω οπτικών ινών. Ωστόσο, αυτές οι προσεγγίσεις είτε προϋποθέτουν πρόσβαση σε πληροφορίες για το συνολικό σύστημα, είτε περιορίζονται σε τοπικές πληροφορίες γειτονικών εξυπηρετητών, περιορίζοντας την επεκτασιμότητα σε μεγάλες τοπολογίες.

Ειδικά σε ό,τι αφορά συγκεκριμένα σε δορυφορικά δίκτυα, οι μελέτες σχετικά με τη συνεργατική εκφόρτωση μεταξύ των δορυφόρων για την επίτευξη εξισορρόπησης φορτίου υιοθετούν μια γενική διατύπωση Διαδικασίας Απόφασης Markov (Markov Decision Process -MDP), παρά μια μοντελοποίηση βασισμένη στη θεωρία παιγνίων ([Shi+24]).

1.3 Κίνητρο και Συνεισφορά

Όπως προκύπτει από την ανάλυση της σχετικής βιβλιογραφίας, παρότι έχει σημειωθεί σημαντική πρόοδος στην κατανομημένη εξισορρόπηση φορτίου σε ποικίλες αρχιτεκτονικές δικτύων, ελάχιστες μελέτες έχουν εστιάσει στην ανάπτυξη παιγνιοθεωρητικών μοντέλων συνεργατικής εκφόρτωσης μεταξύ εξυπηρετητών, ειδικά προσαρμοσμένων στις απαιτήσεις των Δορυφορικών Υπολογιστικών Υποδομών Παρυφής. Τα υφιστάμενα μοντέλα επιτυγχάνουν έμμεση εξισορρόπηση φορτίου των εξυπηρετητών μέσω αποκεντρωμένης λήψης αποφάσεων εκφόρτωσης από τους χρήστες, ενώ οι στρατηγικές που εστιάζουν στους εξυπηρετητές είτε βασίζονται σε ισχυρές υποθέσεις γνώσης της κατάστασης του συστήματος, είτε δεν αξιοποιούν παιγνιοθεωρητικούς μηχανισμούς. Η κατάσταση αυτή αναδεικνύει την ανάγκη για υπολογιστικά αποδοτικές, κατανομημένες και κλιμακώσιμες μεθόδους συνεργασίας εξυπηρετητών, θεμελιωμένες στη θεωρία παιγνίων, ικανές να ανταποκριθούν στις ιδιαιτερότητες των δορυφορικών δικτύων και τη δυναμικότητα του συστήματος.

1.4 Διάρθρωση Διπλωματικής Εργασίας

Στο δεύτερο κεφάλαιο παρουσιάζεται το σχετικό θεωρητικό υπόβαθρο της εργασίας. Περιγράφονται οι βασικές αρχές της Θεωρίας Παιγνίων, με έμφαση στα Πάιγνια Δυναμικού (Potential Games) και στα Στοχαστικά Πάιγνια (Stochastic Games), αλλά και αλγόριθμοι Μετανοητικής Μάθησης (No-Regret Learning) για προβλήματα λήψης αποφάσεων.

Στο τρίτο κεφαλαίο παρουσιάζεται η μοντελοποίηση του συστήματος, ως προς τις παραμέτρους διαδορυφορικής επικοινωνίας, παραγωγής διεργασιών και εκφόρτωσής τους. Αναλύεται επίσης η στρατηγική εκφόρτωσης των δορυφόρων, ειδικά ως προς το ποσό εκφόρτωσης.

Στο τέταρτο κεφάλαιο συνεχίζεται η μοντελοποίηση της στρατηγικής εκφόρτωσης των δορυφόρων, αυτή τη φορά ως προς το στόχο εκφόρτωσης, με χρήση παιγνιο-θεωρητικών μηχανισμών. Καταδεικνύεται επίσης η σύγκλιση επιλεγμένων παιγνίων με χρήση αλγορίθμων μετανοητικής μάθησης στα επιθυμητά σημεία ισορροπίας.

Στο πέμπτο κεφάλαιο παρατίθενται τα πειραματικά αποτελέσματα των προσομοιώσεων και μια πρώτη αξιολόγησή τους.

Τα τελικά συμπεράσματα συνοψίζονται στο έκτο κεφάλαιο.

2.1 Βασικές Αρχές Θεωρίας Παιγνίων

2.1.1 Γενικοί Ορισμοί

Η θεωρία παιγνίων είναι ένα σύνολο αναλυτικών εργαλείων που μας βοηθούν να κατανοήσουμε τα φαινόμενα που παρατηρούνται όταν αλληλεπιδρούν διάφοροι λήπτες αποφάσεων [OR94].

Αποτελεί ένα θεωρητικό πλαίσιο που βασίζεται στη δημιουργία αυστηρών μοντέλων, τα οποία περιγράφουν καταστάσεις σύγκρουσης ή/και συνεργασίας μεταξύ ορθολογικών ατόμων. Σύμφωνα με την κλασική θεωρία αποφάσεων και την οικονομική επιστήμη, η ορθολογική συμπεριφορά ορίζεται ως η επιλογή ενεργειών που μεγιστοποιούν την απόδοση (ή κάποια μορφή ωφελιμότητας), λαμβάνοντας υπόψη τους περιορισμούς που αντιμετωπίζει κάθε παίκτης. Η θεωρία παιγνίων έχει εφαρμοστεί με επιτυχία σε διάφορους τομείς, όπως ο επιχειρηματικός ανταγωνισμός, η λειτουργία των αγορών, οι πολιτικές εκστρατείες και οι δημοπρασίες, ενώ έχει συμβάλει σημαντικά στην κατανόηση άλλων επιστημονικών πεδίων, όπως η εξελικτική βιολογία και η ψυχολογία [Tad13].

Σε όλα τα μοντέλα της θεωρίας παιγνίων, η βασική μονάδα είναι ο παίκτης. Ο όρος «παίκτης» μπορεί να αναφέρεται είτε σε ένα άτομο είτε σε μια ομάδα ατόμων που λαμβάνουν από κοινού αποφάσεις.

Ορισμός 2.1 (Παίγνιο). Ένα παίγνιο (*game*) είναι μια αναπαράσταση στρατηγικής αλληλεπίδρασης, η οποία περιλαμβάνει τόσο τους περιορισμούς στις επιλογές που έχουν οι παίκτες όσο και τα συμφέροντα ή τις προτιμήσεις τους. Ωστόσο, δεν καθορίζει ποιες ενέργειες θα επιλέξουν τελικά οι παίκτες [OR94].

Ορισμός 2.2 (Λύση). Μια λύση (*solution*) είναι μια συστηματική περιγραφή των αποτελεσμάτων που μπορούν να προκύψουν σε μια κατηγορία παιγνίων [OR94].

Ορισμός 2.3 (Καθαρή Στρατηγική). Μια καθαρή στρατηγική (*pure strategy*) για τον παίκτη i είναι ένα καθορισμένο, ντετερμινιστικό σχέδιο δράσης. Το σύνολο όλων των καθαρών στρατηγικών του

παίκτη i συμβολίζεται με S_i . Ένα προφίλ καθαρών στρατηγικών $s = (s_1, s_2, \dots, s_n)$, $s_i \in S_i$ για κάθε $i = 1, 2, \dots, n$, περιγράφει ένα συγκεκριμένο συνδυασμό καθαρών στρατηγικών που έχουν επιλεγεί από όλους τους n παίκτες του παιχνιδιού [Tad13].

Ορισμός 2.4 (Παίγνιο σε Κανονική Μορφή). Ένα παίγνιο σε κανονική μορφή (*normal-form game*) περιλαμβάνει τα εξής τρία βασικά συστατικά:

1. Ένα πεπερασμένο σύνολο παικτών, $N = \{1, 2, \dots, n\}$.
2. Μια συλλογή από σύνολα καθαρών στρατηγικών, $\{S_1, S_2, \dots, S_n\}$.
3. Ένα σύνολο συναρτήσεων απόδοσης, $\{v_1, v_2, \dots, v_n\}$, όπου κάθε συνάρτηση αποδίδει μια τιμή απόδοσης για κάθε πιθανό συνδυασμό επιλεγμένων στρατηγικών, δηλαδή μια συνάρτηση της μορφής

$$v_i : S_1 \times S_2 \times \dots \times S_n \rightarrow \mathbb{R}$$

για κάθε $i \in N$.

Προκειμένου να απλοποιηθούν οι συμβολισμοί, θα χρησιμοποιούμε την ακόλουθη συντομογραφία: Ορίζουμε $S_{-i} \equiv S_1 \times S_2 \times \dots \times S_{i-1} \times S_{i+1} \times \dots \times S_n$ ως το σύνολο των συνόλων στρατηγικών των υπόλοιπων παικτών εκτός του παίκτη i . Έπειτα, ορίζουμε ότι $s_{-i} \in S_{-i}$ είναι ένα συγκεκριμένο προφίλ στρατηγικών για όλους τους παίκτες εκτός του i . Συνεπώς, μπορούμε να επαναδιατυπώσουμε την απόδοση του παίκτη i από τη στρατηγική s ως $v_i(s_i, s_{-i})$, όπου $s = (s_i, s_{-i})$ [Tad13].

Ορισμός 2.5 (Λύση κυρίαρχης στρατηγικής). Ένα διανύσμα στρατηγικών $s \in S$ είναι μια λύση κυρίαρχης στρατηγικής, αν για κάθε παίκτη i , και για κάθε εναλλακτικό διανύσμα στρατηγικών $s' \in S$, ισχύει ότι

$$u_i(s_i, s'_{-i}) \geq u_i(s'_i, s'_{-i}).$$

Αξίζει να σημειωθεί ότι μια λύση κυρίαρχης στρατηγικής ενδέχεται να μην προσφέρει τη βέλτιστη απόδοση σε κανέναν από τους παίκτες. Επίσης, η ύπαρξη μίας κυρίαρχης στρατηγικής για κάθε παίκτη είναι μια εξαιρετικά αυστηρή προϋπόθεση για ένα παίγνιο και πολύ λίγα παιχνίδια την ικανοποιούν [Nis+07].

Λεδομένου λοιπόν ότι τα παιχνίδια σπάνια διαθέτουν λύσεις κυρίαρχης στρατηγικής, είναι απαραίτητο να αναζητήσουμε μια λιγότερο αυστηρή έννοια λύσης.

Ορισμός 2.6 (δ-Ισορροπία). Ένα διανύσμα στρατηγικών $s \in S$ αποτελεί δ-ισορροπία, αν για όλους τους παίκτες i και για κάθε εναλλακτική στρατηγική $s'_i \in S_i$, ισχύει ότι [CHM17]:

$$u_i(s_i, s_{-i}) + \delta \geq u_i(s'_i, s_{-i}).$$

Ορισμός 2.7 (Ισορροπία Nash). Ένα διανύσμα στρατηγικών $s \in S$ αποτελεί ισορροπία Nash, αν για όλους τους παίκτες i και για κάθε εναλλακτική στρατηγική $s'_i \in S_i$, ισχύει ότι

$$u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i}).$$

Με άλλα λόγια, κανένας παίκτης i δεν μπορεί να αλλάξει τη στρατηγική του από s_i σε s'_i και να βελτιώσει έτσι την απόδοσή του, υποθέτοντας ότι όλοι οι υπόλοιποι παίκτες διατηρούν τις στρατηγικές που έχουν επιλέξει.

Προφανώς, μια λύση κυρίαρχης στρατηγικής είναι ισορροπία Nash, ενώ, αν η λύση είναι αυστηρά κυρίαρχη (δηλαδή, η μετάβαση σε αυτή βελτιώνει πάντα αυστηρά το αποτέλεσμα), τότε αποτελεί και τη μοναδική ισορροπία Nash. Ωστόσο, οι ισορροπίες Nash γενικότερα ενδέχεται να μην είναι μοναδικές. Εκτός αυτών, γνωρίζουμε ήδη ότι οι ισορροπίες Nash ενδέχεται να μην είναι βέλτιστες για τους παίκτες, δεδομένου ότι οι λύσεις κυρίαρχης στρατηγικής είναι ισορροπίες Nash. Σε παίγνια με πολλαπλές ισορροπίες Nash, διαφορετικές ισορροπίες μπορεί να έχουν (σημαντικά) διαφορετικές αποδόσεις για τους παίκτες [Nis+07].

Ένα παίγνιο δεν είναι απαραίτητο να διαθέτει καθαρές στρατηγικές ισορροπίες Nash. Ωστόσο, παρουσιάζονται επεκτάσεις αν επιτρέψουμε τυχαιοποιημένες στρατηγικές. Για να ορίσουμε τις τυχαιοποιημένες στρατηγικές με συστηματικό τρόπο, διευρύνουμε τις επιλογές των παικτών έτσι ώστε ο καθένας να μπορεί να επιλέξει μια κατανομή πιθανότητας επί του συνόλου των πιθανών στρατηγικών του· μια τέτοια επιλογή ονομάζεται μεικτή στρατηγική. Υποθέτουμε δηλαδή ότι οι παίκτες επιλέγουν στρατηγικές ανεξάρτητα, βάσει των αντίστοιχων κατανομών πιθανότητας, με τις ανεξάρτητες τυχαίες επιλογές των παικτών να οδηγούν σε μια κατανομή πιθανότητας επί των διανυσμάτων στρατηγικών s . Ο Nash (1951) απέδειξε ότι υπό αυτή την επέκταση, κάθε παιχνίδι με πεπερασμένο αριθμό παικτών, καθένας από τους οποίους έχει πεπερασμένο σύνολο στρατηγικών, έχει τουλάχιστον μία ισορροπία Nash.

Θεώρημα 2.1. Κάθε παιχνίδι με πεπερασμένο σύνολο παικτών και πεπερασμένο σύνολο στρατηγικών έχει μια ισορροπία Nash με μεικτές στρατηγικές [Nis+07].

2.1.2 Παίγνια σε Δυναμικά Περιβάλλοντα

Έστω τώρα ότι οι αποδόσεις βασίζονται σε μια τυχαία μεταβλητή που μπορεί να λάβει οποιαδήποτε τιμή σε ένα συνεχές διάστημα X . Ο ορισμός της αναμενόμενης απόδοσης είναι ανάλογος και διατυπώνεται ως εξής [Tad13]:

Ορισμός 2.8 (2.4). Έστω $u(x)$ η συνάρτηση απόδοσης του παίκτη για την τυχαία μεταβλητή στο διάστημα $X = [\underline{x}, \bar{x}]$, και έστω ότι η κατανομή περιγράφεται από τη συνάρτηση πυκνότητας πιθανότητας $f(x)$. Τότε ορίζουμε την αναμενόμενη απόδοση του παίκτη ως

$$E[u(x)] = \int_{\underline{x}}^{\bar{x}} u(x) f(x) dx.$$

Μπορούμε στη συνέχεια να τροποποιήσουμε τον προηγούμενο ορισμό της ισορροπίας Nash

ώστε να λαμβάνει υπόψη τη δυναμική φύση του προβλήματος:

Ένα διάνυσμα στρατηγικών $s \in S$ αποτελεί ισορροπία Nash με τη στοχαστική έννοια, αν για όλους τους παίκτες i και για κάθε εναλλακτική στρατηγική $s'_i \in S_i$, ισχύει ότι

$$E[u_i(s_i, s_{-i})] \geq E[u_i(s'_i, s_{-i})].$$

Με άλλα λόγια, κανένας παίκτης i δεν μπορεί να αλλάξει τη στρατηγική του από s_i σε s'_i και να βελτιώσει έτσι την αναμενόμενη απόδοσή του, υποθέτοντας ότι όλοι οι υπόλοιποι παίκτες διατηρούν τις στρατηγικές που έχουν επιλέξει.

2.1.3 Παίγνια Δυναμικού (Potential Games)

Βάσει του [Nis+07], για οποιοδήποτε πεπερασμένο παίγνιο, μια ακριβής συνάρτηση δυναμικού Φ είναι μια συνάρτηση που αντιστοιχίζει κάθε διάνυσμα στρατηγικών S σε κάποια πραγματική τιμή και ικανοποιεί την ακόλουθη συνθήκη: Αν $S = (s_1, s_2, \dots, s_k)$, $s'_i \neq s_i$ είναι μια εναλλακτική στρατηγική για κάποιον παίκτη i , και $S' = (s_{-i}, s'_i)$, τότε ισχύει

$$\Phi(S') - \Phi(S) = u_i(s'_i) - u_i(s_i).$$

Με άλλα λόγια, αν η τρέχουσα κατάσταση του παιγνίου είναι S , και ο παίκτης i αλλάξει από τη στρατηγική s_i στη στρατηγική s'_i , τότε η μεταβολή στην απόδοση του i αντιστοιχεί ακριβώς στη μείωση της τιμής της συνάρτησης δυναμικού.

Κάθε παίγνιο έχει το πολύ μία συνάρτηση δυναμικού (με την πιθανή πρόσθεση μιας σταθεράς). Ένα παίγνιο που διαθέτει μια ακριβή συνάρτηση δυναμικού ονομάζεται παίγνιο δυναμικού. Η δομή αυτή έχει ισχυρές συνέπειες όσον αφορά την ύπαρξη και τη σύγκλιση σε σημεία ισορροπίας.

Θεώρημα 2.2. Κάθε παίγνιο δυναμικού έχει τουλάχιστον μία καθαρή ισορροπία Nash, συγκεκριμένα τη στρατηγική S που ελαχιστοποιεί τη $\Phi(S)$.

Επεκτείνοντας το συλλογισμό, παρατηρούμε ότι οποιοδήποτε σύνολο στρατηγικών S με την ιδιότητα ότι η Φ δεν μπορεί να μειωθεί με την αλλαγή κάποιας μεμονωμένης στρατηγικής στο S , αποτελεί ισορροπία Nash βάσει του ίδιου επιχειρήματος.

2.1.4 Στοχαστικά Παίγνια (Stochastic Games)

Ενώ η σχετική απλότητα των παιχνιδιών κανονικής μορφής είναι χρήσιμη για τη μελέτη των αλληλεπιδράσεων μεταξύ παικτών, δε μοντελοποιείται σε αυτά η έννοια του περιβάλλοντος που επηρεάζεται από τις ενέργειες των παικτών. Τα στοχαστικά παίγνια ορίζουν ένα περιβάλλον βασισμένο σε καταστάσεις, στο οποίο η κατάσταση εξελίσσεται με την πάροδο του χρόνου βάσει των ενεργειών των παικτών και των πιθανοτικών μεταβάσεων μεταξύ καταστάσεων [ACS24]. Πιο συγκεκριμένα, βάσει του [Sol22] έχουμε:

Ορισμός 2.9 (4.1). Ένα στοχαστικό παίγνιο είναι ένα διάνυσμα $\Gamma = \langle I, S, (A^i(s))_{s \in S}^{i \in I}, q, r \rangle$, όπου

- $I = \{1, 2, \dots, n\}$ είναι ένα πεπερασμένο σύνολο παικτών.
- S είναι ένα πεπερασμένο σύνολο καταστάσεων.
- Για κάθε $i \in I$ και κάθε κατάσταση $s \in S$, $A^i(s)$ είναι ένα πεπερασμένο σύνολο ενεργειών διαθέσιμων στον παίκτη i στην κατάσταση s . Συμβολίζουμε το σύνολο όλων των προφίλ ενεργειών στην κατάσταση s με $A(s) = \prod_{i \in I} A^i(s)$, και το σύνολο όλων των προφίλ ενεργειών σε όλες τις καταστάσεις με

$$SA = \{(s, a) : s \in S, a \in A(s)\}.$$

- $q : SA \rightarrow \Delta(S)$ είναι ένας κανόνας μετάβασης.
- Για κάθε παίκτη $i \in I$, $r^i : SA \rightarrow \mathbf{R}$ είναι η συνάρτηση απόδοσης.

Το παίγνιο ξεκινά από μια αρχική κατάσταση $s_1 \in S$ και παίζεται σε στάδια. Σε κάθε στάδιο $t \geq 1$ συμβαίνουν τα εξής:

- Η τρέχουσα κατάσταση s_t ανακοινώνεται στους παίκτες.
- Κάθε παίκτης $i \in I$ επιλέγει μια ενέργεια $a_t^i \in A^i(s_t)$. Οι επιλογές των παικτών γίνονται ταυτόχρονα και ανεξάρτητα.
- Το προφίλ ενεργειών $a_t = (a_t^i)_{i \in I}$ ανακοινώνεται δημόσια σε όλους τους παίκτες.
- Κάθε παίκτης $i \in I$ λαμβάνει απόδοση σταδίου $r^i(s_t, a_t)$.
- Μια νέα κατάσταση $s_{t+1} \in S$ επιλέγεται σύμφωνα με την κατανομή πιθανοτήτων $q(\cdot \mid s_t, a_t)$, και το παίγνιο συνεχίζεται στο στάδιο $t + 1$.

