

Πανεπιστήμιο Δυτικής Μακεδονίας
Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών
Υπολογιστών

Αλγόριθμοι ενισχυτικής μάθησης για το
πρόβλημα χωροθέτησης μονάδων

Γεώργιος Τσακίρης (ΑΜ: 1242)
Επιβλέπων Καθηγητής: Νικόλαος Πλόσκας

Εργαστήριο Ευφρών Συστημάτων & Βελτιστοποίησης
11 Οκτωβρίου 2021

Περίληψη

Το Πρόβλημα Χωροθέτησης Μονάδων εμφανίστηκε για πρώτη φορά τη δεκαετία του 1970. Ανήκει στον κλάδο της επιστήμης της χωροταξίας και έχει ως στόχο την εύρεση της βέλτιστης τοποθεσίας εγκατάστασης μιας ή περισσότερων μονάδων. Λόγω της ραγδαίας εξέλιξης της Επιστήμης των Υπολογιστών και της Τεχνητής Νοημοσύνης, τα τελευταία χρόνια παρουσιάζεται μεγάλο ενδιαφέρον στην εύρεση νέων μεθόδων επίλυσης του προβλήματος.

Στην ενισχυτική μάθηση, ένας πράκτορας εκπαιδεύεται βάση ερεθισμάτων που λαμβάνει από το περιβάλλον για να ανακαλύψει τη βέλτιστη πολιτική που θα τον οδηγήσει στην επίτευξη του στόχου του. Με αυτόν τον τρόπο επιτυγχάνεται ταυτόχρονη εξερεύνηση και εκμετάλλευση του χώρου. Αυτός είναι άλλωστε και ο λόγος για τον οποίο η ενισχυτική μάθηση χρησιμοποιείται για την επίλυση προβλημάτων βελτιστοποίησης όπως είναι το πρόβλημα χωροθέτησης μονάδων.

Σε αυτήν τη διπλωματική επικεντρωνόμαστε στην επίλυση του προβλήματος χωροθέτησης μονάδων μέσω της ενισχυτικής μάθησης. Σκοπός της παρούσας διπλωματικής είναι η ανάλυση αλγορίθμων ενισχυτικής μάθησης ως προς την ακρίβεια των αποτελεσμάτων και των χρόνων εκπαίδευσής και υλοποίησής τους. Στα πλαίσια της διπλωματικής πραγματοποιείται σύγκριση των αλγορίθμων σε μια πληθώρα προβλημάτων.

Λέξεις κλειδιά: Πρόβλημα Χωροθέτησης Μονάδων, Αλγόριθμοι, Μηχανική Μάθηση, Ενισχυτική Μάθηση, Ευρετικοί Αλγόριθμοι

Abstract

The Facility Location Problem was first introduced in the 1970s. It belongs to the field of spatial planning science and aims to find the optimal location for one or more units. In recent years, due to the rapid development of Computer Science and Artificial Intelligence, there has been great interest in finding new methods to solve the problem.

In reinforcement learning, an agent is trained on the basis of the experiences he receives from the environment he interacts with, to discover the best policy that will lead him to achieve his goal. In this way, simultaneous exploration and exploitation of space is achieved. This is also the reason why reinforcement learning is used to solve various combinatorial optimization problems, such as the facility location problem.

In this thesis, we focus on solving the facility location problem using reinforcement learning algorithms. The purpose of this thesis is the analysis of reinforcement learning algorithms in terms of the accuracy of the results and the times of their training and evaluation. In addition to that, the algorithms are compared in a variety of problems.

Keywords: Facility Location Problem, Algorithms, Machine Learning, Reinforcement Learning, Heuristics

Δήλωση Πνευματικών Δικαιωμάτων

Δήλωση Πνευματικών Δικαιωμάτων Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα Διπλωματική Εργασία με τίτλο "Αλγόριθμοι ενισχυτικής μάθησης για το πρόβλημα χωροθέτησης μονάδων" καθώς και τα ηλεκτρονικά αρχεία και πηγαίοι κώδικες που αναπτύχθηκαν ή τροποποιήθηκαν στα πλαίσια αυτής της εργασίας και αναφέρονται ρητώς μέσα στο κείμενο που συνοδεύουν, και η οποία έχει εκπονηθεί στο Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών του Πανεπιστημίου Δυτικής Μακεδονίας, υπό την επίβλεψη του μέλους του Τμήματος κ. Νικόλαου Πλόσκα αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο, αρχεία ή / και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή.

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 μόνο.