Τα στοχαστικά παίγνια περιλαμβάνουν ως ειδική περίπτωση τα επαναλαμβανόμενα παίγνια κανονικής μορφής, στα οποία υπάρχει μόνο μία κατάσταση στο S και δεν υφίστανται τερματικές καταστάσεις (δηλαδή, $\bar{S} = \emptyset$). Πιο γενικά, αν ορίσουμε τις συναρτήσεις απόδοσης ως $\mathcal{R}_I(s, a)$, τότε κάθε κατάσταση $s \in S$ στο στοχαστικό παίγνιο μπορεί να θεωρηθεί ως ένα παίγνιο κανονικής μορφής με αποδόσεις που δίνονται από τη $\mathcal{R}_I(s, \cdot)$. Με αυτή την έννοια, τα παίγνια κανονικής μορφής αποτελούν το βασικό δομικό στοιχείο των στοχαστικών παιγνίων [ACS24].

2.2 Μετανοητική Μάθηση (No-Regret Learning)

2.2.1 Εξωτερική Μετάνοια (External Regret)

Ακολουθώντας τη θεμελίωση του [Nis+07], έστω ένα σύστημα με N διαθέσιμες ενέργειες $X = \{1, \dots, N\}$ και ένας αλγόριθμος H που επιλέγει σε κάθε στάδιο/χρονική στιγμή μία από αυτές. Πιο συγκεκριμένα, σε κάθε χρονική στιγμή t , ο αλγόριθμος διαθέτει μια κατανομή p^t επί των N ενεργειών που καθοδηγεί την επιλογή του. Στη συνέχεια, ένας αντίπαλος (adversary)

επιλέγει έναν διάνυσμα απωλειών (loss vector) $\ell^t \in [0, 1]^N$, όπου $\ell_i^t \in [0, 1]$ είναι η απώλεια της i -οστής ενέργειας στη χρονική στιγμή t . Στο μοντέλο πλήρους πληροφορίας, ο αλγόριθμος H λαμβάνει το διάνυσμα απωλειών ℓ^t και υφίσταται απώλεια $\ell_H^t = \sum_{i=1}^N p_i^t \ell_i^t$. (Αυτό μπορεί να θεωρηθεί ως η αναμενόμενη απώλεια όταν ο αλγόριθμος επιλέγει την ενέργεια $i \in X$ με πιθανότητα p_i^t .) Αντιθέτως, στο μοντέλο μερικής πληροφορίας, ο αλγόριθμος λαμβάνει $(\ell_{k^t}^t, k^t)$, όπου k^t κατανέμεται σύμφωνα με την p^t , και η απώλεια του είναι $\ell_H^t = \ell_{k^t}^t$. Η απώλεια της i -οστής ενέργειας κατά τα πρώτα T βήματα χρόνου είναι $L_i^T = \sum_{t=1}^T \ell_i^t$, και η απώλεια του H είναι $L_H^T = \sum_{t=1}^T \ell_H^t$.

Ο στόχος στο πλαίσιο της εξωτερικής μετάνοιας (external regret) είναι να σχεδιαστεί ένας αλγόριθμος που θα μπορεί να προσεγγίσει την απόδοση του καλύτερου αλγορίθμου από μια δοσμένη κλάση αλγορίθμων G , δηλαδή να έχει απώλεια κοντά στο $L_g^{T, \min} = \min_{g \in G} L_g^T$. Φορμαλιστικά, επιθυμούμε να ελαχιστοποιήσουμε την εξωτερική μετάνοια $R_G = L_H^T - L_g^{T, \min}$, και η κλάση G ονομάζεται κλάση σύγκρισης. Η πιο μελετημένη κλάση σύγκρισης G είναι αυτή που αποτελείται από όλες τις μεμονωμένες ενέργειες, δηλαδή $G = X$. Εδώ θα μελετηθεί αυτή ακριβώς η κλάση, δηλαδή θέλουμε η απώλεια του αλγορίθμου να είναι κοντά στο $L_{\min}^T = \min_i L_i^T$, και ορίζουμε την εξωτερική μετάνοια ως $R = L_H^T - L_{\min}^T$.

2.2.2 Μέθοδος "Ακολουθήσε-τον-Ηγέτη" (Follow the Leader - FTL)

Όπως σημειώνεται στο [Sha12], η πιο διαισθητική προσέγγιση είναι να χρησιμοποιείται σε κάθε γύρο οποιαδήποτε ενέργεια συγκεντρώνει τη χαμηλότερη απώλεια σε όλους τους περασμένους γύρους:

$$x_{t+1} = \arg \min_{x \in \mathcal{X}} \sum_{\tau=1}^t f_{\tau}(x),$$

όπου $f_{\tau}(x)$ η συνάρτηση που μοντελοποιεί την απώλεια. Αυτή η μορφή στρατηγικής είναι γνωστή ως "εικονικό παίξιμο" (fictitious play) στα οικονομικά, και έχει ονομαστεί "Ακολουθήσε τον Ηγέτη" (Follow the Leader - FTL) στη μηχανική μάθηση. Η στρατηγική αυτή αποτυγχάνει πλήρως στη χειρότερη περίπτωση, δηλαδή η μετάνοιά της μπορεί να είναι γραμμική ως προς τον αριθμό των επαναλήψεων, όπως δείχνει το παρακάτω παράδειγμα: Έστω $\mathcal{X} = [-1, 1]$, έστω $f_1(x) = \frac{1}{2}x$, και έστω ότι τα f_{τ} για $\tau = 2, \dots, T$ εναλλάσσονται μεταξύ $-x$ και x . Έτσι,

$$\sum_{\tau=1}^t f_{\tau}(x) = \begin{cases} \frac{1}{2}x, & t \text{ περιττός} \\ -\frac{1}{2}x, & \text{αλλιώς.} \end{cases}$$

Η στρατηγική FTL θα εναλλάσσει συνεχώς τις δράσεις μεταξύ $x_t = -1$ και $x_t = 1$, κάνοντας πάντα τη λανθασμένη επιλογή.

2.2.3 Μέθοδος "Ακολουθήσε-τον-Κανονικοποιημένο-Ηγέτη" (Follow the Regularized Leader - FTRL)

Η κατά τα άλλα αρκετά διαισθητική στρατηγική FTL αποτυγχάνει στο παραπάνω παράδειγμα επειδή είναι ασταθής. Μπορούμε να την τροποποιήσουμε ώστε να μην αλλάζει συχνά

αποφάσεις, και επομένως να επιτυγχάνει χαμηλή μετάνοια, με τη μέθοδο "Ακολουθήσε τον Κανονικοποιημένο Ηγέτη" (Follow the Regularized Leader - FTRL). Για μια συνάρτηση κανονικοποίησης $R : X \rightarrow \mathbb{R}$, έχουμε [Sha12]:

$$x_{t+1} = \arg \min_{x \in \mathcal{X}} \left[\sum_{\tau=1}^t f_{\tau}(x) + R(x) \right].$$

Η μέθοδος FTRL επιτυγχάνει μετάνοια

$$\text{Regret}_T = O(\sqrt{T}),$$

όπου T ο αριθμός επαναλήψεων, η οποία είναι βέλτιστη με προσέγγιση σταθεράς. Αυτό αποτελεί σημαντική βελτίωση σε σχέση με τη μέθοδο FTL, η οποία μπορεί να παρουσιάσει γραμμική μετάνοια στη χειρότερη περίπτωση.

Η μέθοδος FTRL σε παιγνιο-θεωρητικά πλαίσια

Στο σημείο αυτό θα συνδέσουμε την προηγούμενη ανάλυση με τη θεωρία παιγνίων, ακολουθώντας τη σημειολογία του [Flo+20].

Έστω ένα παίγνιο που αποτελείται από:

1. ένα πεπερασμένο σύνολο από παίκτες $i \in \mathcal{N} = \{1, \dots, N\}$.
2. ένα πεπερασμένο σύνολο από ενέργειες (ή καθαρές στρατηγικές) $\mathcal{A}_i = \{\alpha_1, \dots, \alpha_{n_i}\}$ για κάθε παίκτη $i \in \mathcal{N}$ και
3. κάθε συνάρτηση απόδοσης του παίκτη $u_i : \mathcal{A} \rightarrow \mathbb{R}$, όπου $\mathcal{A} := \prod_i \mathcal{A}_i$ υποδηλώνει το σύνολο όλων των δυνατών προφίλ ενεργειών $\alpha = (\alpha_1, \dots, \alpha_N)$.

Οι παίκτες μπορούν επίσης να παίζουν μεικτές στρατηγικές, δηλαδή, κατανομές πιθανοτήτων $x_i = (x_{i\alpha_i})_{\alpha_i \in \mathcal{A}_i} \in \Delta(\mathcal{A}_i)$ επί των καθαρών στρατηγικών $\alpha_i \in \mathcal{A}_i$. Συνολικά, θα γράφουμε $\mathcal{X}_i := \Delta(\mathcal{A}_i)$ για το χώρο μεικτών στρατηγικών του παίκτη i και $\mathcal{X} := \prod_i \mathcal{X}_i$ για το χώρο όλων των προφίλ μεικτών στρατηγικών $x = (x_1, \dots, x_N)$.

Δεδομένου ενός μεικτού προφίλ $x \in \mathcal{X}$, η αντίστοιχη αναμενόμενη απόδοση του παίκτη i δίνεται από τη σχέση:

$$u_i(x) = \sum_{\alpha_1 \in \mathcal{A}_1} \cdots \sum_{\alpha_N \in \mathcal{A}_N} x_{1,\alpha_1} \cdots x_{N,\alpha_N} u_i(\alpha_1, \dots, \alpha_N). \quad (2.1)$$

Για να παρακολουθούμε την απόδοση κάθε ατομικής ενέργειας, θα γράφουμε επίσης

$$v_{i\alpha_i}(x) := u_i(\alpha_i; x_{-i}) \quad (2.2)$$

για την απόδοση της καθαρής στρατηγικής $\alpha_i \in \mathcal{A}_i$ στο μεικτό προφίλ $x = (x_i; x_{-i}) \in \mathcal{X}$. Άρα, γράφοντας $v_i(x) := (v_{i\alpha_i}(x))_{\alpha_i \in \mathcal{A}_i} \in \mathbb{R}^{\mathcal{A}_i}$ για το διάνυσμα απόδοσης του παίκτη i , έχουμε τη

σχέση

$$u_i(x) = \langle v_i(x), x_i \rangle = \sum_{\alpha_i \in \mathcal{A}_i} x_{i\alpha_i} v_{i\alpha_i}(x) \quad (2.3)$$

όπου, στην καθιερωμένη σημειογραφία, $\langle v, x \rangle = v^\top x$ δηλώνει το κανονικό εσωτερικό γινόμενο μεταξύ των v και x .

Υπό το πλαίσιο αυτό, σε κάθε χρονική στιγμή $t \geq 0$, η μέθοδος FTRL προτείνει μία μεικτή στρατηγική που μεγιστοποιεί το σωρευτικό κέρδος του παίκτη έως τη χρονική στιγμή t μείον μία ποινή κανονικοποίησης, η οποία ενθαρρύνει την εξερεύνηση. Τυπικά, αυτό περιγράφεται από τις δυναμικές:

$$y_{i\alpha_i}(t) = y_{i\alpha_i}(0) + \int_0^t v_{i\alpha_i}(x(s)) ds \quad (\text{συγκεντρωτικά κέρδη})$$

$$x_{i\alpha_i}(t) = Q_{i\alpha_i}(y_i(t)) \quad (\text{επιλογή στρατηγικής})$$

ή, σε πιο συμπαγή μορφή:

$$\dot{y}(t) = v(Q(y(t))) \quad (\text{FTRL})$$

Στα παραπάνω, κάθε $y_{i\alpha_i}$ έχει το ρόλο μιας βοηθητικής μεταβλητής σκορ, η οποία μετρά τη συνολική απόδοση της καθαρής στρατηγικής $\alpha_i \in A_i$ με την πάροδο του χρόνου. Αυτά τα σκορ μετατρέπονται στη συνέχεια σε μεικτές στρατηγικές μέσω ενός ειδικού για κάθε παίκτη μετασχηματιστή $y_i \mapsto x_i = Q_i(y_i)$, ο οποίος ορίζεται ως:

$$Q_i(y_i) = \arg \max_{x_i \in A_i} \{ \langle y_i, x_i \rangle - h_i(x_i) \} \quad \text{για κάθε } y_i \in \mathcal{Y}_i := \mathbb{R}^{n_i}.$$

όπου $h_i(x_i)$ ένας κυρτός όρος κανονικοποίησης.

2.2.4 Αλγόριθμος Εκθετικών Βαρών (Exponential Weights Algorithm)

Όπως αναφέρεται στο [Flo+20], ένας από τους πιο διαδεδομένους κανονικοποιητές είναι η (αρνητική) εντροπία Gibbs-Shannon $h_i(x_i) = \sum_{\alpha_i} x_{i\alpha_i} \log x_{i\alpha_i}$. Μπορεί να αποδειχθεί τότε ότι οδηγούμαστε σε ένα χάρτη πιθανοτικής επιλογής (logit choice map), ο οποίος γράφεται σε διανυσματική μορφή ως

$$\Lambda_i(y_i) = \frac{\exp(y_i)}{\sum_{\alpha_i \in \mathcal{A}_i} \exp(y_{i\alpha_i})}.$$

Αυτός, με τη σειρά του, οδηγεί στις δυναμικές εκθετικών βαρών:

$$\dot{y}_i(t) = v_i(x(t)),$$

$$x_i(t) = \Lambda_i(y_i(t)).$$

Οι δύο εξισώσεις παραπάνω περιγράφουν τη δυναμική του Αλγορίθμου Πολλαπλασιαστι-

κής Ενημέρωσης Βαρών (Multiplicative Weights Update - MWU). Πιο διαισθητικά, η βασική ιδέα είναι ότι κάθε παίκτης αθροίζει τα σωρευτικά κέρδη κάθε μιας από τις ενέργειές του και στη συνέχεια εφαρμόζει μια καθαρή στρατηγική $\alpha_i \in \mathcal{A}_i$ με πιθανότητα χονδρικά ανάλογη προς αυτά τα σκορ απόδοσης. Εστιάζοντας στην παραλλαγή “ ε -HEDGE” [CHM17], η διαδικασία αυτή μπορεί να περιγραφεί σε ψευδοκώδικα ως εξής:

Αλγόριθμος 1 (ε -HEDGE)

Απαιτείται: ακολουθία βημάτων $\gamma_n > 0$, παράγοντας εξερεύνησης $\varepsilon \in [0, 1]$, αρχικά σκορ $Y_i \in \mathbb{R}^{\mathcal{A}_i}$.

Για $n = 1, 2, \dots$:

Για κάθε παίκτη $i \in \mathcal{N}$:

 όρισε μεικτή στρατηγική: $X_i \leftarrow \varepsilon \text{unif}_i + (1 - \varepsilon)\Lambda_i(Y_i)$;

 επίλεξε ενέργεια $\alpha_i \sim X_i$;

 απόκτησε εκτίμηση $\hat{v}_i$ του διανύσματος αποδόσεων $v_i(\alpha_i; \alpha_{-i})$;

 ενημέρωσε τα σκορ: $Y_i \leftarrow Y_i + \gamma_n \hat{v}_i$;

Τέλος Επανάληψης

Τέλος Επανάληψης

Μαθηματικά, ο παραπάνω αλγόριθμος αναπαριστά την αναδρομή

$$X_i(n) = \varepsilon \text{unif}_i + (1 - \varepsilon)\Lambda_i(Y_i(n)),$$

$$Y_i(n+1) = Y_i(n) + \gamma_{n+1} \hat{v}_i(n+1),$$

όπου

$$\text{unif}_i = \frac{1}{|\mathcal{A}_i|} (1, \dots, 1)$$

δηλώνει την ομοιόμορφη κατανομή πάνω στο $\mathcal{A}_i$ και η $\Lambda_i : \mathbb{R}^{\mathcal{A}_i} \rightarrow \mathcal{X}_i$ είναι ο χάρτης πιθανοτικής επιλογής

$$\Lambda_i(y_i) = \frac{(\exp(y_i \alpha_i))_{\alpha_i \in \mathcal{A}_i}}{\sum_{\alpha_i \in \mathcal{A}_i} \exp(y_i \alpha_i)},$$

ο οποίος αποδίδει εκθετικά μεγαλύτερη πιθανότητα σε καθαρές στρατηγικές με υψηλότερα σκορ. Έτσι, οι πιθανότητες επιλογής ενέργειας στο ε -Hedge αποτελούν κυρτό συνδυασμό της ομοιόμορφης εξερεύνησης (με συνολικό βάρος ε) και της εκθετικής στάθμισης (με συνολικό βάρος $1 - \varepsilon$). Ως αποτέλεσμα, για $\varepsilon \approx 1$, η επιλογή ενέργειας είναι ουσιαστικά ομοιόμορφη. Αντιθέτως, όταν $\varepsilon = 0$, ανακτούμε τον κλασικό αλγόριθμο Hedge με ακολουθία ανάδρασης $\hat{v}(n)$ και χωρίς ρητή εξερεύνηση.

2.2.5 Σύγκλιση για πάιγνια δυναμικού

Σύμφωνα με το [CHM17], διακρίνουμε δύο περιπτώσεις σύγκλισης, ανάλογα με τη μορφή της ανάδρασης της απόδοσης.

Ημιανατροφοδοτούμενη Περίπτωση

Στην ημιανατροφοδοτούμενη περίπτωση (semi-bandit case), κάθε παίκτης έχει πρόσβαση σε μια (πιθανώς θορυβώδη) εκτίμηση του συνολικού διανύσματος αποδόσεων στο γύρο n , δηλαδή μια εκτίμηση για την απόδοση που θα απέφερε η επιλογή οποιασδήποτε πιθανής δράσης στον τρέχοντα γύρο. Θεωρούμε ότι η ακολουθία ανατροφοδότησης $\hat{v}_i(n)$ προς τον Αλγόριθμο ε -Hedge είναι της γενικής μορφής

$$\hat{v}_i(n) = v_i(\alpha_i(n); \alpha_{-i}(n)) + \xi_i(n),$$

όπου η $(\xi_i(n))_{i \in \mathcal{N}}$ είναι μια στοχαστική διαδικασία θορύβου που αναπαριστά το σφάλμα εκτίμησης των παικτών και ικανοποιεί τις ακόλουθες στατιστικές υποθέσεις:

1. Μηδενική μέση τιμή:

$$\mathbb{E}[\xi_i(n) \mid \mathcal{F}_{n-1}] = 0 \quad \text{για όλα τα } n = 1, 2, \dots \quad (\text{σ.σ.}).$$

όπου $\mathcal{F}_{n-1}$ το ιστορικό του παιχνιδιού. Η ακολουθία ανατροφοδότησης $\hat{v}(n)$, δηλαδή, των παικτών είναι υπό συνθήκη αμερόληπτη σε σχέση με το ιστορικό του παιχνιδιού, δηλ.

$$\mathbb{E}[\hat{v}_i(n) \mid \mathcal{F}_{n-1}] = v_i(X(n-1)), \quad \text{για όλα τα } n = 1, 2, \dots \quad (\text{σ.σ.}).$$

2. Ήπιες ουρές κατανομής:

$$\mathbb{P}(\|\xi_i(n)\|_\infty^2 \geq z \mid \mathcal{F}_{n-1}) \leq \frac{A}{z^q} \quad \text{για κάποιο } q > 2, A > 0, \text{ και όλα τα } n = 1, 2, \dots \quad (\text{σ.σ.}).$$

Επομένως, η διασπορά του εκτιμητή $\hat{v}$ είναι υπό συνθήκη φραγμένη, δηλ.

$$\text{Var}[\hat{v}_i(n) \mid \mathcal{F}_{n-1}] \leq \sigma^2 \quad \text{για όλα τα } n = 1, 2, \dots \quad (\text{σ.σ.}).$$

Θεώρημα. Έστω Γ ένα παίγνιο δυναμικού και υποθέτουμε ότι ο Αλγόριθμος ε -Hedge εκτελείται με i) ημιανατροφοδοτούμενη πληροφόρηση που ικανοποιεί τις δύο υποθέσεις παραπάνω, ii) έναν μη αρνητικό παράγοντα εξερεύνησης $\varepsilon \geq 0$, και iii) μια ακολουθία βημάτων της μορφής $\gamma_n \propto 1/n^\beta$ για κάποιο $\beta \in (1/q, 1]$. Τότε η ακολουθία $X(n)$ συγκλίνει (σ.σ.) σε μια δ -ισορροπία του Γ με $\delta \equiv \delta(\varepsilon) \rightarrow 0$ καθώς $\varepsilon \rightarrow 0$.

Πόρισμα. Αν ο Αλγόριθμος ε -Hedge εκτελείται με τις ίδιες υποθέσεις και χωρίς εξερεύνηση ($\varepsilon = 0$), τότε το $X(n)$ συγκλίνει σε ισορροπία Nash με πιθανότητα 1.

Μη ανατροφοδοτούμενη περίπτωση (Bandit Case)

Στη μη ανατροφοδοτούμενη περίπτωση, οι παίκτες γνωρίζουν την απόδοση μόνο της ενέργειας που επέλεξαν, και όχι των αποδόσεων που θα είχαν λάβει εάν είχαν επιλέξει τις υπόλοιπες εναλλακτικές ενέργειες:

$$\hat{u}_i(n) = u_i(\alpha_i(n); \alpha_{-i}(n)).$$

Σε αυτή την περίπτωση, δεδομένου ότι οι παίκτες δεν έχουν καμία πληροφορία για τις αποδόσεις των ενεργειών που δεν επέλεξαν, κατασκευάζουν έναν εκτιμητή για αυτές (bandit estimator):

$$\hat{v}_{i\alpha_i}(n) = \frac{1(\alpha_i(n) = \alpha_i)}{\mathbb{P}(\alpha_i(n) = \alpha_i \mid \mathcal{F}_{n-1})} \cdot \hat{u}_i(n) = \begin{cases} \hat{u}_i(n)/X_{i\alpha_i}(n-1) & \text{αν } \alpha_i = \alpha_i(n), \\ 0 & \text{διαφορετικά.} \end{cases}$$

Θεώρημα 2.3. Έστω Γ ένα παίγνιο δυναμικού και υποθέτουμε ότι εκτελείται ο Αλγόριθμος ε -Hedge με *i)* τον παραπάνω εκτιμητή, *ii)* θετικό συντελεστή εξερεύνησης $\varepsilon > 0$, και *iii)* ακολουθία βημάτων της μορφής $\gamma_n \propto 1/n^\beta$ για κάποιο $\beta \in (0, 1]$. Τότε η ακολουθία $X(n)$ συγκλίνει (σ.σ.) σε δ -ισορροπία του Γ , όπου $\delta \equiv \delta(\varepsilon) \rightarrow 0$ καθώς $\varepsilon \rightarrow 0$.

Το παραπάνω θεώρημα καταδεικνύει ότι το παίγνιο συγκλίνει πλησιέστερα στην ισορροπία Nash του παιχνιδιού όταν ο συντελεστής εξερεύνησης ε λαμβάνει όσο το δυνατόν μικρότερες τιμές. Βέβαια, δεν επιτρέπεται να μηδενιστεί πλήρως ($\varepsilon = 0$) ώστε να οδηγούμασταν σε αμιγή ισορροπία Nash. Ο λόγος για αυτό είναι ότι η διασπορά του εκτιμητή $\hat{v}(n)$ μπορεί να αυξάνεται απεριόριστα αν οι πιθανότητες επιλογής ενεργειών γίνουν αυθαίρετα μικρές.

Μοντελοποίηση Συστήματος

3.1 Δορυφορικό Δίκτυο

Η Δορυφορική Υποδομή Υπολογιστικής Παρυφής που θα μελετηθεί αποτελείται από συνολικά n δορυφόρους, διατεταγμένους σε έναν Αστερισμό Χαμηλής Τροχιάς (Low Earth Orbit - LEO).

Η φυσική μοντελοποίηση είναι βασισμένη στην αντίστοιχη μοντελοποίηση του [Ton+23]. Οι διαδορυφορικές ζεύξεις (Inter-Satellite Links - ISLs) μεταξύ ενός δορυφόρου και των γειτονικών του δορυφόρων χρησιμοποιούν τη ζώνη Ka. Λαμβάνονται υπόψη οι Απώλειες Διάδοσης Ελεύθερου Χώρου (Free-Space Path Loss - FSPL) και ο Προσθετικός Λευκός Γκαουσιανός Θόρυβος (Additive White Gaussian Noise - AWGN). Ο επιτεύξιμος ρυθμός μετάδοσης της ζεύξης μεταξύ του εξεταζόμενου δορυφόρου (δορυφόρος 1) και του γειτονικού του δορυφόρου j , δίνεται από τη σχέση:

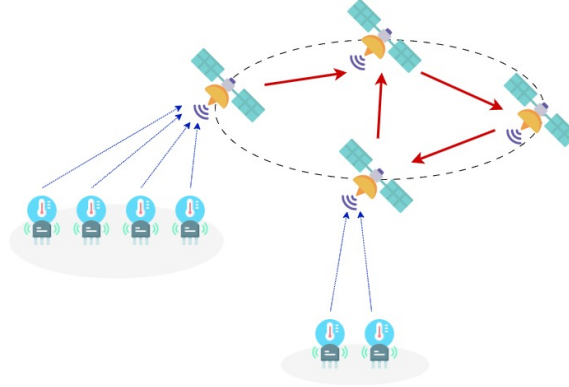
$$r_{1,j} = B_{1,j} \log_2 \left(1 + \frac{PGH_{1,j}}{N_{01,j}} \right), \quad (3.1)$$

όπου $B_{1,j}$ είναι το εύρος ζώνης του υποδιαύλου, P είναι η ισχύς μετάδοσης του εξεταζόμενου δορυφόρου, G είναι το κέρδος κεραίας του εξεταζόμενου δορυφόρου, $H_{1,j}$ είναι το κέρδος καναλιού της ζεύξης, $N_{01,j}$ είναι η ισχύς θορύβου της ζεύξης. Έχουμε $G \approx (4D_{TA}f/c)^2$, και D_{TA} είναι η διάμετρος της κεραίας εκπομπής του εξεταζόμενου δορυφόρου. Επίσης έχουμε $H_{1,j} = (h_{1,j})^2$, όπου ο συντελεστής καναλιού ισούται με $h_{1,j} = 1/\sqrt{L_{1,j}^f}$. Ακόμα, $L_{1,j}^f$ είναι οι απώλειες διάδοσης ελεύθερου χώρου, $L_{1,j}^f = (4\pi l_{1,j}f/c)^2$, και $l_{1,j}$ είναι η απόσταση μεταξύ του εξεταζόμενου δορυφόρου και του γειτονικού δορυφόρου j (σε μέτρα).

Συνήθως, κάθε δορυφόρος έχει τέσσερις γειτονικούς δορυφόρους, δύο στην ίδια τροχιά και δύο στις αριστερές και δεξιές τροχιές. Η απόσταση μεταξύ του εξεταζόμενου δορυφόρου και του γειτονικού δορυφόρου στην ίδια τροχιά είναι $l_{1,j} = \sqrt{2R} \sqrt{1 - \cos(\frac{360^\circ}{J})}$, όπου R είναι η ακτίνα της τροχιάς, και J είναι ο αριθμός δορυφόρων σε κάθε τροχιά. Η απόσταση μεταξύ του τοπικού δορυφόρου και του γειτονικού δορυφόρου σε διαφορετικές τροχιές είναι $l_{1,j} = \omega \cos(lat)$, όπου $\omega = \sqrt{2R} \sqrt{1 - \cos(\frac{360^\circ}{2I})}$, lat είναι το γεωγραφικό πλάτος στο οποίο βρίσκεται η διαδορυφορική ζεύξη του εξεταζόμενου δορυφόρου, και I είναι ο αριθμός των τροχιών του αστερι-

σμού δορυφόρων.

Στο παρόν πρόβλημα επικεντρωνόμαστε στο μέσο ρυθμό μετάδοσης μεταξύ κάθε ζεύγους δορυφόρων i, j , ο οποίος θα συμβολιστεί ως $R_{i,j}$.



Σχήμα 3.1: Δορυφορική Υπολογιστική Υποδομή Παρυφής

3.2 Υπολογιστικό Σύστημα

Κάθε δορυφόρος στο σύστημα εξυπηρετεί μια περιοχή επίγειων χρηστών Διαδικτύου των Πραγμάτων (Internet of Things - IoT), οι οποίοι εκφορτώνουν διεργασίες προς εκτέλεση σε αυτόν (βλ. ενδεικτικά Σχήμα 3.1). Οι διεργασίες αποθηκεύονται στην ουρά του εξυπηρετητή του δορυφόρου και στη συνέχεια εκτελούνται ή εκφορτώνονται με τη σειρά τους σε άλλους δορυφόρους.

Οι χρήστες χαρακτηρίζονται από δυναμικότητα. Μπορεί να διαθέτουν παραδείγματος χάρη πυκνή γεωγραφική συγκέντρωση ή υψηλό ρυθμό παραγωγής νέων διεργασιών προς εκτέλεση. Αυτό μεταφράζεται στο δορυφόρο σε υψηλό ρυθμό ανάθεσης νέων διεργασιών. Αντίστοιχα, δυναμικότητα χαρακτηρίζει και το μέγεθος αυτών των διεργασιών, τόσο ως προς το πλήθος Bytes όσο και ως προς τον υπολογιστικό φόρτο που προσθέτουν στις υπολογιστικές μονάδες. Προκειμένου να συστηματοποιήσουμε αυτές τις παραμέτρους, εισάγουμε τις εξής μεταβλητές:

- t : Γύρος εκφόρτωσης. Χωρίζουμε τη συνεχή διαδικασία παραγωγής διεργασιών και εκφόρτωσης σε διακριτές περιόδους.
- a_i^t : πλήθος νέων διεργασιών που ανατίθενται στο δορυφόρο i από τους IoT χρήστες στο γύρο t . Υποθέτουμε ότι το πλήθος αυτό ακολουθεί κατανομή Poisson με παράμετρο λ_i για κάθε δορυφόρο.
- d_i [Bytes]: μέσο μέγεθος των διεργασιών που ανατίθενται στο δορυφόρο i από τους IoT χρήστες. Υποθέτουμε ότι το μέγεθος αυτό ακολουθεί ομοιόμορφη κατανομή μεταξύ 0 και $2 * d_i$.

- I [Κύκλοι/Byte]: μέσος υπολογιστικός φόρτος που προστίθεται στο σύστημα ανά Byte διεργασίας (αναλογία μεγέθους-φόρτου).
- Q [Κύκλοι]: μέγιστος υπολογιστικός φόρτος που μπορεί να αποθηκευτεί στην ουρά των εξυπηρετητών των δορυφόρων.
- l_i^t [Κύκλοι]: συνολικός υπολογιστικός φόρτος που απομένει στην ουρά του δορυφόρου i προς εκτέλεση στην αρχή του γύρου t .
- $l_{i,new}^t$ [Κύκλοι]: υπολογιστικός φόρτος που προστίθεται στην ουρά του δορυφόρου i προς εκτέλεση στο γύρο t . Προφανώς πρέπει, προστιθέμενος στην ουρά του εξυπηρετητή του δορυφόρου, να μην υπερβαίνει τη χωρητικότητα της ουράς, δηλαδή $0 \leq l_i^t + l_{i,new}^t \leq Q$.
- ϕ_i [Κύκλοι/Γύρο]: υπολογιστική ισχύς του εξυπηρετητή του δορυφόρου i .

Βέβαια, εκτός από το στατιστικό προφίλ παραγωγής διεργασιών από τους IoT χρήστες και εκφόρτωσης σε δορυφόρους, έχουμε και την διαδορυφορική εκφόρτωση διεργασιών που χαρακτηρίζεται από την ακόλουθη μεταβλητή:

- $b_{i,j}^t$ [Κύκλοι]: υπολογιστικός φόρτος διεργασιών που εκφορτώνεται από το δορυφόρο i στο δορυφόρο j στο γύρο t .

Τελικά, σε κάθε γύρο και για κάθε δορυφόρο i έχουμε τα εξής γεγονότα:

- Η ουρά του εξυπηρετητή περιέχει φόρτο διεργασιών l_i^t .
- Παράγονται από τους IoT χρήστες και προστίθενται στην ουρά a_i^t νέες διεργασίες συνολικού φόρτου $l_{i,new}^t$.
- Αφαιρείται από την ουρά φόρτος $\sum_{j=1, j \neq i}^N b_{i,j}^t$ ο οποίος εκφορτώνεται σε άλλους δορυφόρους.
- Προστίθεται στην ουρά φόρτος $\sum_{j=1, j \neq i}^N b_{j,i}^t$ από άλλους δορυφόρους.

Επομένως, στο τέλος του γύρου t , ο συνολικός χρόνος που χρειάζεται ο δορυφόρος i για να εκτελέσει όλες τις διεργασίες που του έχουν ανατεθεί ισούται με:

$$T_i = \frac{\sum_{j=1, j \neq i}^N (b_{j,i}^t) + (l_i^t + l_{i,new}^t - \sum_{j=1, j \neq i}^N b_{i,j}^t)}{\phi_i}$$

3.3 Στρατηγική εκφόρτωσης υπολογισμών

Δεδομένου ότι στόχος είναι η εξισορρόπηση του φορτίου στις ουρές των εξυπηρετητών των δορυφόρων, πρέπει να προσδιοριστούν οι κατάλληλες τιμές των μεταβλητών $b_{i,j}^t$ για όλους

τους δορυφόρους και για όλους τους γύρους, ώστε τελικά η τυπική απόκλιση του φόρτου l_i^t στις ουρές να ελαχιστοποιηθεί. Πρέπει να σημειωθεί ότι παρά τη διαφορετική υπολογιστική ικανότητα των δορυφόρων, απαιτούμε οι ουρές τους να διαθέτουν όσο το δυνατόν παρόμοιο φόρτο, διότι όλοι οι δορυφόροι χαρακτηρίζονται από την ίδια μεταβλητή μέγιστου αποθηκεύσιμου φόρτου Q .

Ο προσδιορισμός των μεταβλητών αυτών θα γίνει δυναμικά, με μια διαδικασία εκμάθησης. Αυτό που μπορεί να διαφέρει όμως, είναι το πλήθος των πληροφοριών που συλλέγονται για το σύστημα κατά τη διαδικασία αυτή, και η κεντροποίηση ή μη της διαδικασίας λήψης αποφάσεων.

Πλήρως κεντροποιημένη μέθοδος

Ένας πλήρως κεντροποιημένος ορχηστρωτής λαμβάνει πληροφορίες από όλους τους συμμετέχοντες δορυφόρους σχετικά με το τρέχον φορτίο στις ουρές των εξυπηρετητών τους και προτείνει επαναληπτικά στρατηγικές εκφόρτωσης σε κάθε έναν. Ωστόσο, αυτό συνεπάγεται υψηλό κόστος α) για τις ενεργειακές απαιτήσεις του ορχηστρωτή, β) για τη συνεχή επικοινωνία μεταξύ όλων των δορυφόρων και του ορχηστρωτή, καθώς και γ) εγείρει σοβαρούς περιορισμούς σε περίπτωση αποτυχίας του ορχηστρωτή, ειδικά εάν έχουμε περιορισμένη ανοχή σε σφάλματα. Αυτές οι σκέψεις συνηγορούν υπέρ της χρήσης μιας αποκεντρωμένης προσέγγισης.

Μερικώς αποκεντρωμένη μέθοδος

Σε προηγούμενη βιβλιογραφία ([Zha+21]) συναντούμε μεθόδους στις οποίες κάθε δορυφόρος συλλέγει πληροφορίες για όλο το σύστημα και λαμβάνει αποφάσεις για τη στρατηγική του με κατανεμημένο τρόπο. Αυτή η προσέγγιση αντιμετωπίζει τα προβλήματα (α) και (γ), όπως περιγράφηκαν παραπάνω, αλλά επιδεινώνει περαιτέρω το πρόβλημα (β), καθώς για τη συλλογή των πληροφοριών οι δορυφόροι πρέπει να επικοινωνούν συνεχώς μεταξύ τους, και όχι με έναν μεμονωμένο ορχηστρωτή, οδηγώντας έτσι σε μια ανταλλαγή πακέτων τύπου "πλημμύρας" στο σύστημα.

Πλήρως αποκεντρωμένη μέθοδος

Στην παρούσα εργασία, θα αναλυθεί μια πλήρως αποκεντρωμένη μοντελοποίηση, στην οποία κάθε δορυφόρος-πομπός ορίζει έναν μοναδικό δορυφόρο-δέκτη σε κάθε γύρο, στον οποίο και εκφορτώνει. Κατά την εκφόρτωση λαμβάνει πληροφορίες για το φορτίο του δορυφόρου-δέκτη, επομένως η πληροφορία και η απόφαση παραμένει κατανεμημένη, αλλά η απόκτηση μιας καθολικής εικόνας για το σύστημα (δηλαδή για τα φορτία όλων των δορυφόρων) επιτυγχάνεται μέσω 1-προς-1 αλληλεπιδράσεων με διαφορετικούς δορυφόρους με την πάροδο του χρόνου.

Καταρχάς, η παρούσα μοντελοποίηση βασίζεται σε μια στάσιμη κατανομή πλήθους και μεγέθους νέων παραγόμενων διεργασιών. Σε περίπτωση αλλαγής του στατιστικού προφίλ του συστήματος, η μοντελοποίηση πραγματοποιείται εκ νέου. Υπό αυτό το πρίσμα, η στρατηγική εκφόρτωσης $b_{i,j}^t$ μπορεί και πρέπει να παραμένει σταθερή στη διάρκεια του χρόνου, καθώς αυτό

για το οποίο ενδιαφερόμαστε είναι η μέση (μακροχρόνια) συμπεριφορά του συστήματος όσον αφορά στην πρόθεση για εξισορρόπηση φορτίου. Με άλλα λόγια, στάσιμη κατανομή διεργασιών συνεπάγεται και στατική στρατηγική εξισορρόπησης φορτίου, εάν ο στόχος είναι η μακροπρόθεσμη και όχι η βραχυπρόθεσμη εξισορρόπηση. Επομένως, η στρατηγική από $b_{i,j}^t$ εκφυλίζεται σε $b_{i,j}$.

Η παραδοχή ότι σε κάθε γύρο επιλέγεται ένας μόνο δορυφόρος ως προορισμός εκφόρτωσης γίνεται χωρίς βλάβη της γενικότητας. Πράγματι, έστω ότι η βέλτιστη στρατηγική για το δορυφόρο i έγκειται στην εκφόρτωση φορτίου $b_{i,1}^{OPT}, \dots, b_{i,i-1}^{OPT}, b_{i,i+1}^{OPT}, \dots, b_{i,n}^{OPT}$ προς τους δορυφόρους $1, \dots, i-1, i+1, \dots, n$ αντίστοιχα σε κάθε γύρο. Το μέσο φορτίο που εκφορτώνεται από τον i στο j σε κάθε γύρο ταυτίζεται προφανώς με την τιμή $b_{i,j}^{OPT}$. Ορίζουμε τη στρατηγική κατά την οποία ο δορυφόρος i επιλέγει το δορυφόρο j με πιθανότητα $p_{i,j}, j \neq i, \sum_{j \neq i} p_{i,j} = 1$ και εκφορτώνει σε αυτόν φορτίο $b'_{i,j} = \frac{b_{i,j}^{OPT}}{p_{i,j}}$. Τότε, το μέσο φορτίο που εκφορτώνεται από τον i στο j σε κάθε γύρο ισούται και πάλι με $p_{i,j} * b'_{i,j} = b_{i,j}^{OPT}$. Επομένως, υπάρχει βέλτιστη στρατηγική (ως προς την αναμενόμενη συμπεριφορά) στην οποία επιλέγεται μόνο ένας δορυφόρος-στόχος κάθε φορά.

Αξίζει να παρατηρήσουμε ότι στην προαναφερθείσα στρατηγική, οι πιθανότητες $p_{i,j}$ μπορούν να οριστούν αυθαίρετα, και απλώς να επιλεγθούν κατάλληλα οι τιμές $b_{i,j}$. Στην πραγματικότητα, το πρόβλημα είναι πιο σύνθετο. Οι τιμές των $b_{i,j}$ φράσσονται από ένα όριο (ας συμβολιστεί ως c_i), το οποίο οφείλεται σε χρονικούς και ενεργειακούς περιορισμούς: δε μπορεί να εκφορτώνεται ένα αυθαίρετα μεγάλο φορτίο μεταξύ δύο δορυφόρων στη μονάδα του χρόνου. Το πρόβλημα αυτό θα λυνόταν (ή, τουλάχιστον, θα μειωνόταν) εάν επιτρέπαμε πολλαπλούς δορυφόρους-στόχους σε κάθε γύρο, όμως, όπως προαναφέρθηκε, η ιδέα αυτή απορρίπτεται στην παρούσα εργασία, καθώς ο στόχος είναι να μην επιβαρυνθεί το σύστημα με πακέτα από όλους προς όλους τύπου πλημμύρας.