Copyright (C) Γεώργιος Τσακίρης & Νικόλαος Πλόσκας, 2021, Κοζάνη

Υπογραφή Φοιτητή

Περιεχόμενα

1	Εισαγωγή	10
1.1	Ορισμός του Προβλήματος	10
1.2	Κίνητρα και Στόχοι Υλοποίησης	10
1.3	Διάρθρωση Κειμένου	11
2	Πρόβλημα Χωροθέτησης Μονάδων	12
2.1	Εισαγωγή	12
2.2	Συναρτήσεις Απόστασης	14
2.3	Επεκτάσεις του Προβλήματος Χωροθέτησης Μονάδων	15
2.4	Μέθοδοι Επίλυσης του Προβλήματος Χωροθέτησης Μονάδων	17
3	Βιβλιογραφική Ανασκόπηση	20
3.1	Μοντέλο Ενισχυτικής Μάθησης	20
3.2	Εφαρμογή Μοντέλου Βαθιάς Ενισχυτικής Μάθησης στο Πρόβλημα K-Server	22
3.3	Αύξηση Γράφου για Ίση Πρόσβαση με τη Χρήση Ενισχυτικής Μάθησης	24
3.4	Επίλυση του Προβλήματος Συνδυαστικής Βελτιστοποίησης με τη Χρήση Ενισχυτικής Μάθησης Πολλαπλών Πρακτόρων	28
3.5	Μοντέλο Βελτιστοποίησης της Εγγύησης Προσβασιμότητας για την Τοποθέτηση Μονάδων Έκτακτης Ανάγκης με Ενίσχυση Συνδέσμων . .	30
3.6	Εφαρμογή Μεθόδου Βελτιστοποίησης Πολλαπλών Πρακτόρων για το Πρόβλημα Τετραγωνικής Ανάθεσης	33
3.7	Βελτίωση της Εκπαίδευσης των Ακμών με τη Χρήση Ενισχυτικής Μάθησης	37
3.8	Μάθηση από Άκρο σε Άκρο και Βελτιστοποίηση Γράφων	39

4	Ανάλυση Μοντέλων	45
4.1	CLUSTERNET System	45
4.2	Equitable Mechanism Design in Markov Reward Process (EMD-MRP) .	49
5	Παρουσίαση Αποτελεσμάτων	55
5.1	Συγκρινόμενα Μοντέλα	55
5.1.1	Μοντέλο Clusternet	55
5.1.2	Μοντέλο EMD-MRP	56
5.2	Σύγκριση Μοντέλων Ενισχυτικής Μάθησης	58
5.2.1	Εκπαίδευση	58
5.2.2	Αξιολόγηση	60
5.2.3	Σύνοψη	64
6	Συμπεράσματα	69

Κατάλογος σχημάτων

3.1	Απεικόνιση μοντέλου ενισχυτικής μάθησης	21
3.2	Γραφική παράσταση του προϋπολογισμού ως προς το δείκτη Gini [57]	26
3.3	Απεικόνιση καλύτερης μέσης τιμής για το ftv44 ως προς το λ [48] . . .	30
3.4	Δικτύου μεταφοράς του Sioux Falls [74]	32
4.1	Απεικόνιση μοντέλου CLUSTERNET και Two-Stage	46
5.1	Μέση απόσταση μοντέλου Clusternet για τα σύνολα δεδομένων Cora, Citeseer, Protain Vidal, 15 nodes	57
5.2	Μέση απόσταση μοντέλου Clusternet για τους γράφους του συνόλου δεδομένων Synthetic SPA	57
5.3	Μέση απόσταση μοντέλου EMD-MRP για τα σύνολα δεδομένων Cora, Citeseer, Protain Vidal, 15 nodes	59
5.4	Μέση απόσταση μοντέλου EMD-MRP για τους γράφους του συνόλου δεδομένων Synthetic SPA	59
5.5	Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στα σύνολα δεδομένων Cora, Citeseer, Protain Vidal για μια μονάδα	60
5.6	Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στους γράφους του συνόλου δεδομένων Synthetic SPA, 15 nodes για μια μονάδα	61
5.7	Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στα σύνολα δεδομένων Cora, Citeseer, Protain Vidal για δύο μονάδες	61
5.8	Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στους γράφους του συνόλου δεδομένων Synthetic SPA, 15 nodes για δύο μονάδες	62
5.9	Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στα σύνολα δεδομένων Cora, Citeseer, Protain Vidal για τρεις μονάδες	62
5.10	Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στους γράφους του συνόλου δεδομένων Synthetic SPA, 15 nodes για τρεις μονάδες	63

5.11 Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στα σύνολα δεδομένων Cora, Citeseer, Protain Vidal για τέσσερις μονάδες	63
5.12 Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στους γράφους του συνόλου δεδομένων Synthetic SPA, 15 nodes για τέσσερις μονάδες	64
5.13 Μέση απόσταση μοντέλου Clusternet και EMD-MDP ανά σύνολο δεδομένων για μια μονάδα	65
5.14 Μέση απόσταση μοντέλου Clusternet και EMD-MDP ανά σύνολο δεδομένων για δύο μονάδες	65
5.15 Μέση απόσταση μοντέλου Clusternet και EMD-MDP ανά σύνολο δεδομένων για τρεις μονάδες	66
5.16 Μέση απόσταση μοντέλου Clusternet και EMD-MDP ανά σύνολο δεδομένων για τέσσερις μονάδες	66