Δεδομένων λοιπόν των περιορισμών c_i , οι πιθανότητες δεν πρέπει πλέον να οριστούν αυθαίρετα. Από την άλλη, η αυθαίρετη επιλογή των $b_{i,j}$ δεν είναι απαραίτητο ότι θα μπορεί να διορθωθεί με κατάλληλη επιλογή των $p_{i,j}$, και πάλι λόγω των περιορισμών c_i . Έτσι, οι μεταβλητές απόφασης διακρίνονται (α) στις μεταβλητές $p_{i,j}, j \neq i, 1 \leq i, j \leq N$ και (β) στις μεταβλητές $b_{i,j}, j \neq i, 1 \leq i, j \leq N$. Για κάθε δορυφόρο, ο κοινός αυτός χώρος απόφασης (πιθανότητες+ποσά) κλιμακώνεται εκθετικά, ακόμα και αν υποθέσουμε διακριτές πιθανές τιμές. Είναι λοιπόν φανερό ότι η στρατηγική πρέπει να αποσυνζευκτεί σε δύο ξεχωριστές στρατηγικές: τη στρατηγική επιλογής του στόχου εκφόρτωσης (δηλαδή τις μεταβλητές $p_{i,j}$) και τη στρατηγική ποσού εκφόρτωσης (δηλαδή τις μεταβλητές $b_{i,j}$).

3.3.1 Στρατηγική ποσού εκφόρτωσης

Ας ξεκινήσουμε από τον καθορισμό του ποσού εκφόρτωσης. Υποθέτουμε δηλαδή ότι ο δορυφόρος i (με εναπομείναν φορτίο l_i) έχει επιλέξει στο γύρο t ως προορισμό εκφόρτωσης το δορυφόρο j (με εναπομείναν φορτίο l_j), και απομένει ο προσδιορισμός της μεταβλητής $b_{i,j}$.

Ο στόχος είναι η εξισορρόπηση φορτίου, δηλαδή η μετάβαση σε μια κατάσταση ισορροπίας στην οποία το εναπομείναν φορτίο στις ουρές των εξυπηρετητών των δορυφόρων είναι ίσο.

Μια πρώτη ιδέα, λοιπόν, είναι να τεθεί το ποσό εκφόρτωσης ίσο με το ποσό που χρειάζεται προκειμένου το l_i να εξισωθεί με το μέσο όρο των φορτίων στο σύστημα, δηλαδή να εκφορτωθεί ποσό $l_i - \sum_i \frac{l_i}{N}$.

Η γνώση όμως του μέσου φορτίου στο σύστημα θα απαιτούσε την επικοινωνία μεταξύ όλων των δορυφόρων, οπότε θα οδηγούσε και πάλι σε πακέτα τύπου πλημμύρας. Επομένως, είναι απαραίτητη μια άλλη στρατηγική προδιορισμού του ποσού.

Ας εξετάσουμε το σενάριο στο οποίο υπάρχει στρατηγική ισορροπίας με 1-προς-1 στοχαστικές εκφορτώσεις, όπως περιγράφηκε παραπάνω, για την οποία μάλιστα έχουν βρεθεί οι κατάλληλες τιμές $p_{i,j}$ για όλους τους δορυφόρους, καθώς και οι κατάλληλες τιμές $b_{i,j}$ για όλους τους δορυφόρους εκτός του δορυφόρου k . Τότε, έστω ότι όλοι οι δορυφόροι διαθέτουν ίσα αρχικά φορτία l_{eq} στις ουρές τους, και έστω το προστιθέμενο φορτίο $\Delta l_m^t = l_A^t$ για το γύρο t στην ουρά ενός δορυφόρου $m, m \neq k$ και $\Delta l_k^t = l_B^t \neq l_A^t$ το αντίστοιχο φορτίο για το δορυφόρο k . Κατά την επιτυχή εξισορρόπηση φορτίου θα έπρεπε τα προστιθέμενα φορτία να είναι ίσα, άρα η τιμή $b_{k,m}^t$ πρέπει να ισούται με $\frac{l_B^t - l_A^t}{2} = \frac{(l_B^t + l_{eq}) - (l_A^t + l_{eq})}{2} = \frac{l_m^t - l_k^t}{2}$. Τελικά, για δορυφόρους που ξεκινούν με ίσα φορτία στις ουρές τους, εάν υπάρχει κατάλληλη στρατηγική εξισορρόπηση φορτίου με 1-προς-1 στοχαστικές εκφορτώσεις, αυτή θα βρίσκεται θέτοντας τις τιμές $b_{i,j}^t$ ίσες με $\frac{l_i^t - l_j^t}{2}$. Βέβαια, η εκφόρτωση αυτής της ποσότητας μπορεί να μην είναι πάντα δυνατή:

1. Αν $l_i^t \leq l_j^t$, η ποσότητα εκφόρτωσης είναι αρνητική (αδύνατο στην πραγματικότητα).
2. Δεδομένου ότι εκφορτώνουμε ακέραιο αριθμό εργασιών, ενδέχεται να μην καταφέρουμε να εκφορτώσουμε ακριβώς το φορτίο $\frac{l_i^t - l_j^t}{2}$. Αντίθετα, εκφορτώνουμε συνεχώς εργασίες μέχρι να φτάσουμε όσο το δυνατόν πιο κοντά στη ζητούμενη κλασματική τιμή.
3. Οι δορυφόροι έχουν περιορισμένες δυνατότητες επικοινωνίας. Όπως αναφέρθηκε και προηγουμένως, το εκφορτώμενο φορτίο σε ένα γύρο περιορίζεται από μια μεταβλητή c_i , η οποία είναι συμβατή με τις δυνατότητες του δορυφόρου i . Σημειώνεται ότι το c_i πρέπει να είναι αρκετά μεγάλο ώστε να εξασφαλίζεται σταθερή ουρά, δηλαδή ο κάθε δορυφόρος πρέπει να έχει τη δυνατότητα να εκφορτώνει τουλάχιστον το ποσό που χρειάζεται ώστε ο εξυπηρετητής του να μην υπερφορτώνεται.

Επομένως, το φορτίο που εκφορτώνεται από το δορυφόρο i στο δορυφόρο j σε ένα γύρο t προσεγγίζεται ως εξής:

$$b_{i,j}(t) \approx \max(\min(\frac{l_i^t - l_j^t}{2}, c_i), 0)$$

Η παραπάνω σχέση αποτελεί προσέγγιση και όχι ισότητα, ακριβώς λόγω του επιχειρήματος (2) παραπάνω. Για να είναι λοιπόν η ποσότητα $b_{i,j}$ όσο το δυνατόν πλησιέστερη στην τιμή $\frac{l_i^t - l_j^t}{2}$ που εξασφαλίζει την εξισορρόπηση, τα φορτία l_i^t, l_j^t πρέπει να διαθέτουν κατάλληλες τιμές. Δε διαθέτουμε έλεγχο των φορτίων, ωστόσο διαθέτουμε έλεγχο των πιθανοτήτων $p_{i,j}$, οι οποίες πρέπει να ρυθμιστούν κατάλληλα, ώστε εκφορτώσεις που συνδέονται με μεγάλα φορτία να πραγματοποιούνται συχνότερα και η απαιτούμενη ποσότητα προς εκφόρτωση να ελαττώνεται (με σκοπό να βρίσκεται εντός των επιτρεπόμενων ορίων).

3.3.2 Στρατηγική στόχου εκφόρτωσης

Στην προηγούμενη ενότητα αναλύθηκε η διαδικασία προσδιορισμού του ποσού εκφόρτωσης αφού έχει επιλεγθεί ο δορυφόρος-στόχος. Για τις πιθανότητες επιλογής ενός δορυφόρου ως στόχο σε κάθε γύρο, το πρόβλημα βελτιστοποίησης, λοιπόν, μπορεί να διατυπωθεί ως εξής:

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbb{E}[\sigma(l_1, l_2, \dots, l_N)] \\ \text{s.t.} \quad & \sum_{j=1, j \neq i}^N p_{i,j} = 1, \quad \forall i \in \{1, \dots, N\} \end{aligned}$$

όπου με σ συμβολίζεται η τυπική απόκλιση, και με l_i το αναμενόμενο φορτίο ουράς σε κατάσταση ισορροπίας για τον δορυφόρο i , το οποίο προκύπτει από τις στάσιμες διαδικασίες άφιξης νέων διεργασιών και διαδορυφορικής εκφόρτωσης.

Η διαδικασία απόφασης των κατάλληλων τιμών $p_{i,j}$ είναι σύνθετη, λόγω της (άγνωστης) δυναμικότητας των φορτίων στις ουρές. Η παιγνιοθεωρητική προσέγγιση επίλυσης εδώ παρουσιάζει σημαντικά πλεονεκτήματα: είναι εγγενώς κατανεμημένη, αποφεύγοντας μοναδικά σημεία αστοχίας (single points of failure) και λαμβάνοντας αποφάσεις που βασίζονται αποκλειστικά σε τοπικές πληροφορίες. Παράλληλα, προσφέρει θεωρητικές εγγυήσεις σύγκλισης σε σημεία ισορροπίας, ενώ οι αλγόριθμοι επίλυσης των παιγνίων χαρακτηρίζονται από χαμηλή υπολογιστική πολυπλοκότητα. Αυτός ο συνδυασμός αξιοπιστίας, επεκτασιμότητας και υπολογιστικής αποδοτικότητας την καθιστά ιδανική για αποκεντρωμένα συστήματα όπως το μελετούμενο. Η ακριβής διατύπωση του προβλήματος απόφασης ως παίγνιο αναλύεται στο επόμενο κεφάλαιο.

Παιγνιο-θεωρητική μοντελοποίηση

4.1 Γενική μοντελοποίηση του προβλήματος απόφασης

Το πρόβλημα απόφασης των μεταβλητών $p_{i,j}$ μπορεί να αντιμετωπιστεί ως ένα παίγνιο μεταξύ δορυφόρων, συγκεκριμένα επαναλαμβανόμενο, διότι διεξάγεται σε διακριτούς γύρους εκφόρτωσης.

Κατά τη μετάβαση από ένα γύρο στον επόμενο, η κατάσταση του παιχνιδιού εν γένει μπορεί να θεωρηθεί ότι μεταβάλλεται, καθώς οι αφίξεις νέων διεργασιών και οι εκφορτώσεις αλλάζουν τα περιοχόμενα των ουρών των εξυπηρετητών των δορυφόρων. Η πλήρης συμπερίληψη της δυναμικότητας του περιβάλλοντος γίνεται δυνατή μόνο στην περίπτωση της μοντελοποίησης του παιγνίου ως στοχαστικό παίγνιο. Ακολουθώντας τις συμβάσεις και τους συμβολισμούς του κεφαλαίου 2, έχουμε:

$$\Gamma = \langle I, S, (A^i(s))_{s \in S}^{i \in I}, q, r \rangle, \text{ όπου}$$

- $I = \{1, 2, \dots, n\}$ είναι το σύνολο των δορυφόρων.
- S είναι το σύνολο καταστάσεων. Κάθε κατάσταση αφορά συγκεκριμένα εναπομείναντα φορτία προς επεξεργασία στις ουρές των δορυφόρων, δηλαδή κάθε κατάσταση προσδιορίζεται από το διάνυσμα $\mathbf{I}^t = [l_1^t, l_2^t, \dots, l_n^t]$, $\mathbf{I}^t \in S$, των τρέχοντων φορτίων στις ουρές.
- Για κάθε $i \in I$ και κάθε κατάσταση $s \in S$, $A^i(s)$ είναι ένα πεπερασμένο σύνολο ενεργειών διαθέσιμων στον παίκτη i στην κατάσταση s , δηλαδή οι δορυφόροι που μπορεί να επιλέξει προς εκφόρτωση για τον τρέχοντα γύρο. Έτσι, έχουμε $A^i(s) = \{1, \dots, i-1, i+1, \dots, n\}$.
- $q : SA \rightarrow \Delta(S)$ είναι ο κανόνας μετάβασης, ο οποίος ορίζει τα φορτία που θα απομένουν στις ουρές στον επόμενο γύρο. Προφανώς, εξαρτάται από την τρέχουσα κατάσταση στις ουρές (S) και από τους στόχους εκφορτώσεων (A), όπως αυτοί εκφράζονται από τις αντίστοιχες μεταβλητές $b_{i,j}^t, \forall i, j \in I, i \neq j$. Είναι ένας μη ντετερμινιστικός κανόνας, καθώς εξαρτάται επίσης από τις στοχαστικές εκβάσεις των αφίξεων νέων διεργασιών και των αντιστοιχων μεγεθών τους, δηλαδή τις τιμές $l_{i,new}^t, \forall i \in I$. Έτσι, με βάση το κεφάλαιο 3, ο κανόνας αυτός διατυπώνεται ως εξής:

$$l_i^{t+1} = \sum_{j=1, j \neq i}^N (b_{j,i}^t) + (l_i^t + l_{i,new}^t - \sum_{j=1, j \neq i}^N b_{i,j}^t) - \phi_i, \forall i, j \in I$$

- Για κάθε παίκτη $i \in I$, $r^i : SA \rightarrow \mathbf{R}$ είναι η συνάρτηση απόδοσης. Ισοδύναμα θα συμβολιστεί ως $u_{i,j}^t$ στα πλαίσια της παιγνιο-θεωρητικής μοντελοποίησης, όπου j ο δορυφόρος-στόχος που επιλέχθηκε από το δορυφόρο i στο γύρο t . Ανάλογα με το πως θα οριστεί, μπορεί να εξαρτάται από την κατάσταση των ουρών (S), τους στόχους των εκφορτώσεων (A), ή και τη στοχαστικότητα των νέων διεργασιών.

Προφανώς, η μοντελοποίηση αυτή διαθέτει αυξημένη δυναμικότητα και πολυπλοκότητα, που καθιστά δύσκολη την εύρεση εγγυήσεων σύγκλισης. Το στοχαστικό παίγνιο όμως αποτελείται από επιμέρους μη-στοχαστικά παίγνια, στα οποία η συνάρτηση απόδοσης μπορεί να οριστεί κατάλληλα ώστε να περιοριστεί η πολυπλοκότητα αυτή. Για τα επιμέρους αυτά παίγνια ενός γύρου, η επιλογή στόχου εκφόρτωσης των δορυφόρων μπορεί να διατυπωθεί σε στρατηγική μορφή:

$$\mathcal{G} := (\{N\}, \{\mathcal{S}_i\}_{i \in N}, \{u_i\}_{i \in N})$$

όπου:

1. N είναι το σύνολο των παικτών, δηλαδή των δορυφόρων.
2. $\mathcal{S}_i$ είναι το σύνολο των διαθέσιμων στρατηγικών για τον παίκτη i , δηλαδή οι δορυφόροι που μπορεί να επιλέξει προς εκφόρτωση για τον τρέχοντα γύρο. Έτσι, έχουμε $\mathcal{S}_i = \{1, \dots, i-1, i+1, \dots, n\}$
3. u_i είναι η συνάρτηση απόδοσης για τον παίκτη i , η οποία μπορεί να οριστεί με εναλλακτικούς τρόπους.

Στη συνέχεια θα παρουσιαστούν διαφορετικές μοντελοποιήσεις:

- Το παίγνιο G_1 , το οποίο ισοδυναμεί με ένα στοχαστικό παίγνιο, μη περαιτέρω απλοποιήσιμο.
- Το παίγνιο G_2 , το οποίο είναι ισοδύναμο με ένα θορυβώδες παίγνιο δυναμικού G_2^{EXP} . Διακρίνεται σε δύο περιπτώσεις ανάδρασης: ημιανατροφοδοτούμενο (semi-bandit), ή μη ανατροφοδοτούμενο (bandit).

Για τα τρία αυτά παίγνια (G_1 , ημιανατροφοδοτούμενο G_2/G_2^{EXP} , μη ανατροφοδοτούμενο G_2/G_2^{EXP}), τέλος, θα παρουσιαστεί μια επέκτασή τους που λαμβάνει υπόψη την ενεργειακή κατανάλωση των εκφορτώσεων.

4.2 Πάιγνιο G_1 (Στοχαστικό Πάιγνιο)

Όπως αναφέρθηκε πριν, δεδομένων των περιορισμών c_i , δεν είναι δυνατή η εκφόρτωση ενός αυθαίρετα μεγάλου αριθμού διεργασιών σε έναν γύρο. Για το λόγο αυτό, δορυφόροι-στόχοι στους οποίους είναι απαραίτητη η εκφόρτωση ενός μεγάλου φορτίου πρέπει να επιλέγονται πιο συχνά σε σύγκριση με άλλους.

Έτσι, η συνάρτηση απόδοσης για το στατικό παιχνίδι ενός γύρου πρέπει να συνδέεται άμεσα με το εκφορτώμενο ποσό, το οποίο, όπως αναφέρθηκε προηγουμένως, για την αποστολή από το δορυφόρο i στο δορυφόρο j θα έπρεπε να ισούται με $b_{i,j}^t = \frac{l_i^t - l_j^t}{2}$.

Επομένως, η αντίστοιχη συνάρτηση απόδοσης του δορυφόρου i για την επιλογή του δορυφόρου j ως στόχο σε έναν γύρο μπορεί να οριστεί ως:

$$u_{i,j}^t = \frac{l_i^t - l_j^t}{2}$$

Παρατηρούμε ότι η συνάρτηση απόδοσης εξαρτάται από την κατάσταση των ουρών στο γύρο t . Επομένως, η ισορροπία Nash για το πάιγνιο αυτό ορίζεται ως εξής:

Σε ένα γύρο, για συγκεκριμένα φορτία στις ουρές των δορυφόρων, η στρατηγική $s_N^* = [s_i^*]_{i \in N} = (s_1^*, s_2^*, \dots, s_n^*)$ αποτελεί το Σημείο Ισορροπίας κατά Nash (NE) του παιχνιδιού G_1 , αν και μόνο αν κανένας δορυφόρος δε μπορεί να αυξήσει τη συνάρτηση απόδοσής του παρεκκλίνοντας από την τρέχουσα στρατηγική s_i^* , δηλαδή,

$$u_i^t(s_i^*, s_{-i}^*) \geq u_i^t(s_i, s_{-i}^*), \quad \forall i \in N, \quad \forall s_i \in S_i.$$

Η επιλογή αυτής της συνάρτησης απόδοσης είναι αρκετά φυσική. Ωστόσο, επειδή εξαρτάται από την κατάσταση των ουρών των δορυφόρων, εάν επιθυμήσουμε να μοντελοποιήσουμε το πάιγνιο για πολλούς γύρους, καταφεύγουμε αναγκαστικά σε ένα στοχαστικό πάιγνιο, το οποίο δεν απλοποιείται.

Ακριβώς λόγω της δυναμικότητας του αντίστοιχου στοχαστικού παιχνιδιού, η χρήση του G_1 για τη μοντελοποίηση του παιχνιδιού ενός γύρου δε διαθέτει εγγυήσεις σύγκλισης.

4.3 Πάιγνιο G_2/G_2^{EXP} (Πάιγνιο Δυναμικού)

Ο βασικός παράγοντας που καθιστά το παραπάνω πάιγνιο εξαρτώμενο από την κατάσταση των ουρών, και συνεπώς στοχαστικό στην περίπτωση πολλαπλών γύρων, είναι οι μεταβλητές l_i, l_j στον ορισμό της $u_{i,j}$, που εξαρτώνται από ολόκληρη την ιστορία του παιχνιδιού, αντί των πληροφοριών του τρέχοντος γύρου μόνο. Πρέπει λοιπόν να αντικατασταθούν από άλλες, που να είναι ανεξάρτητες της κατάστασης.

Εντός ενός γύρου t , το φορτίο Δl_i^t που προστίθεται στην ουρά ενός δορυφόρου i που έχει αποφασίσει να έχει το δορυφόρο j ως στόχο σε αυτό το γύρο καθορίζεται από:

1. Το φορτίο λόγω των νέων διεργασιών που παράχθηκαν για το δορυφόρο i στον τρέχοντα γύρο ($l_{i,new}^t$).

2. Το φορτίο που εξυπηρετήθηκε από την υπολογιστική μονάδα του δορυφόρου στον τρέχοντα γύρο και αφαιρείται, δηλαδή ϕ_i .
3. Το φορτίο που εκφορτώνεται στο δορυφόρο j , δηλαδή $b_{i,j}^t$.
4. Το φορτίο λόγω των διεργασιών που καταφθάνουν στο δορυφόρο i στον τρέχοντα γύρο εξαιτίας εκφορτώσεων άλλων δορυφόρων, δηλαδή $\sum_{target(j)=i} b_{j,i}^t$.

Συνολικά λοιπόν:

$$\Delta l_i^t = l_{i,new}^t - \phi_i - b_{i,j}^t + \sum_{target(j)=i} b_{j,i}^t$$

Αποδείχθηκε στο Κεφάλαιο 3 ότι οι μεταβλητές $b_{i,j}$ μπορούν να είναι σταθερές στην κατάσταση ισορροπίας. Επομένως, στην κατάσταση ισορροπίας:

$$\Delta l_i^t = l_{i,new}^t - \phi_i - b_{i,j}^{OPT} + \sum_{target(j)=i} b_{j,i}^{OPT}$$

Λαμβάνοντας τη μέση τιμή:

$$\begin{aligned} E[\Delta l_i^t] &= E[l_{i,new}^t] - \phi_i - \sum_{j \neq i} p_{i,j}^{OPT} b_{i,j}^{OPT} + \sum_{j \neq i} p_{j,i}^{OPT} b_{j,i}^{OPT} = \\ &= \lambda_i d_i - \phi_i - \sum_{j \neq i} p_{i,j}^{OPT} b_{i,j}^{OPT} + \sum_{j \neq i} p_{j,i}^{OPT} b_{j,i}^{OPT} \end{aligned}$$

Στην κατάσταση ισορροπίας, όλες οι ουρές κατά μέσο όρο έχουν το ίδιο φορτίο. Άρα, κατά μέσο όρο, το νέο φορτίο που προστίθεται στην ουρά κάθε δορυφόρου σε κάθε γύρο είναι μηδενικό:

$$E[\Delta l_i^t] = 0 \Rightarrow \lambda_i d_i - \phi_i = \sum_{j \neq i} p_{i,j}^{OPT} b_{i,j}^{OPT} - \sum_{j \neq i} p_{j,i}^{OPT} b_{j,i}^{OPT}$$

Το φαινόμενο αυτό μπορεί να ερμηνευθεί ως εξής: Τα φορτία ισορροπούν κατά μέσο όρο σε μια τιμή l_{eq} , κοινή για όλους τους δορυφόρους στην κατάσταση ισορροπίας. Σε κάθε γύρο, κάθε δορυφόρος i παρεκκλίνει από αυτή την τιμή κατά μέσο όρο κατά $\lambda_i d_i - \phi_i$. Οι εκφορτώσεις κατά μέσο όρο αντισταθμίζουν την επίδραση του $\lambda_i d_i - \phi_i$, που είναι διαφορετικό για κάθε δορυφόρο, ώστε το φορτίο να επανέλθει στην τιμή ισορροπίας. Έτσι, η μέση διαφορά μεταξύ των φορτίων l_i, l_j δύο δορυφόρων i, j πριν πραγματοποιηθεί η εκφόρτωση στην κατάσταση ισορροπίας ισούται με:

$$E[l_i - l_j] = (\lambda_i d_i - \phi_i) - (\lambda_j d_j - \phi_j)$$

Σύμφωνα λοιπόν με τη στρατηγική ποσού εκφόρτωσης που αναλύθηκε στο Κεφάλαιο 3, η ποσότητα που εκφορτώνεται από το δορυφόρο i στο δορυφόρο j , εάν ο τελευταίος επιλεχθεί ως στόχος, ισούται με:

$$b_{i,j}^{OPT} = \frac{(\lambda_i d_i - \phi_i) - (\lambda_j d_j - \phi_j)}{2}$$

Επιστρέφουμε στην προηγούμενη σχέση:

$$\begin{aligned}
E[\Delta l_i^t] = 0 &\Rightarrow \lambda_j d_j - \phi_j = \sum_{k \neq j} p_{j,k}^{OPT} b_{j,k}^{OPT} - \sum_{k \neq j} p_{k,j}^{OPT} b_{k,j}^{OPT} \Rightarrow \\
&\Rightarrow \lambda_j d_j - \phi_j = \sum_{k \neq j} p_{j,k}^{OPT} b_{j,k}^{OPT} - \sum_{k \neq i,j} p_{k,j}^{OPT} b_{k,j}^{OPT} - p_{i,j} b_{i,j} \Rightarrow \\
&\Rightarrow p_{i,j}^{OPT} b_{i,j}^{OPT} = -\lambda_j d_j + \phi_j + \sum_{k \neq j} p_{j,k}^{OPT} b_{j,k}^{OPT} - \sum_{k \neq i,j} p_{k,j}^{OPT} b_{k,j}^{OPT} = \\
&= -\lambda_j d_j + \phi_j + \sum_{k \neq j} p_{j,k}^{OPT} \frac{(\lambda_j d_j - \phi_j) - (\lambda_k d_k - \phi_k)}{2} - \sum_{k \neq i,j} p_{k,j}^{OPT} \frac{(\lambda_k d_k - \phi_k) - (\lambda_j d_j - \phi_j)}{2}
\end{aligned}$$

Ένας δορυφόρος μπορεί να ελέγξει κατά πόσο η στρατηγική του, δεδομένων των στρατηγικών των άλλων δορυφόρων, συμβάλλει στη διατήρηση των φορτίων σε ισορροπία, ελέγχοντας την παρακάτω ποσότητα:

$$u_{i,j}^{virtual} = -\lambda_j d_j + \phi_j + \sum_{k \neq j} p_{j,k}^{OPT} \frac{(\lambda_j d_j - \phi_j) - (\lambda_k d_k - \phi_k)}{2} - \sum_{k \neq i,j} p_{k,j}^{OPT} \frac{(\lambda_k d_k - \phi_k) - (\lambda_j d_j - \phi_j)}{2}$$

Λόγω των πραγματικών περιορισμών (μη αρνητικές ποσότητες εκφόρτωσης, ανώτατο όριο c_j), τελικά έχουμε:

$$\begin{aligned}
u_{i,j}^{virtual} &= -\lambda_j d_j + \phi_j + \sum_{k \neq j} p_{j,k}^{OPT} * \max\{0, \min\{\frac{(\lambda_j d_j - \phi_j) - (\lambda_k d_k - \phi_k)}{2}, c_j\}\} - \\
&- \sum_{k \neq i,j} p_{k,j}^{OPT} * \max\{0, \min\{\frac{(\lambda_k d_k - \phi_k) - (\lambda_j d_j - \phi_j)}{2}, c_k\}\}
\end{aligned}$$

Πράγματι, όσο πιο θετική είναι η παραπάνω ποσότητα, τόσο σημαντικότερη είναι η συμβολή του δορυφόρου i στην ισορροπία του δορυφόρου j , ενώ, αντιθέτως, εάν είναι αρνητική, ο δορυφόρος i προσθέτει στο δορυφόρο j φορτίο που ο τελευταίος δε μπορεί κατά μέσο όρο να εξυπηρετήσει.

Στα πλαίσια ενός γύρου, οι δορυφόροι μπορούν να χρησιμοποιήσουν αντίστοιχα την παρακάτω ποσότητα:

$$\begin{aligned}
u_{i,j}^t &= -l_{j,new}^t + \phi_j + \sum_{k \neq j, \text{ target}(j,t)=k} \max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{k,new}^t - \phi_k)}{2}, c_j\}\} - \\
&- \sum_{k \neq i,j, \text{ target}(k,t)=j} \max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\}\} \Rightarrow
\end{aligned}$$

$$\Rightarrow u_{i,j}^t = -l_{j,new}^t + \phi_j + \max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{target(j,t),new}^t - \phi_{target(j,t)})}{2}, c_j\}\} -$$

$$- \sum_{k \neq i,j, target(k,t)=j} \max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\}\}$$

Άρα, η ποσότητα αυτή μπορεί να χρησιμοποιηθεί ως συνάρτηση απόδοσης. Προς απλοποίηση, θα αφαιρεθεί ο όρος $\max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{target(j,t),new}^t - \phi_{target(j,t)})}{2}, c_j\}\}$, καθώς μετακυλίζει την επίδραση μιας επιλογής δορυφόρου-στόχου και στο στόχο του δορυφόρου-στόχου. Άλλωστε, η τιμή της ποσότητας αυτής φράσσεται μεταξύ 0 και c_j . Το c_j με τη σειρά του, στην πιο συντηρητική θεώρηση, μπορεί να τεθεί ίσο με $\lambda_j d_j - \phi_j$, ώστε να επαρκεί ακριβώς για την ισορροπία της ουράς και να μην επιτρέπει άσκοπες εκφορτώσεις. Στην περίπτωση αυτή η ποσότητα $\max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{target(j,t),new}^t - \phi_{target(j,t)})}{2}, c_j\}\}$ φράσσεται μεταξύ 0 και $\frac{l_{j,new}^t - \phi_j}{2}$, άρα διαθέτει μικρότερη συμβολή στη συνάρτηση απόδοσης τόσο από τον πρώτο όρο (την ποσότητα $-l_{j,new}^t + \phi_j$, που είναι τουλάχιστον διπλάσια), όσο και από το δεύτερο όρο (την ποσότητα $-\sum_{k \neq i,j, target(k,t)=j} \max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\}\}$, που αυξάνεται αυθαίρετα).

Επίσης, για να διατηρηθεί συμμετρία μεταξύ των όρων των συναρτήσεων για όλους τους δορυφόρους, αντικαθιστούμε τον όρο

$$\sum_{k \neq i,j, target(k,t)=j} \max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\}\}$$

με τον όρο:

$$\sum_{k \neq i,j, target(k,t)=j} [\frac{1}{2}(\max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\}\} +$$

$$+ \max\{0, \min\{\frac{(l_{i,new}^t - \phi_i) - (l_{j,new}^t - \phi_j)}{2}, c_i\}\})]$$

Τελικά, η συνάρτηση απόδοσης ορίζεται ως:

$$u_{i,j}^t = -l_{j,new}^t + \phi_j -$$

$$- \sum_{k \neq i,j, target(k,t)=j} [\frac{1}{2}(\max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\}\} +$$

$$+ \max\{0, \min\{\frac{(l_{i,new}^t - \phi_i) - (l_{j,new}^t - \phi_j)}{2}, c_i\}\})]$$

Αυτή η συνάρτηση απόδοσης δεν εξαρτάται από την ιστορία του παιχνιδιού και δε χρειάζεται μοντελοποίηση με βάση καταστάσεις. Είναι βέβαια μη ντετερμινιστική, αφού εξαρτάται από τις τυχαίες μεταβλητές $l_{i,new}^t, i \in 1, \dots, n$. Έτσι, η ισορροπία Nash για αυτό το παιχνίδι ορίζεται ως εξής:

Το σύνολο των (πιθανώς μεικτών) στρατηγικών $s_N^* = [s_i^*]_{i \in N} = (s_1^*, s_2^*, \dots, s_n^*)$ αποτελεί

το Σημείο Ισορροπίας κατά Nash (NE) του παιχνιδιού G_2 , αν και μόνο αν κανένας δορυφόρος δε μπορεί να αυξήσει την αναμενόμενη της συνάρτησης απόδοσής του παρεκκλίνοντας από την τρέχουσα στρατηγική s_i^* , δηλαδή,

$$\tilde{u}_i(s_i^*, s_{-i}^*) \geq \tilde{u}_i(s_i, s_{-i}^*), \quad \forall i \in N, \quad \forall s_i \in \Delta(S_i).$$

Παρατηρούμε λοιπόν ότι η χρήση μιας συνάρτησης απόδοσης που δε βασίζεται σε ιστορία/καταστάσεις μετατρέπει το σύνολο καταστάσεων σε μονοσύνολο, και το στοχαστικό παίγνιο σε ένα επαναλαμβανόμενο παίγνιο με στοχαστικές ανταμοιβές.

Σύμφωνα με τη μοντελοποίηση του G_2 , ο κάθε δορυφόρος i λαμβάνει για την επιλογή του j μια ανταμοιβή σύμφωνα με την παρακάτω συνάρτηση απόδοσης:

$$\begin{aligned} u_{i,j}^t = & -l_{j,new}^t + \phi_j - \\ & - \sum_{k \neq i,j, \text{ target}(k,t)=j} \left[\frac{1}{2} \left(\max\{0, \min\left\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\}\right\} + \right. \right. \\ & \left. \left. + \max\{0, \min\left\{\frac{(l_{i,new}^t - \phi_i) - (l_{j,new}^t - \phi_j)}{2}, c_i\}\right\} \right) \right] \end{aligned}$$

Η απόδοση αυτή είναι στοχαστική, διότι οι μεταβλητές $l_{i,new}^t, i \in 1, \dots, n$ είναι τυχαίες. Ισχύει προφανώς ότι:

$$u_{i,j}^t = E[u_{i,j}] + (u_{i,j}^t - E[u_{i,j}])$$

Ισοδύναμα, λοιπόν, μπορούμε να αντιμετωπίσουμε το παίγνιο G_2 ως ένα παίγνιο στο οποίο η συνάρτηση απόδοσης $u'_{i,j}$ είναι η αναμενόμενη τιμή $E[u_{i,j}]$, την οποία όμως οι πάικτες δε λαμβάνουν αυτούσια, αλλά ύστερα από την προσθήκη ενός θορύβου $u_{i,j}^t - E[u_{i,j}]$. Θα δούμε στη συνέχεια ότι η μοντελοποίηση αυτή είναι αρκετά χρήσιμη για την απόδειξη σύγκλισης.

Έτσι, για το ισοδύναμο παίγνιο G_2^{EXP} έχουμε:

$$\begin{aligned} u'_{i,j} = & E[u_{i,j}] = -\lambda_j d_j + \phi_j - \\ & - \sum_{k \neq i,j, \text{ target}(k,t)=j} E \left[\frac{1}{2} \left(\max\{0, \min\left\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\}\right\} + \right. \right. \\ & \left. \left. + \max\{0, \min\left\{\frac{(l_{i,new}^t - \phi_i) - (l_{j,new}^t - \phi_j)}{2}, c_i\}\right\} \right) \right] \end{aligned}$$

Έστω ότι στο γύρο t οι δορυφόροι $1, \dots, n$, ύστερα από δειγματοληψία της στρατηγικής επιλογής στόχου τους, επέλεξαν ως στόχο τους δορυφόρους $s_1^t, \dots, s_n^t$ αντιστοίχα, που συνολικά σχηματίζουν το διάνυσμα $s_N^t = [s_i^t]_{i \in N}$. Ορίζουμε την ακόλουθη συνάρτηση δυναμικού:

$$\Phi(s_N^t, l_{N,new}^t) = \sum_{i, j=s_i^t} [-\lambda_j d_j + \phi_j] -$$

$$\begin{aligned}
& - \sum_{i \neq j, s_i^t = s_j^t = k} E[\frac{1}{4} (\max\{0, \min\{\frac{(l_{i,new}^t - \phi_i) - (l_{k,new}^t - \phi_k)}{2}, c_i\}\} + \\
& \quad + \max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{k,new}^t - \phi_k)}{2}, c_j\}\})]
\end{aligned}$$

Έστω τώρα ότι ο δορυφόρος q αλλάζει την επιλογή του, από s_q^t σε $s_{q,alt}^t$, με τις επιλογές των υπόλοιπων δορυφόρων να παραμένουν ίδιες και με το νέο σύνολο επιλογών να συμβολίζεται ως $s_{N,alt}^t$. Η μεταβολή στη συνάρτηση απόδοσης ισούται με:

$$\begin{aligned}
\Delta u'_q &= u'_{q,s_{q,alt}^t} - u'_{q,s_q^t} = \\
& -\lambda_{s_{q,alt}^t} d_{s_{q,alt}^t} + \phi_{s_{q,alt}^t} - \sum_{k \neq q, s_k^t = s_{q,alt}^t} E[\frac{1}{2} (\max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{s_{q,alt}^t,new}^t - \phi_{s_{q,alt}^t}^t)}{2}, c_k\}\} + \\
& \quad + \max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{s_{q,alt}^t,new}^t - \phi_{s_q^t}^t)}{2}, c_q\}\})] + \\
& + \lambda_{s_q^t} d_{s_q^t} - \phi_{s_q^t} + \sum_{k \neq q, s_k^t = s_q^t} E[\frac{1}{2} (\max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{s_q^t,new}^t - \phi_{s_q^t}^t)}{2}, c_k\}\} + \\
& \quad + \max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{s_q^t,new}^t - \phi_{s_q^t}^t)}{2}, c_q\}\})]
\end{aligned}$$

Η αντίστοιχη μεταβολή στη συνάρτηση δυναμικού ισούται με:

$$\begin{aligned}
\Delta \Phi &= \Phi(s_{N,alt}^t, l_{N,new}^t) - \Phi(s_N^t, l_{N,new}^t) = \\
& = \sum_{i, j = s_{i,alt}^t} [-\lambda_j d_j + \phi_j] - \sum_{i \neq j, s_{i,alt}^t = s_{j,alt}^t = k} E[\frac{1}{4} (\max\{0, \min\{\frac{(l_{i,new}^t - \phi_i) - (l_{k,new}^t - \phi_k)}{2}, c_i\}\} + \\
& \quad + \max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{k,new}^t - \phi_k)}{2}, c_j\}\})] - \\
& - \sum_{i, j = s_i^t} [-\lambda_j d_j + \phi_j] + \sum_{i \neq j, s_i^t = s_j^t = k} E[\frac{1}{4} (\max\{0, \min\{\frac{(l_{i,new}^t - \phi_i) - (l_{k,new}^t - \phi_k)}{2}, c_i\}\} + \\
& \quad + \max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{k,new}^t - \phi_k)}{2}, c_j\}\})] =
\end{aligned}$$

$$\begin{aligned}
&= -\lambda_{s_{q,alt}^t} d_{s_{q,alt}^t} + \phi_{s_{q,alt}^t} - \sum_{q \neq j, s_{q,alt}^t = s_{j,alt}^t = k} E\left[\frac{1}{4} \left(\max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{k,new}^t - \phi_k)}{2}, c_q\}\} + \right. \right. \\
&\quad \left. \left. + \max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{k,new}^t - \phi_k)}{2}, c_j\}\} \right) \right] - \\
&\quad - \sum_{i \neq q, s_{i,alt}^t = s_{q,alt}^t = k} E\left[\frac{1}{4} \left(\max\{0, \min\{\frac{(l_{i,new}^t - \phi_i) - (l_{k,new}^t - \phi_k)}{2}, c_i\}\} + \right. \right. \\
&\quad \left. \left. + \max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{k,new}^t - \phi_k)}{2}, c_q\}\} \right) \right] + \\
&\quad + \lambda_{s_q^t} d_{s_q^t} - \phi_{s_q^t} + \sum_{q \neq j, s_q^t = s_j^t = k} E\left[\frac{1}{4} \left(\max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{k,new}^t - \phi_k)}{2}, c_q\}\} + \right. \right. \\
&\quad \left. \left. + \max\{0, \min\{\frac{(l_{j,new}^t - \phi_j) - (l_{k,new}^t - \phi_k)}{2}, c_j\}\} \right) \right] - \\
&\quad + \sum_{i \neq q, s_i^t = s_q^t = k} E\left[\frac{1}{4} \left(\max\{0, \min\{\frac{(l_{i,new}^t - \phi_i) - (l_{k,new}^t - \phi_k)}{2}, c_i\}\} + \right. \right. \\
&\quad \left. \left. + \max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{k,new}^t - \phi_k)}{2}, c_q\}\} \right) \right] = \\
&= -\lambda_{s_{q,alt}^t} d_{s_{q,alt}^t} + \phi_{s_{q,alt}^t} - \sum_{k \neq q, s_k^t = s_{q,alt}^t} E\left[\frac{1}{2} \left(\max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{s_{q,alt}^t,new}^t - \phi_{s_{q,alt}^t}^t)}{2}, c_k\}\} + \right. \right. \\
&\quad \left. \left. + \max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{s_{q,alt}^t,new}^t - \phi_{s_q^t}^t)}{2}, c_q\}\} \right) \right] + \\
&\quad + \lambda_{s_q^t} d_{s_q^t} - \phi_{s_q^t} + \sum_{k \neq q, s_k^t = s_q^t} E\left[\frac{1}{2} \left(\max\{0, \min\{\frac{(l_{k,new}^t - \phi_k) - (l_{s_q^t,new}^t - \phi_{s_q^t}^t)}{2}, c_k\}\} + \right. \right. \\
&\quad \left. \left. + \max\{0, \min\{\frac{(l_{q,new}^t - \phi_q) - (l_{s_q^t,new}^t - \phi_{s_q^t}^t)}{2}, c_q\}\} \right) \right] = \Delta u'_q
\end{aligned}$$

Εφόσον έχουμε $\Delta\Phi = \Delta u'_q$, συμπεραίνουμε ότι το στατικό παίγνιο ενός γύρου αποτελεί παίγνιο δυναμικού.