Κατάλογος αλγορίθμων

1	Άπληστη βελτίωση της ισότητας των κέντρων	51
2	Σχεδιασμός μηχανισμού ισότητας μέσω της Μαρκοβιανής διαδικασίας κέρδους	53

Κατάλογος πινάκων

3.1	Παρουσίαση απόδοσης για διαφορετικό αριθμό διαστάσεων [44]	23
3.2	Δείκτης Gini [57]	27
3.3	Λύσεις για διαφορετικές τιμές των G , ω [74]	43
3.4	Εκτίμηση της πολιτικής ανά πρόβλημα [38]	43
3.5	Επίδοση στο πρόβλημα ανίχνευσης κοινότητας [71]	44
3.6	Επίδοση στο πρόβλημα χωροθέτησης μονάδων [71]	44
3.7	Σύγκριση επιδόσεων ως προς την τεχνική Finetune [71]	44
5.1	Πίνακας αποτελεσμάτων μέσης απόστασης, χρόνων εκτέλεσης και εκ- παίδευσης των μοντέλων ανά σύνολο δεδομένων	67

Κεφάλαιο 1

Εισαγωγή

1.1 Ορισμός του Προβλήματος

Το Πρόβλημα Χωροθέτησης Μονάδων (Facility Location Problem - FLP) αποτελεί ένα κλασικό πρόβλημα βελτιστοποίησης που έχει ως στόχο την εύρεση της/των καλύτερης/ων θέσης/εων για την τοποθέτηση μιας/πολλών μονάδας/ων από ένα σύνολο διαθέσιμων τοποθεσιών. Η επιλογή γίνεται βάση διαφόρων κριτηρίων, όπως οι γεωγραφικές απαιτήσεις, το κόστος ανοικοδόμησης και οι αποστάσεις μεταφοράς. Τα προβλήματα χωροθέτησης μονάδων βρίσκουν εφαρμογή στον ιδιωτικό τομέα στοχεύοντας στην αύξηση του κέρδους καθώς και στον δημόσιο τομέα για την ίση και δίκαιη πρόσβαση σε αγαθά και υπηρεσίες.

Παραδείγματα εφαρμογών του προβλήματος αποτελούν η τοποθέτηση νέων αποθηκών, εργοστασίων, τραπεζών, εμπορικών καταστημάτων, στάσεων μέσων μαζικής μεταφοράς, νοσοκομείων και σχολείων. Εκτός από καινούργιες μονάδες, το πρόβλημα χωροθέτησης μονάδων βρίσκει εφαρμογή και στη μεταφορά υπαρχόντων μονάδων.

1.2 Κίνητρα και Στόχοι Υλοποίησης

Το πρόβλημα χωροθέτησης μονάδων ανήκει στην κατηγορία συνδυαστικών προβλημάτων NP-hard. Σε αυτήν, δεν υπάρχει αλγόριθμος που να εγγυάται επίλυση σε πολυωνυμικό χρόνο. Λόγω αυτού, η επίλυση με γραμμικό προγραμματισμό μέσω περιορισμών καθιστά αδύνατη την εύρεση λύσης για μεγάλο αριθμό κόμβων σε σύντομο χρονικό διάστημα.

Η ενισχυτική μάθηση αποτελεί μία κατηγορία της μηχανικής μάθησης η οποία

εφαρμόζεται για την επίλυση σύνθετων προβλημάτων. Στην παρούσα διπλωματική εργασία παρουσιάζονται μοντέλα ενισχυτικής μάθησης για την επίλυση του προβλήματος χωροθέτησης μονάδων καθώς και η ανάλυσή τους ως προς την ελαχιστοποίηση του χρόνου επίλυσης.

1.3 Διάρθρωση Κειμένου

Τα κεφάλαια της διπλωματικής εργασίας περιγράφονται ως εξής: Στο κεφάλαιο 2 παρουσιάζεται το πρόβλημα χωροθέτησης μονάδων, επεκτάσεις του καθώς και μέθοδοι επίλυσής του. Στο Κεφάλαιο 3 παρουσιάζονται οι βασικές έννοιες της ενισχυτικής μάθησης καθώς και μοντέλα που έχουν προταθεί για την επίλυση του προβλήματος με τη συγκεκριμένη μέθοδο. Στο Κεφάλαιο 4 παρουσιάζονται εκτενώς δύο μοντέλα ενισχυτικής μάθησης, τα οποία χρησιμοποιούνται στην υπολογιστική μελέτη της παρούσας διπλωματικής εργασίας. Στο Κεφάλαιο 5 παρουσιάζονται τα αποτελέσματα των δύο μοντέλων καθώς και τα συμπεράσματα που προκύπτουν από τη σύγκρισή τους. Τέλος, στο Κεφάλαιο 6 παρουσιάζονται τα συμπεράσματα της παρούσας εργασίας.