Όπως αναλύθηκε προηγουμένως, όταν οι δορυφόροι παίζουν σύμφωνα με το παίγνιο G_2 , ουσιαστικά παίζουν σύμφωνα με το παίγνιο G_2^{EXP} , λαμβάνοντας όμως μια θορυβώδη εκδοχή της συνάρτησης απόδοσης, με θόρυβο $u_{i,j}^t - E[u_{i,j}]$.

Εξετάζοντας αυτό το θόρυβο, παρατηρούμε ότι:

- Έχει μηδενική μέση τιμή:

$$E[u_{i,j}^t - E[u_{i,j}]] = 0$$

- Είναι φραγμένος:

$$\begin{aligned} 0 &\leq l_{j,new}^t \leq Q, \\ -Q + \phi_j - (n-1)\left[\frac{1}{2}(c_k + c_i)\right] &\leq u_{i,j}^t \leq \phi_j, \\ -Q + \phi_j - (n-1)\left[\frac{1}{2}(c_k + c_i)\right] &\leq E[u_{i,j}^t] \leq \phi_j, \\ -Q - (n-1)\left[\frac{1}{2}(c_k + c_i)\right] &\leq u_{i,j}^t - E[u_{i,j}] \leq Q + (n-1)\left[\frac{1}{2}(c_k + c_i)\right], \end{aligned}$$

Επομένως, ο θόρυβος ικανοποιεί τις προϋποθέσεις της εργασίας [CHM17] (βλ. Κεφάλαιο 2). Επιπλέον, εφόσον το παίγνιο G_2^{EXP} είναι παίγνιο δυναμικού, έχουμε, και πάλι με βάση την αντίστοιχη ενότητα του κεφαλαίου 2, ότι:

- Μη ανατροφοδοτούμενη περίπτωση: Είναι το πραγματικό, πλήρως αποκεντρωμένο σενάριο που εξετάζουμε. Κάθε δορυφόρος σε κάθε γύρο επιλέγει έναν μοναδικό δορυφόρο-στόχο και λαμβάνει, ύστερα από την επικοινωνία μαζί του, μια απόδοση για την ενέργεια αυτή. Σύμφωνα με τη θεωρία, έχουμε:

Σύγκλιση. Υποθέτοντας ότι ο Αλγόριθμος ε -Hedge εκτελείται με i) τον εκτιμητή bandit (μη ανατροφοδοτούμενη πληροφόρηση), ii) θετικό συντελεστή εξερεύνησης $\varepsilon > 0$, και iii) ακολουθία βημάτων της μορφής $\gamma_n \propto 1/n^\beta$ για κάποιο $\beta \in (0, 1]$, τότε οι στρατηγικές των δορυφόρων συγκλίνουν (σ.σ.) σε δ -ισορροπία του Γ , όπου $\delta \equiv \delta(\varepsilon) \rightarrow 0$ καθώς $\varepsilon \rightarrow 0$.

- Ημιανατροφοδοτούμενη περίπτωση: Αποτελεί ένα λιγότερο ρεαλιστικό σενάριο, στο οποίο κάθε δορυφόρος επιλέγει μοναδικό δορυφόρο-στόχο σε κάθε γύρο, αλλά γνωρίζει την απόδοση που θα λάμβανε εάν είχε επιλέξει οποιονδήποτε άλλο δορυφόρο. Προφανώς, αυτό συνιστά μια μερικώς αποκεντρωμένη μέθοδο, που απαιτεί ευρεκπομπή πακέτων για την ενημέρωση του διανύσματος αποδόσεων για όλους τους δορυφόρους. Παρά τη μη αποδοτική φύση του, έχει ενδιαφέρον η θεωρητική μελέτη του, καθώς συγκλίνει εγγυημένα σε ισορροπία Nash:

Σύγκλιση. Υποθέτοντας ότι ο Αλγόριθμος ε -Hedge εκτελείται με i) ημιανατροφοδοτούμενη πληροφόρηση, ii) μηδενικό παράγοντα εξερεύνησης $\varepsilon = 0$, και iii) μια ακολουθία βημάτων της μορφής $\gamma_n \propto 1/n^\beta$ για κάποιο $\beta \in (0, 1]$, τότε οι στρατηγικές των δορυφόρων συγκλίνουν σε ισορροπία Nash με πιθανότητα 1.

4.4 Ενεργειακά αποδοτικά παίγνια

Στη μοντελοποίηση που παρουσιάστηκε παραπάνω, υποθέσαμε ότι η εκφόρτωση εργασιών είναι πάντα δυνατή μεταξύ των δορυφόρων, αλλά περιορίζεται από μια μεταβλητή c_i που αντιστοιχεί στις δυνατότητες μετάδοσης του κάθε δορυφόρου. Στην πραγματικότητα, η ενεργειακή απόδοση μπορεί επίσης να αποτελεί ζήτημα. Για κάθε δορυφόρο-πηγή, ο δορυφόρος-προορισμός θα πρέπει να επιλέγεται με επίγνωση του καναλιού: διαδορυφορικές ζεύξεις που χαρακτηρίζονται από χαμηλό ρυθμό μετάδοσης και άρα απαιτούν σημαντική ενέργεια για την εκφόρτωση θα πρέπει να αποφεύγονται.

Η μέση ενέργεια που απαιτείται για την εκφόρτωση φορτίου $b_{i,j}^t$ από τον δορυφόρο i στον δορυφόρο-στόχο j δίνεται από τη σχέση:

$$E_{i,j}^t = \frac{P_i^{LB} * b_{i,j}^t}{R_{i,j}},$$

όπου $R_{i,j}$ ο μέσος ρυθμός μετάδοσης της διαδορυφορικής ζεύξης και το P_i^{LB} είναι η ισχύς εκπομπής του δορυφόρου i .

Η ενεργειακή απόδοση μπορεί να ενσωματωθεί σε όλα τα παίγνια G_1, G_2, G_2^{EXP} που αναλύθηκαν προηγουμένως, μέσω ενός προσθετικού όρου στη συνάρτηση απόδοσης. Ενδεικτικά, θα την εφαρμόσουμε στην περίπτωση του G_2^{EXP} . Παρόμοια με τη μέθοδο σταθμισμένου αθροίσματος στη βελτιστοποίηση Pareto, αφαιρούμε τον όρο $c * \frac{1}{R_{i,j}} * b_{i,j}^{OPT}$ από την αρχική συνάρτηση χρησιμότητας του G_2/G_2^{EXP} :

$$\begin{aligned} & u'_{i,j} = -\lambda_j d_j + \phi_j - \\ & - \sum_{k \neq i,j, \text{ target}(k,t)=j} E\left[\frac{1}{2} \left(\max\{0, \min\left\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\right\}\} + \right. \right. \\ & \left. \left. + \max\{0, \min\left\{\frac{(l_{i,new}^t - \phi_i) - (l_{j,new}^t - \phi_j)}{2}, c_i\right\}\} \right) \right] - c * \frac{1}{R_{i,j}} * b_{i,j}^{OPT} \Rightarrow \\ & \Rightarrow u'_{i,j} = -\lambda_j d_j + \phi_j - \\ & - \sum_{k \neq i,j, \text{ target}(k,t)=j} E\left[\frac{1}{2} \left(\max\{0, \min\left\{\frac{(l_{k,new}^t - \phi_k) - (l_{j,new}^t - \phi_j)}{2}, c_k\right\}\} + \right. \right. \\ & \left. \left. + \max\{0, \min\left\{\frac{(l_{i,new}^t - \phi_i) - (l_{j,new}^t - \phi_j)}{2}, c_i\right\}\} \right) \right] - c * \frac{1}{R_{i,j}} * \frac{(\lambda_i d_i - \phi_i) - (\lambda_j d_j - \phi_j)}{2} \end{aligned}$$

Ο όρος αυτός είναι ανάλογος με την ενεργειακή απόδοση και σταθμίζεται με έναν σταθερό συντελεστή c . Για το λόγο αυτό, τροποποιούμε και τη συνάρτηση δυναμικού:

$$\Phi(s_N^t, l_{N,new}^t) = \sum_{i, j=s_i^t} \left[-\lambda_j d_j + \phi_j - c * \frac{1}{R_{i,j}} * \frac{(\lambda_i d_i - \phi_i) - (\lambda_j d_j - \phi_j)}{2} \right] -$$

$$\begin{aligned}
& - \sum_{i \neq j, s_i^t = s_j^t = k} E \left[\frac{1}{4} \left(\max\{0, \min\left\{ \frac{(l_{i,new}^t - \phi_i) - (l_{k,new}^t - \phi_k)}{2}, c_i \right\} \} + \right. \right. \\
& \quad \left. \left. + \max\{0, \min\left\{ \frac{(l_{j,new}^t - \phi_j) - (l_{k,new}^t - \phi_k)}{2}, c_j \right\} \} \right) \right]
\end{aligned}$$

Μάλιστα, ομοίως με πριν, είναι εύκολο να αποδειχτεί ότι:

$$\Delta\Phi = \Delta u'_q$$

Επομένως, η ενεργειακά αποδοτική εκδοχή του παιχνιδιού παραμένει παίγνιο δυναμικού.

Ο σταθμιστικός παράγοντας c μπορεί να ρυθμιστεί ώστε να δίνεται περισσότερη ή λιγότερη έμφαση στην ενεργειακή απόδοση σε σύγκριση με τον στόχο εξισορρόπησης φορτίου.

Ομοίως προσθέτουμε τον όρο αυτό και στις συναρτήσεις απόδοσης των υπόλοιπων παιγνίων, ώστε να τα καταστήσουμε πιο ενεργειακά αποδοτικά.

4.5 Σύνοψη μοντελοποιήσεων

Όπως αναλύθηκε παραπάνω, οι στρατηγικές επιλογής στόχου εκφόρτωσης για τους δορυφόρους μπορούν να μοντελοποιηθούν ως ένα στοχαστικό παίγνιο πολλών γύρων. Το παίγνιο αυτό αναλύεται σε επιμέρους παίγνια ενός γύρου, με τους εξής εναλλακτικούς τρόπους:

1. Πάιγνιο G_1 , του οποίου η συνάρτηση απόδοσης βασίζεται με έναν αρκετά διαισθητικό τρόπο στην τρέχουσα κατάσταση των ουρών των δορυφόρων. Ως αποτέλεσμα, ωστόσο, το αντίστοιχο στοχαστικό παίγνιο πολλών γύρων δεν απλοποιείται και δε συνοδεύεται από εγγυήσεις σύγκλισης.
2. Πάιγνιο G_2 , του οποίου η συνάρτηση απόδοσης ακολουθεί μια τυχαία κατανομή. Το παίγνιο αυτό είναι ισοδύναμο με μια θορυβώδη εκδοχή του παιγνίου G_2^{EXP} και στη ρεαλιστική, μη ανατροφοδοτούμενη περίπτωση, οδηγεί σε μια δ-ισορροπία, όπου δ ανάλογο του (μη-μηδενικού) παράγοντα εξερεύνησης.
3. Πάιγνιο G_2 , ομοίως με παραπάνω, ισοδύναμο με θορυβώδη εκδοχή του G_2^{EXP} , αλλά με η-μιανατροφοδοτούμενο περιβάλλον (μερικώς αποκεντρωμένο), το οποίο οδηγεί εγγυημένα σε ισορροπία Nash.

Για κάθε ένα από αυτά τα παίγνια, μπορεί να οριστεί το αντίστοιχο ενεργειακά αποδοτικό παίγνιο, με τις ίδιες εγγυήσεις σύγκλισης, προσθέτοντας το σχετικό όρο ενέργειας, σταθμισμένο με ένα συντελεστή c .

Αξιολόγηση και Αριθμητικά Αποτελέσματα

5.1 Παραμετροποίηση

Προκειμένου να επιβεβαιωθεί η αποτελεσματικότητα των αλγορίθμων εκμάθησης της βέλτιστης στρατηγικής εκφόρτωσης, πραγματοποιήθηκε προσομοίωση ενός αστερισμού δορυφόρων με τις παραμέτρους που σημειώνονται στον πίνακα 5.1.

Για τις παραμέτρους c_i του μέγιστου δυνατού όγκου εκφόρτωσης ανά γύρο ακολουθήθηκε μια σχετικά συντηρητική εκτίμηση, θέτοντάς τες στο $c_i = 1.2 * (\lambda_i d_i - \phi_i)$, $i \in \{1, \dots, n\}$.

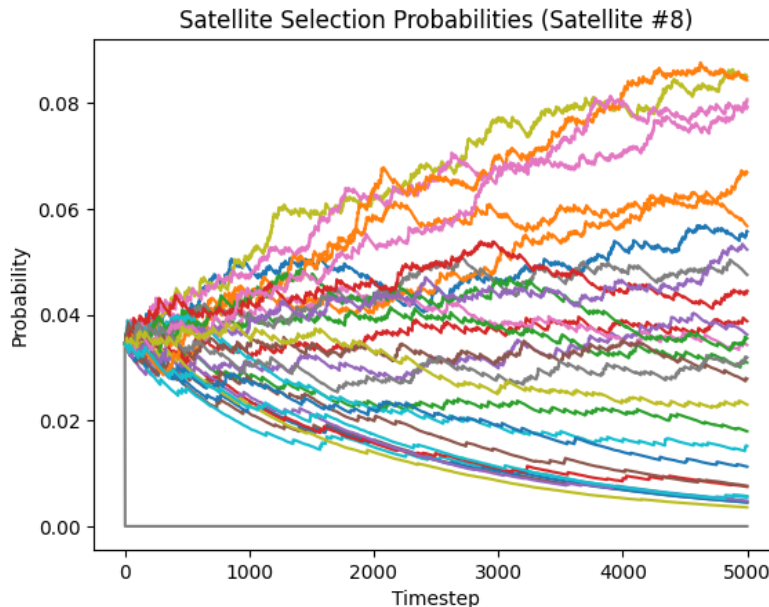
5.2 Αποτελέσματα

5.2.1 Σύγκριση μεταξύ των τριών παιγνιο-θεωρητικών μοντέλων

Αρχικά εξετάστηκε η σύγκλιση του αλγορίθμου για τις τρεις διαφορετικές μοντελοποιήσεις. Οι δορυφόροι ξεκινούν με άδειες ουρές και γεμίζουν σταδιακά με φορτίο προς εκτέλεση, βάσει των αντίστοιχων στοχαστικών ρυθμών άφιξης διεργασιών και των στοχαστικών μεγεθών αυτών. Ταυτόχρονα όμως πραγματοποιούν εκφορτώσεις μεταξύ τους, με την αναλυθείσα στρατηγική καθορισμού ποσού εκφόρτωσης (Ενότητα 3.3.1), ενώ ως προς τη στρατηγική καθορισμού στόχου εκφόρτωσης ακολουθούν (α) τυχαία επιλογή δορυφόρου-στόχου, (β) στρατηγική που επικαιροποιείται με βάση το παίγνιο G_1 , (γ) στρατηγική που επικαιροποιείται με βάση

Πίνακας 5.1: Παράμετροι Προσομοίωσης

Παράμετρος	Τιμή
Αριθμός δορυφόρων	30
Μέσος αριθμός νέων εργασιών ανά δορυφόρο (λ_i)	[1, 15]
Μέσο μέγεθος δεδομένων εργασιών (MB)	[1, 16]
Υπολογιστική ένταση εργασιών I (κύκλοι CPU/byte)	100
Υπολογιστική ικανότητα δορυφόρων ϕ_i (GHz)	[0.04, 1.6]
Χωρητικότητα ουρών δορυφόρων Q (κύκλοι CPU)	$8 * 10^{10}$
Ρυθμός μετάδοσης μεταξύ δορυφόρων $R_{i,j}$ (Gbps)	[2, 6]

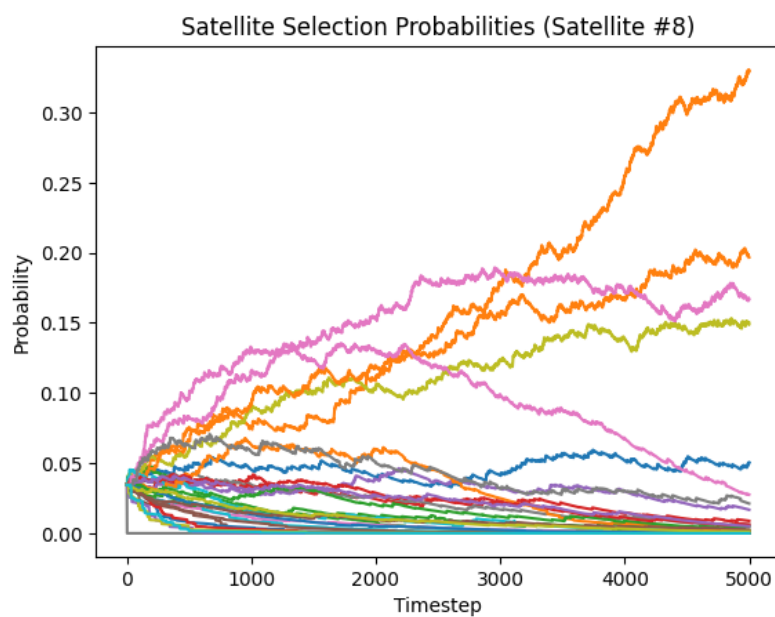


Σχήμα 5.1: Πείραμα 1 - Πάιγνιο G_1 : Σύγκλιση πιθανοτήτων επιλογής δορυφόρου-στόχου για το δορυφόρο (πομπό) νο.8.

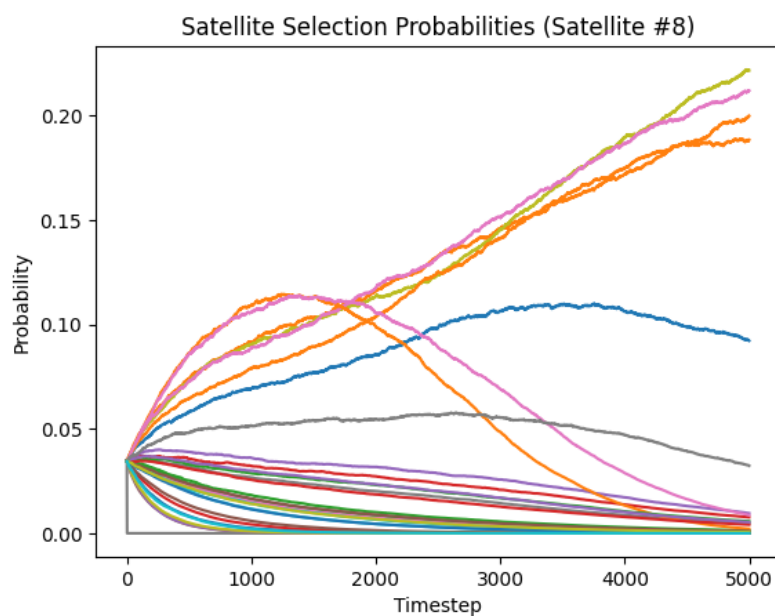
το πάιγνιο G_2/G_2^{EXP} σε μη ανατροφοδοτούμενο περιβάλλον, (δ) στρατηγική που επικαιροποιείται με βάση το πάιγνιο G_2/G_2^{EXP} σε ημιανατροφοδοτούμενο περιβάλλον. Για το πείραμα αυτό χρησιμοποιήθηκε παράγοντας εξερεύνησης $\varepsilon = 0.05$ και η ακολουθία ρυθμού μάθησης $\gamma_n = 10^{-6} * (1/n^{0.2})$. Υπενθυμίζεται ότι οι εγγυήσεις σύγκλισης βασίζονται στην υπόθεση $\gamma_n \propto 1/n^\beta$ για κάποιο $\beta \in (0, 1]$.

Αρχικά ως παρατηρήσουμε τη σύγκλιση των πιθανοτήτων επιλογής δορυφόρου-στόχου για έναν δορυφόρο-πηγή, ενδεικτικά για το δορυφόρο νο.8. Το πάιγνιο G_1 (βλ. Σχήμα 5.1) ακολουθεί ουσιαστικά μια μη ανατροφοδοτούμενη λογική και σε κάθε γύρο ανανεώνει την πιθανότητα της επιλεγμένης ενέργειας (βέβαια οι υπόλοιπες πιθανότητες ανανεώνονται επίσης, αν και σε μικρότερο βαθμό, λόγω της επακόλουθης κανονικοποίησης του συνολικού διανύσματος πιθανοτήτων). Το αποτέλεσμα είναι μια σχετικά θορυβώδης σύγκλιση, πλην όμως με ξεκάθαρες γενικές τάσεις. Ομοίως, το πάιγνιο G_2 στη μη ανατροφοδοτούμενη περίπτωση (βλ. Σχήμα 5.2), παρουσιάζει μια θορυβώδη σύγκλιση. Αντιθέτως, το πάιγνιο G_2 στην ημιανατροφοδοτούμενη περίπτωση (βλ. Σχήμα 5.3) ανανεώνει ισότιμα τις πιθανότητες όλων των πιθανών ενεργειών σε κάθε γύρο, με αποτέλεσμα μια αρκετά ομαλή σύγκλιση των πιθανοτήτων. Πάντως, οι τρεις αλγόριθμοι συγκλίνουν σε διαφορετικές στρατηγικές.

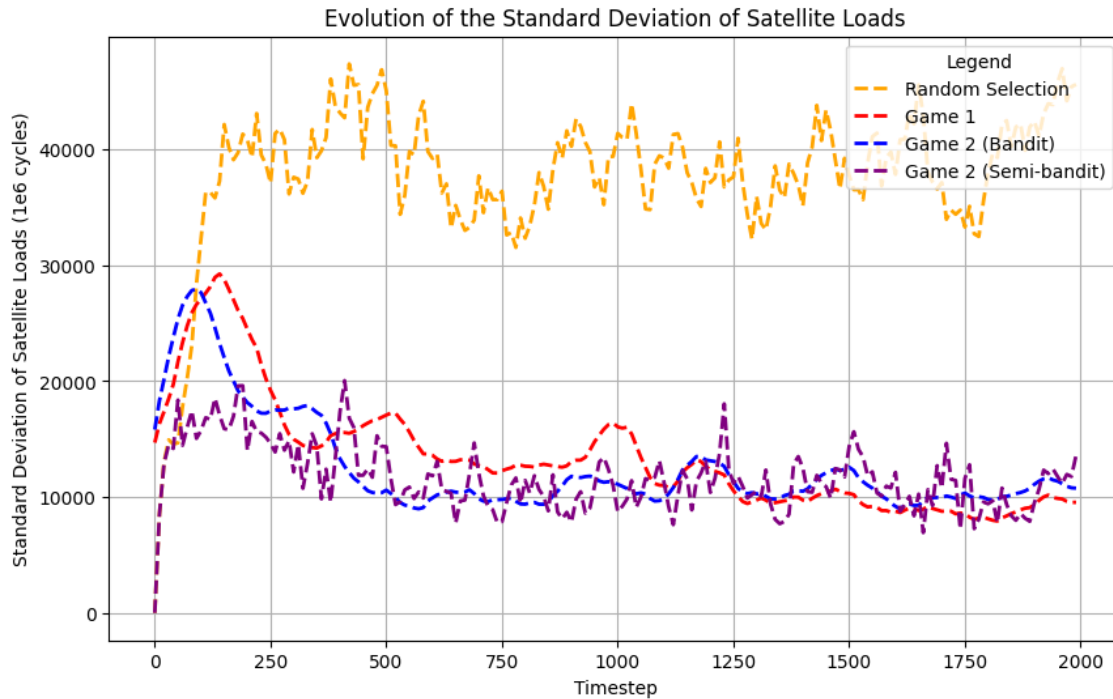
Όσον αφορά την εξέλιξη των ουρών, η συγκριτική αξιολόγηση πραγματοποιείται στο Σχήμα 5.4. Η τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων αυξάνεται στην αρχή καθώς αυτές γεμίζουν με διεργασίες, διότι κάποιοι δορυφόροι υπερφορτώνονται ενώ άλλοι όχι. Η κατάσταση αυτή είναι μη αναστρέψιμη στην περίπτωση της τυχαίας στρατηγικής εκφόρτωσης (κίτρινη καμπύλη), καθώς οι στόχοι εκφόρτωσης δεν είναι κατάλληλα προσαρμοσμένοι στις στατιστικές ιδιότητες κάθε δορυφόρου. Αντιθέτως, οι παιγνιο-θεωρητικοί μηχανισμοί αντιμετωπίζουν την ανισορροπία αρκετά γρήγορα: εντός 100-150 επαναλήψεων, η τυπική απόκλι-



Σχήμα 5.2: Πείραμα 1 - Παίγνιο G_2 (Μη ανατροφοδοτούμενη περίπτωση): Σύγκλιση πιθανοτήτων επιλογής δορυφόρου-στόχου για το δορυφόρο (πομπό) νο.8.



Σχήμα 5.3: Πείραμα 1 - Παίγνιο G_2 (Ημιανατροφοδοτούμενη περίπτωση): Σύγκλιση πιθανοτήτων επιλογής δορυφόρου-στόχου για το δορυφόρο (πομπό) νο.8.



Σχήμα 5.4: Πείραμα 1: Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (σύγκριση των τριών παιγνίων)

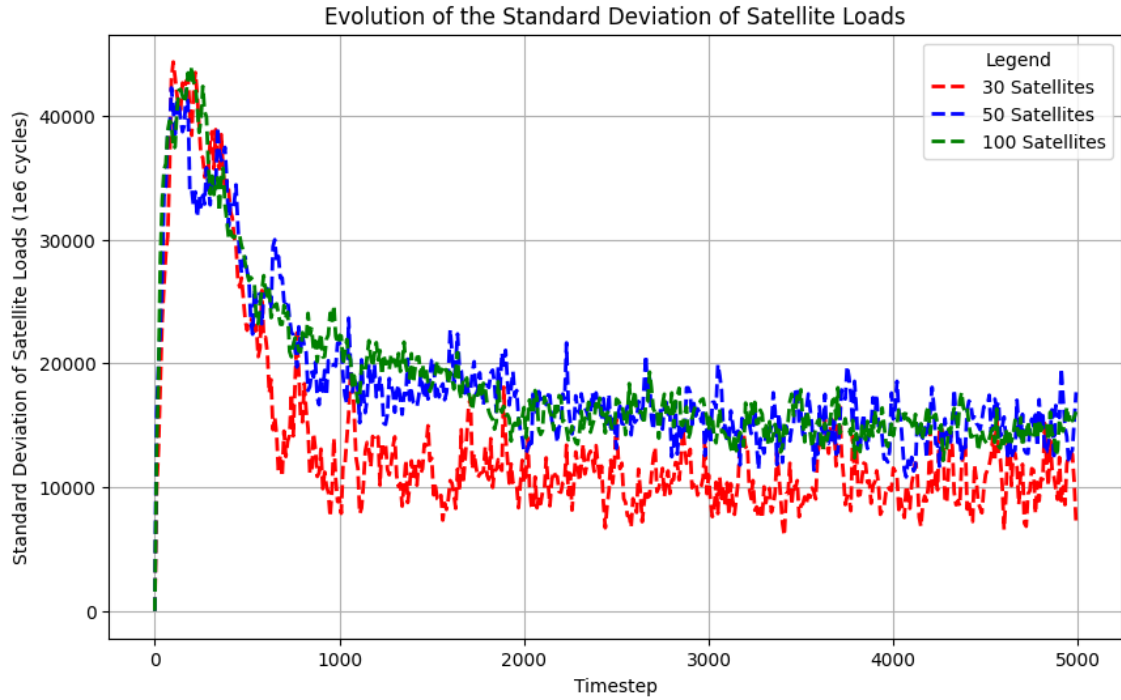
ση των φορτίων ήδη έχει αρχίσει να μειώνεται για το στοχαστικό παίγνιο G_1 και το παίγνιο δυναμικού G_2/G_2^{EXP} στο μη ανατροφοδοτούμενο περιβάλλον, ενώ για το παίγνιο δυναμικού G_2/G_2^{EXP} στο ημιανατροφοδοτούμενο περιβάλλον, η μείωση αυτή αρχίζει ακόμα νωρίτερα. Και στους τρεις μηχανισμούς, η τυπική απόκλιση καταλήγει σε τιμές σχεδόν υποτετραπλάσιες του σεναρίου τυχαίας επιλογής.

Τελικά, μεταξύ των τριών παιγνίων, το παίγνιο δυναμικού G_2/G_2^{EXP} στο μη ανατροφοδοτούμενο περιβάλλον φαίνεται να ξεχωρίζει, καθώς προσφέρει συγκρίσιμη απόδοση με τα άλλα δύο παίγνια, αλλά επιπλέον: (α) σε αντίθεση με το παίγνιο G_1 , παρέχει εγγυήσεις σύγκλισης σε ισορροπία, ενώ (β) σε αντίθεση με το παίγνιο G_2/G_2^{EXP} στο ημιανατροφοδοτούμενο περιβάλλον, είναι πλήρως αποκεντρωμένο και συνεπώς κλιμακώσιμο και εφαρμόσιμο σε πραγματικά συστήματα. Για το λόγο αυτό, τα επόμενα πειράματα θα εστιάσουν στην επίδοση αυτού μόνο του παιγνίου, υπό μεταβολές στις υπόλοιπες παραμέτρους του συστήματος και της εκμάθησης.

5.2.2 Συμπεριφορά του συστήματος υπό μεταβλητό αριθμό δορυφόρων

Εξετάζουμε την επίδοση του παιγνίου δυναμικού G_2/G_2^{EXP} στο μη ανατροφοδοτούμενο περιβάλλον για τοπολογίες με διαφορετικό αριθμό δορυφόρων (35, 50 και 100). Οι τιμές των υπόλοιπων παραμέτρων παραμένουν ίδιες με το πείραμα της ενότητας 5.2.1.

Τα αποτελέσματα παρουσιάζονται στο Σχήμα 5.5. Παρατηρούμε ότι η ταχύτητα σύγκλισης δεν επηρεάζεται ιδιαίτερα από τη μεταβολή του πλήθους των δορυφόρων. Το ακριβές σημείο ισορροπίας στο οποίο καταλήγει ο αλγόριθμος εξαρτάται από τα εκάστοτε χαρακτηριστικά



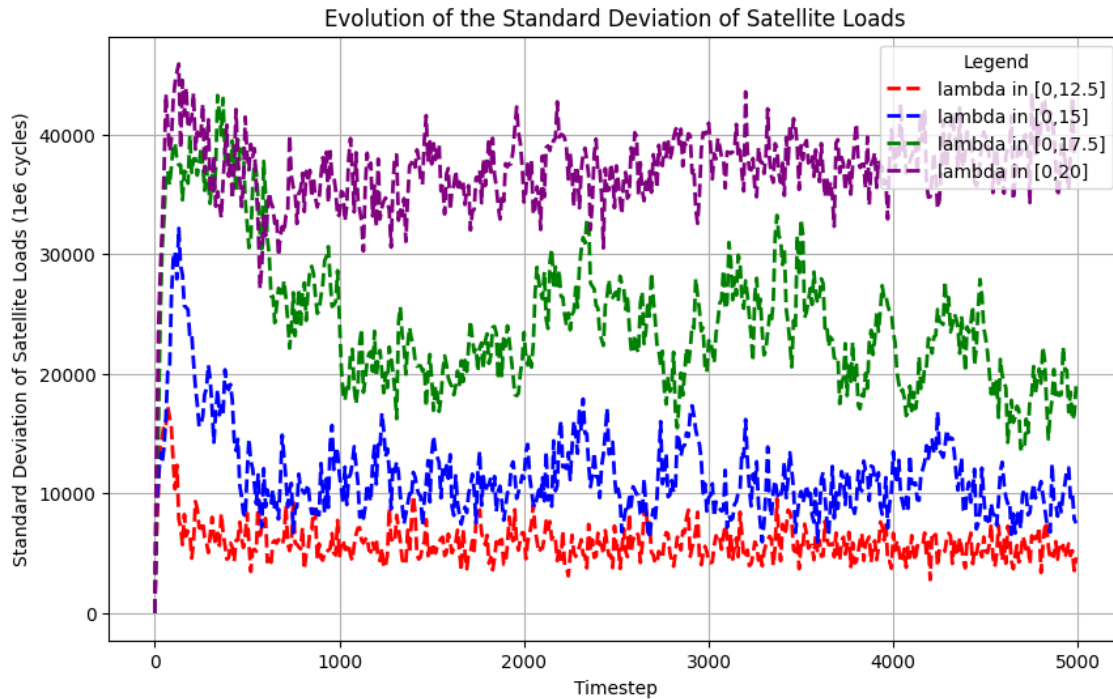
Σχήμα 5.5: Πείραμα 2 - Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (μεταβολή του πλήθους των δορυφόρων).

των δορυφόρων, οπότε οι μικρές διαφορές στις τελικές τιμές της τυπικής απόκλισης μεταξύ των τριών περιπτώσεων (30, 50 και 100 δορυφόροι) είναι αναμενόμενες.

5.2.3 Συμπεριφορά του συστήματος υπό μεταβλητό μέσο ρυθμό παραγωγής εργασιών

Σε αυτό το πείραμα καταγράφουμε την επίδοση του παιγνίου δυναμικού G_2/G_2^{EXP} στο μη ανατροφοδοτούμενο περιβάλλον για τοπολογίες με διαφορετικούς ρυθμούς παραγωγής νέων διεργασιών $\lambda_i, i \in \{0, \dots, n\}$. Στις προσομοιώσεις, ο ρυθμός αυτός για κάθε δορυφόρο επιλέγεται τυχαία (ομοιόμορφα) εντός ενός εύρους $[0, \lambda_{max}]$. Έτσι, για να εξετάσουμε την επίδοση του αλγορίθμου υπό διαφορετικούς συνολικούς όγκους συστήματος, μεταβάλλουμε το άνω όριο λ_{max} μεταξύ των τιμών 12.5, 15, 17.5 και 20. Οι τιμές των υπόλοιπων παραμέτρων παραμένουν ίδιες με το πείραμα της ενότητας 5.2.1.

Τα αποτελέσματα παρουσιάζονται στο Σχήμα 5.6. Για χαμηλές τιμές ρυθμών παραγωγής ($\lambda_{max} = 12.5$), η εξισορρόπηση των φορτίων πραγματοποιείται ταχύτατα, με την τυπική απόκλιση να ισορροπεί σε μικρές τιμές. Αυξάνοντας το λ_{max} σε 15 και 17.5, η ταχύτητα σύγκλισης μειώνεται, ενώ για την τιμή 20, ο συνολικός φόρτος στο σύστημα είναι τόσο έντονος που ακόμα και με τις κατάλληλες εκφορτώσεις η τυπική απόκλιση παραμένει σε υψηλές τιμές.



Σχήμα 5.6: Πείραμα 3 - Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (μεταβολή του εύρους των τιμών λ_i).

5.2.4 Συμπεριφορά του συστήματος υπό μεταβλητό παράγοντα εξερεύνησης

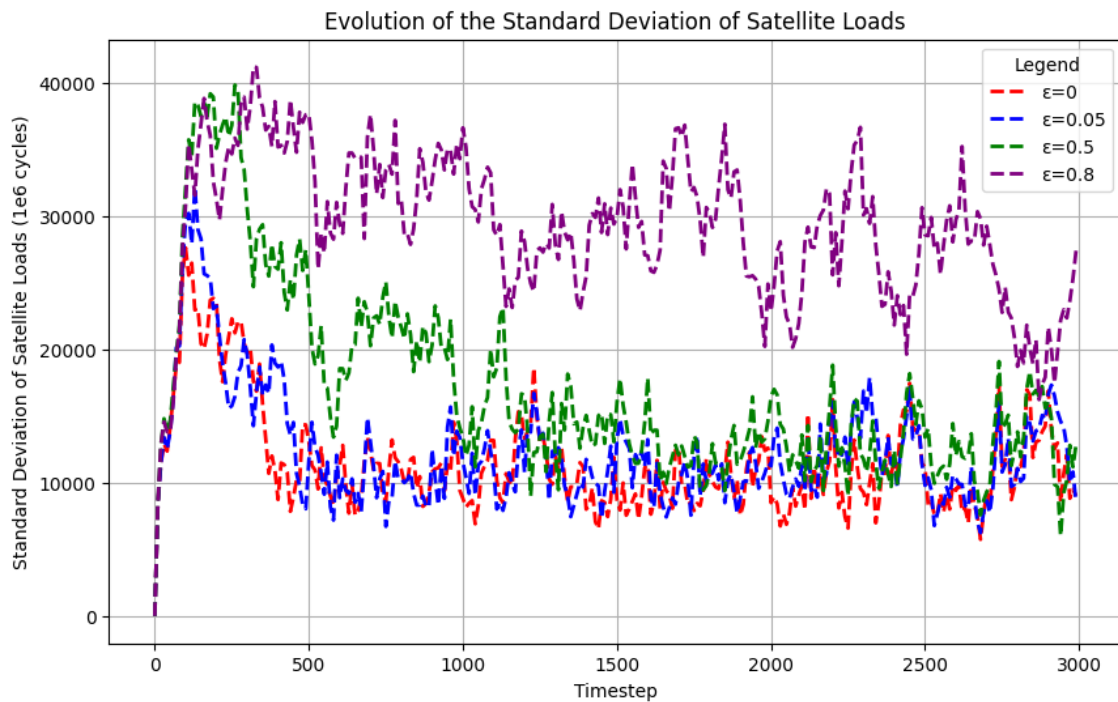
Εδώ διερευνάται η επίδραση της τιμής του παράγοντα εξερεύνησης ε στην επίδοση του παιγνίου δυναμικού G_2/G_2^{EXP} στο μη ανατροφοδοτούμενο περιβάλλον. Οι τιμές των υπόλοιπων παραμέτρων παραμένουν ίδιες με το πείραμα της ενότητας 5.2.1.

Οι εγγυήσεις σύγκλισης απαιτούν $\varepsilon > 0$. Προφανώς προτιμούνται χαμηλές τιμές, ώστε ο αλγοριθμός να επικεντρώνεται περισσότερο στην εκμετάλλευση (exploitation), παρά στην εξερεύνηση (exploration). Αυτό φαίνεται και στο θεώρημα σύγκλισης που περιγράφηκε στις ενότητες 2.2.5/4.4.6: οι στρατηγικές των δορυφόρων συγκλίνουν σε δ-ισορροπία, όπου $\delta \equiv \delta(\varepsilon) \rightarrow 0$ καθώς $\varepsilon \rightarrow 0$, δηλαδή το σύστημα προσεγγίζει όλο και περισσότερο την ισορροπία Nash για χαμηλότερες τιμές του ε .

Οι προσομοιώσεις (Σχήμα 5.7) καταδεικνύουν το συμβιβασμό μεταξύ εκμετάλλευσης και εξερεύνησης, καθώς οι πολύ υψηλές τιμές του ε (π.χ. 0.8) αποτρέπουν την εξισορρόπηση φορτίου, ακόμα και εάν οι δορυφόροι έχουν μάθει την καταλληλή στρατηγική εκφόρτωσης. Στον αντίποδα, οι χαμηλές τιμές του ε οδηγούν σε γρήγορη και αποτελεσματική εξισορρόπηση. Αξίζει να σημειωθεί ότι ακόμα και η τιμή $\varepsilon = 0$ αποδείχθηκε κατάλληλη για το συγκεκριμένο πείραμα (βλ. κόκκινη καμπύλη), παρόλο που δεν υφίστανται οι θεωρητικές εγγυήσεις στη γενική περίπτωση.

5.2.5 Συμπεριφορά του ενεργειακά αποδοτικού παιγνίου

Σύμφωνα με την ανάλυση της ενότητας 4.5, το παίγνιο δυναμικού G_2/G_2^{EXP} μπορεί να συμπεριλάβει την ενεργειακή αποδοτικότητα στη συνάρτηση απόδοσης, με την προσθήκη ενός



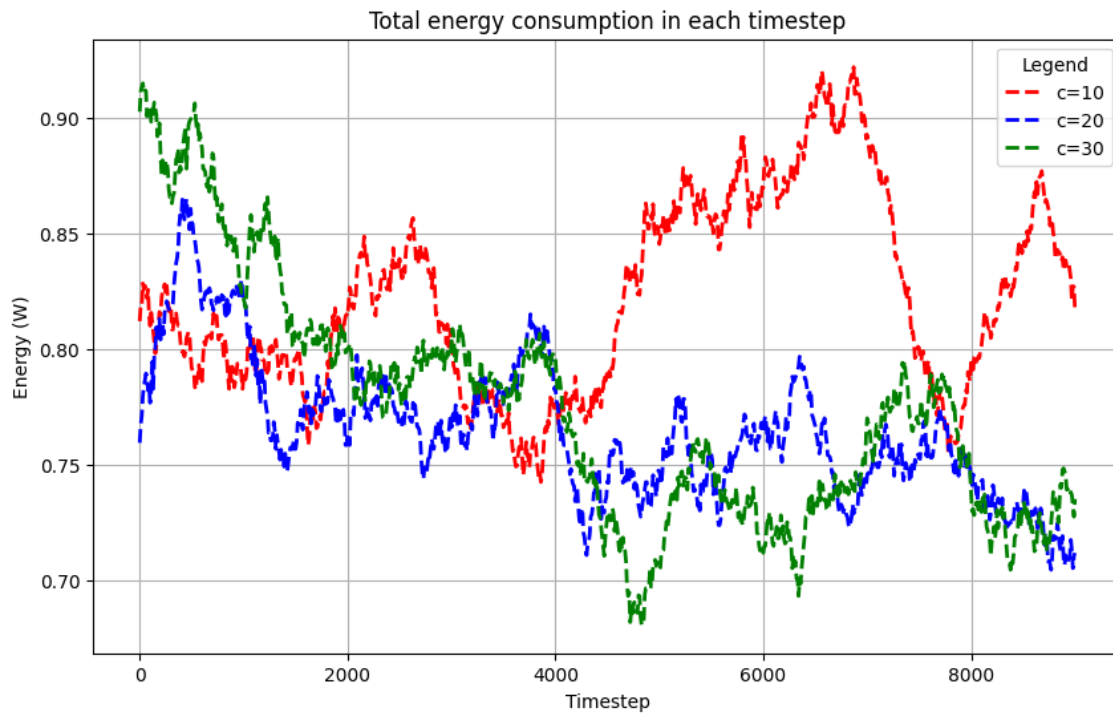
Σχήμα 5.7: Πείραμα 4 - Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (μεταβολή του παράγοντα εξερεύνησης).

αντίστοιχου όρου, σταθμισμένου μέσω ενός συντελεστή c . Σε αυτό το πείραμα αναλύεται η συμπεριφορά του συστήματος στο μη ανατροφοδοτούμενο περιβάλλον για διαφορετικές τιμές του c . Υποθέτουμε ότι οι δορυφόροι εκπέμπουν με ισχύ $P = 50W$, ενώ οι τιμές των υπόλοιπων παραμέτρων λαμβάνονται ίδιες με το πείραμα της ενότητας 5.2.1.

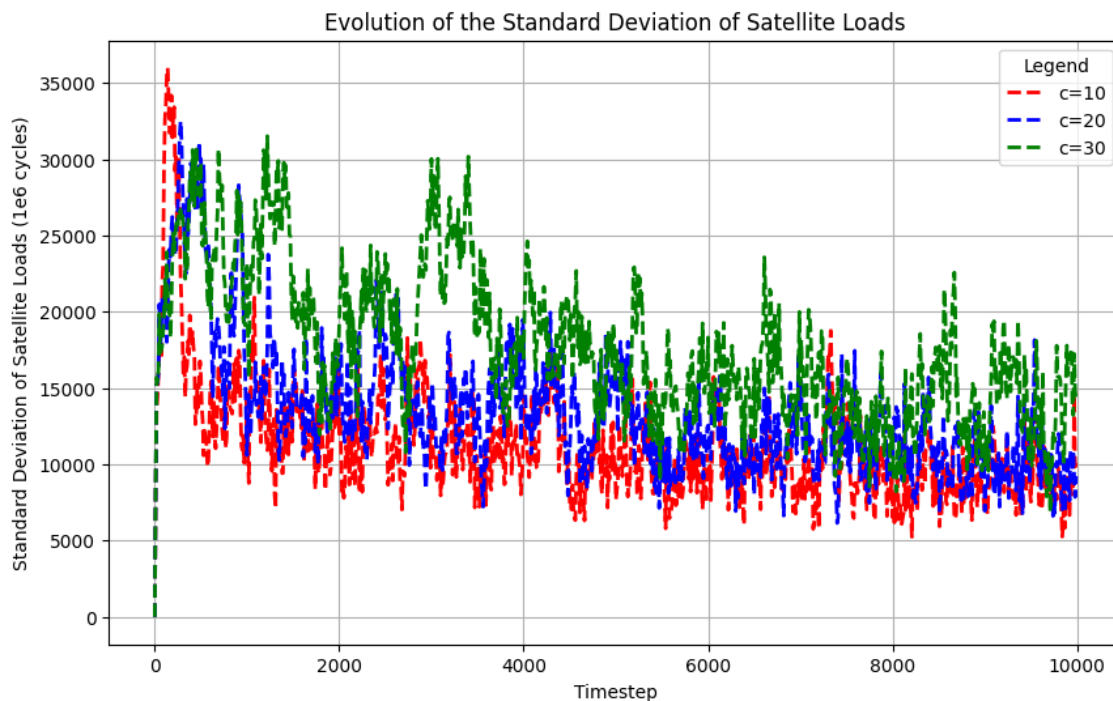
Η στάθμιση της ενεργειακής απόδοσης με έναν υψηλό συντελεστή καθοδηγεί τους δορυφόρους στην επιλογή μιας στρατηγικής που θα μειώσει τη συνολική ενέργεια που καταναλώνεται σε κάθε γύρο για τις αντίστοιχες εκφορτώσεις. Πράγματι, όπως φαίνεται στο Σχήμα 5.8, η στάθμιση με την τιμή $c = 10$ συντελεί στην επιλογή μιας ενεργειακά κοστοβόρας στρατηγικής, ενώ η στάθμιση με υψηλότερες τιμές ($c = 20, 30$) συμβάλλει στη μείωση της καταναλισκόμενης ενέργειας κατά τη διάρκεια της εκμάθησης.

Βέβαια, η εκμάθηση ενεργειακά αποδοτικών στρατηγικών πραγματοποιείται σε βάρος της εκμάθησης στρατηγικών που εξισορροπούν βέλτιστα τα φορτία. Το Σχήμα 5.9 φανερώνει αυτή την εξάρτηση: η αύξηση του συντελεστή c οδηγεί το σύστημα σε σημεία ισορροπίας με αυξημένη τυπική απόκλιση μεταξύ των φορτίων των δορυφόρων.

Δεδομένου αυτού του συμβιβασμού, τεχνο-οικονομικοί παράγοντες πρέπει να ληφθούν υπόψη για την επιλογή του κατάλληλου συντελεστή ενεργειακής αποδοτικότητας με βάση τις πραγματικές προτεραιότητες.



Σχήμα 5.8: Πείραμα 5 - Συνολική ενέργεια που καταναλώνεται σε κάθε γύρο λόγω των εκφορτώσεων (μεταβολή του συντελεστή ενεργειακής απόδοσης).



Σχήμα 5.9: Πείραμα 5 - Τυπική απόκλιση των φορτίων στις ουρές των δορυφόρων με την πάροδο του χρόνου (μεταβολή του συντελεστή ενεργειακής απόδοσης).

Σύνοψη και Συμπεράσματα

Η παρούσα διπλωματική εργασία αντιμετώπισε την ανάγκη εξισορρόπησης φορτίου σε περιβάλλοντα Δορυφορικών Υποδομών Υπολογιστικής Παρυφής, μέσω της εκφόρτωσης διεργασιών μεταξύ δορυφόρων. Διατυπώθηκε ένα πρόβλημα απόφασης για κάθε δορυφόρο, ο οποίος πρέπει να ορίσει μια κατανομή πιθανότητας επιλογής όλων των άλλων δορυφόρων ως στόχο του σε κάθε γύρο, καθώς και το ποσό που θα εκφορτώσει σε καθέναν από αυτούς. Ο δεύτερος στόχος επιλύθηκε μέσω μιας απλής ντετερμινιστικής στρατηγικής που βασίζεται στα τρέχοντα φορτία των ουρών των εξυπηρετητών των δορυφόρων. Ο πρώτος μοντελοποιήθηκε χρησιμοποιώντας α) ένα στοχαστικό παίγνιο και β) ένα παίγνιο δυναμικού. Ενώ το στοχαστικό παίγνιο χαρακτηρίζεται από διαισθητική απλότητα, δεν προσφέρει εγγυήσεις σύγκλισης. Αντίθετα, το παίγνιο δυναμικού προσφέρει εγγυήσεις σύγκλισης για μη ανατροφοδοτούμενα περιβάλλοντα (δ-ισορροπίες) και ακόμη ισχυρότερες εγγυήσεις για τα ημιανατροφοδοτούμενα περιβάλλοντα (ισορροπίες Nash), χρησιμοποιώντας την παραλλαγή ε-Hedge του Αλγορίθμου Εκθετικών Βαρών (EWA). Κάθε ένα από αυτά τα παίγνια μπορεί να εμπλουτιστεί με έναν όρο που σχετίζεται με την καταναλισκόμενη ενέργεια στη συνάρτηση απόδοσής τους, προκειμένου να σταθμιστεί ο στόχος της εξισορρόπησης φορτίου με την ενεργειακή απόδοση σε πραγματικές εφαρμογές.

Οι προσομοιώσεις αποδεικνύουν ότι τόσο το στοχαστικό παίγνιο όσο και το παίγνιο δυναμικού (μη ανατροφοδοτούμενο/ημιανατροφοδοτούμενο) παρουσιάζουν γρήγορη σύγκλιση και συνολική εξισορρόπηση του φορτίου του συστήματος. Όλες οι προσεγγίσεις αποδίδουν συγκρίσιμα καλά, οπότε η μη ανατροφοδοτούμενη περίπτωση του παιγνίου δυναμικού ξεχωρίζει λόγω των ισχυρών επιδόσεων, των εγγυήσεων σύγκλισης και της πλήρως αποκεντρωμένης, κλιμακούμενης φύσης της. Περαιτέρω αξιολογήσεις του εν λόγω παιγνίου επιβεβαιώνουν ότι η εξισορρόπηση φορτίου επιτυγχάνεται ακόμη και υπό διαφορετικό αριθμό δορυφόρων και υψηλότερους ρυθμούς παραγωγής διεργασιών, υπό την προϋπόθεση ότι οι τελευταίοι δεν υπερβαίνουν τη συνολική χωρητικότητα του συστήματος. Η μεταβαλλόμενη επιλογή του συντελεστή ε κατέδειξε το συμβιβασμό εξερεύνησης-εκμετάλλευσης, ενώ η ενεργειακά αποδοτική εκδοχή του παιγνίου κατέστησε φανερό το συμβιβασμό εξισορρόπησης φορτίου-ενεργειακής κατανάλωσης.

Εν ολίγοις, παρουσιάστηκε ένα πλήρως λειτουργικό πλαίσιο εξισορρόπησης φορτίου μεταξύ δορυφόρων, το οποίο λαμβάνει υπόψη πραγματικούς τηλεπικοινωνιακούς και ενεργειακούς

περιορισμούς, προσφέροντας μια πλήρως αποκεντρωμένη, κλιμακούμενη μέθοδο με εγγυήσεις σύγκλισης σε παιγνιο-θεωρητικές ισορροπίες. Φυσικά, η εξισορρόπηση φορτίου αποτελεί μόνο έναν από τους πολλαπλούς παράλληλους στόχους που θα θέλαμε να επιτύχουμε μέσω της διαδορυφορικής επικοινωνίας. Η μοντελοποίηση και άλλων τέτοιων στόχων με πολυκριτήρια λογική, σε μια πιο εκτεταμένη διατύπωση, η οποία για παράδειγμα θα λαμβάνει υπόψη τη φυσική μοντελοποίηση του διαδορυφορικού καναλιού, καθώς και τη δορυφορική τοπολογία, θα μπορούσε να προσφέρει ένα ακόμη πιο εξελιγμένο σχήμα επικοινωνίας/εκφόρτωσης και αποτελεί μια υποσχόμενη μελλοντική ερευνητική κατεύθυνση.

Βιβλιογραφία

- [ACS24] S. V. Albrecht, F. Christianos και L. Schäfer. *Multi-Agent Reinforcement Learning: Foundations and Modern Approaches*. Cambridge, MA: MIT Press, 2024.
- [CHM17] J. Cohen, A. Héliou και P. Mertikopoulos. «Learning with bandit feedback in potential games». Στο: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*. Τόμ. 30. 2017, σσ. 6372–6381.
- [FE25] F. Fardno και R. Etesami. *A game-theoretic framework for distributed load balancing: static and dynamic game models*. 2025. DOI: 10.48550/arXiv.2501.15324. arXiv: 2501.15324 [cs.GT].
- [Flo+20] L. Flokas, E. V. Vlatakis-Gkaragkounis, T. Lianas, P. Mertikopoulos και G. Piliouras. «No-regret learning and mixed Nash equilibria: they do not mix». Στο: *Proceedings of the 34th International Conference on Neural Information Processing Systems (NeurIPS)*. Τόμ. 33. 2020, σσ. 1380–1391.
- [Gao+24a] Y. Gao, Z. Ji, K. Zhao, T. de Cola και W. Li. «Game-Based Computation Offloading and Power Allocation for LEO Constellation Networks in Distributed and Dynamic Environment». Στο: *IEEE Internet of Things Journal* 11.4 (Φεβ. 2024), σσ. 7040–7058. DOI: 10.1109/JIOT.2023.3314650.
- [Gao+24b] Y. Gao, K. Zhao, T. de Cola, W. Li, Y. Cui και P. Hu. «Multi-User Computation Offloading in Dynamic Environment for LEO Constellation Networks». Στο: *IEEE Transactions on Vehicular Technology* 73.3 (Μαρ. 2024), σσ. 4453–4458. DOI: 10.1109/TVT.2023.3327308.
- [GC05] D. Grosu και A. T. Chronopoulos. «Noncooperative load balancing in distributed systems». Στο: *Journal of Parallel and Distributed Computing* 65 (2005), σσ. 1022–1034. DOI: 10.1016/j.jpdc.2005.05.001.
- [GCL02] D. Grosu, A. T. Chronopoulos και M.-Y. Leung. «Load balancing in distributed systems: an approach using cooperative games». Στο: *Proceedings 16th International Parallel*

- and Distributed Processing Symposium*. Ft. Lauderdale, FL, USA, 2002, 10 pp-. doi: 10.1109/IPDPS.2002.1015536.
- [Haz22] E. Hazan. *Introduction to Online Convex Optimization*. 2nd. Cambridge, MA: MIT Press, 2022.
- [Hu+20] J. Hu, K. Li, C. Liu και K. Li. «Game-based task offloading of multiple mobile devices with QoS in mobile edge computing systems of limited computation capacity». Στο: *ACM Transactions on Embedded Computing Systems* 19 (2020), σσ. 1–21. doi: 10.1145/3398038.
- [Jia+22] M. Jia, L. Zhang, J. Wu, Q. Guo και X. Gu. «Joint computing and communication resource allocation for edge computing towards Huge LEO networks». Στο: *China Communications* 19.8 (Αύγ. 2022), σσ. 73–84. doi: 10.23919/JCC.2022.08.006.
- [Lin+25] B. Lin, J. Weng, X. Chen, Y. Ma και C.-H. Hsu. «A Game-Based Computation Offloading With Imperfect Information in Multi-Edge Environments». Στο: *IEEE Transactions on Services Computing* 18.1 (Ιαν. 2025), σσ. 1–14. doi: 10.1109/TSC.2024.3517336.
- [Nis+07] N. Nisan, T. Roughgarden, É. Tardos και V. V. Vazirani. *Algorithmic Game Theory*. New York, NY: Cambridge University Press, 2007.
- [OR94] M. J. Osborne και A. Rubinstein. *A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.
- [Sha12] S. Shalev-Shwartz. *Online Learning and Online Convex Optimization*. Hanover, MA: Now Publishers, 2012.
- [Shi+24] J. Shi, D. Lv, T. Chen και Y. Li. «Learning-Based Inter-Satellite Computation Offloading in Satellite Edge Computing». Στο: *2024 9th International Conference on Signal and Image Processing (ICSIP)*. Nanjing, China, 2024, σσ. 476–480. doi: 10.1109/ICSIP61881.2024.10671510.
- [Sol22] E. Solan. *A Course in Stochastic Game Theory*. Cambridge, UK: Cambridge University Press, 2022.
- [Tad13] S. Tadelis. *Game Theory: An Introduction*. Princeton, NJ: Princeton University Press, 2013.
- [Ton+23] M. Tong, S. Li, X. Wang και P. Wei. «Inter-satellite cooperative offloading decision and resource allocation in mobile edge computing-enabled satellite–terrestrial networks». Στο: *Sensors* 23.2 (Ιαν. 2023), σ. 668. doi: 10.3390/s23020668.
- [Wan+20] Y. Wang, J. Yang, X. Guo και Z. Qu. «A Game-Theoretic Approach to Computation Offloading in Satellite Edge Computing». Στο: *IEEE Access* 8 (2020), σσ. 12510–12520. doi: 10.1109/ACCESS.2019.2963068.
- [Wu+24] G. Wu, H. Wang, H. Zhang, Y. Shen, S. Shen και S. Yu. «Mean-Field Game-Based Task-Offloaded Load Balance for Industrial Mobile Edge Computing Systems Using Software-Defined Networking». Στο: *IEEE Transactions on Mobile Computing* 23.12 (Δεκ. 2024), σσ. 13773–13786. doi: 10.1109/TMC.2024.3437761.

- [YCM24] M. Yahya, A. Conzelmann και S. Maghsudi. «Decentralized Task Offloading and Load-Balancing for Mobile Edge Computing in Dense Networks». Στο: *IEEE Communications Letters* 28.8 (Αύγ. 2024), σσ. 1954–1958. doi: 10.1109/LCOMM.2024.3416833.
- [Zha+21] F. Zhang, R. Deng, X. Zhao και M. M. Wang. «Load Balancing for Distributed Intelligent Edge Computing: A State-Based Game Approach». Στο: *IEEE Transactions on Cognitive Communications and Networking* 7.4 (Δεκ. 2021), σσ. 1066–1077. doi: 10.1109/TCCN.2021.3087178.

Απόδοση

Αντίπαλος
 Αλγόριθμος Εκθετικών Βαρών
 Αλγόριθμος Πολλαπλασιαστικής
 Ενημέρωσης Βαρών
 Απώλειες Διάδοσης Ελεύθερου Χώρου
 Διαδορυφορικές Ζεύξεις
 Διαδικασία Απόφασης Markov
 Διασυνδεδεμένοι Εξυπηρετητές Παρυφής
 Διαδίκτυο των Πραγμάτων
 Δικαιοσύνη
 Δορυφόροι Χαμηλής Τροχιάς
 Δορυφορικές Υποδομές
 Υπολογιστικής Παρυφής
 Εκτιμητής Χωρίς Ανατροφοδότηση
 Εκφόρτωση Υπολογισμών
 Εξερεύνηση-Εκμετάλλευση
 Εξισορρόπηση Φορτίου
 Εικονικό Παίξιμο
 Εξωτερική Μετάνοια
 Επίγεια / Δορυφορικά Δίκτυα
 Ημιανατροφοδοτούμενο Περιβάλλον
 Θεωρία Παιγνίων
 Θεωρία Παιγνίων Μέσου Πεδίου
 Καθαρή Στρατηγική
 Κανονική Μορφή
 Κέρδος Κεραίας
 Λύση Διαπραγματεύσεως Nash

Ξενόγλωσσος όρος

Adversary
 Exponential Weights Algorithm (EWA)
 Multiplicative Weights Update (MWU)
 Free-Space Path Loss (FSPL)
 Inter-Satellite Links (ISLs)
 Markov Decision Process (MDP)
 Intelligent Edge Servers (IES)
 Internet of Things (IoT)
 Fairness
 Low Earth Orbit (LEO) Satellites
 Satellite Edge Computing (SEC)
 Bandit Estimator
 Task Offloading
 Exploration-Exploitation
 Load Balancing
 Fictitious Play
 External Regret
 Terrestrial / Non-Terrestrial Networks
 Semi-Bandit Framework
 Game Theory
 Mean-Field Game Theory
 Pure Strategy
 Normal Form
 Antenna Gain
 Nash Bargaining Solution (NBS)

Λύση Κυρίαρχης Στρατηγικής	Dominant Strategy Solution
Μέθοδος "Ακολούθησε-τον-Ηγέτη"	Follow the Leader (FTL)
Μέθοδος "Ακολούθησε-τον-Κανονικο- ποιημένο-Ηγέτη"	Follow the Regularized Leader (FTRL)
Μεικτή Στρατηγική	Mixed Strategy
Μετανοητική Μάθηση	No-Regret Learning
Μη Ανατροφοδοτούμενο Περιβάλλον	Bandit Framework
Πάιγνια Δυναμικού	Potential Games
Προσθετικός Λευκός Γκαουσιανός Θόρυβος	Additive White Gaussian Noise (AWGN)
Συνάρτηση Απόδοσης	Utility Function
Συνάρτηση Δυναμικού	Potential Function
Υπολογιστικές Υποδομές Παρυφής	Mobile Edge Computing (MEC)
Χάρτης Πιθανοτικής Επιλογής	Logit Choice Map