Κεφάλαιο 2

Πρόβλημα Χωροθέτησης Μονάδων

2.1 Εισαγωγή

Το Πρόβλημα Χωροθέτησης Μονάδων (Facility Location Problem - FLP) ανήκει στον κλάδο της επιστήμης της χωροταξίας (Location Science). Ξεκίνησε να μελετάται στις αρχές της δεκαετίας του 1970 ενώ αναλύθηκε εκτενώς τη δεκαετία του 1980 [17]. Στόχος του προβλήματος αποτελεί η εύρεση ενός συνόλου θέσεων για την τοποθέτηση μονάδων σε έναν γνωστό χώρο.

Η επιλογή των θέσεων εξαρτάται από διάφορους παράγοντες, όπως οι γεωγραφικές απαιτήσεις, το κόστος ανοικοδόμησης και οι αποστάσεις μεταφοράς που προκύπτουν από την κατανομή των σημείων ζήτησης. Στη συνέχεια δίνονται τα βασικά στοιχεία των προβλημάτων χωροθέτησης μονάδων σύμφωνα με τους Eiselt και Laporte [22]:

- Ο χώρος στον οποίο θα τοποθετηθούν οι μονάδες
- Το πλήθος των μονάδων που θα τοποθετηθούν καθώς και αν υπάρχουν ήδη ή είναι μόνο καινούριες
- Η ζήτηση των πελατών
- Η αλληλεπίδραση μεταξύ πελατών και μονάδων
- Η συνάρτηση που χρησιμοποιείται για τον υπολογισμό της απόστασης μεταξύ πελατών και μονάδων
- Επιπλέον περιορισμοί που πρέπει να ικανοποιηθούν

Το πρόβλημα διακρίνεται σε δύο βασικές κατηγορίες, οι οποίες περιγράφονται παρακάτω.

Πρόβλημα Χωροθέτησης Μονάδων βάση Χωρητικότητας (Capacitated Facility Location Problem - CFLP): Αποτελεί τη βάση σε πολλά προβλήματα βελτιστοποίησης όπου η συνολική ζήτηση κάθε μονάδας πληροί κάποιες προϋποθέσεις. Στη συνέχεια περιγράφεται η μοντελοποίηση του προβλήματος.

$$\begin{aligned}
 \min \quad & Z = \sum_{j=1}^m f_j y_j + \sum_{i=1}^n \sum_{j=1}^m c_{ij} x_{ij} \\
 \text{s.t.} \quad & \sum_{j=1}^m x_{ij} = d_i && \text{for } i = 1, \dots, n \\
 & \sum_{i=1}^n x_{ij} \leq M_j y_j && \text{for } j = 1, \dots, m \\
 & x_{ij} \leq d_i y_j && \text{for } i = 1, \dots, n; j = 1, \dots, m \\
 & x_{ij} \geq 0 && \text{for } i = 1, \dots, n; j = 1, \dots, m \\
 & y_j \in \{0, 1\} && \text{for } j = 1, \dots, m
 \end{aligned}$$

Το i δηλώνει τον αριθμό των πελατών, ενώ το j τον αριθμό των μονάδων. Επιπλέον, το x_{ij} συμβολίζει το ποσό που έχει μεταφερθεί από τη μονάδα j στο σημείο i , το f_j δηλώνει το κόστος κατασκευής της μονάδας j , ενώ το c_{ij} περιγράφει το κόστος μεταφοράς από το σημείο j στο i . Το M δηλώνει τον περιορισμό ως προς το χώρο, ενώ το d_i τη ζήτηση του πελάτη i . Τέλος, το y_j περιγράφει έναν δείκτη βάση του οποίου δηλώνεται η τοποθέτηση μιας μονάδας.

Πρόβλημα Χωροθέτησης Μονάδων χωρίς τον Περιορισμό της Χωρητικότητας (Uncapacitated Facility Location Problem - UFLP): Μπορεί να χαρακτηριστεί ως ένα χαλαρό CFLP. Στη συνέχεια δίνεται η μαθηματική έκφραση του προβλήματος για μεγάλες τιμές του M . Σε αυτή, το M δηλώνει τον περιορισμό ως προς το χώρο.

$$\sum_{i=1}^n x_{ij} \leq M_j y_j \quad \text{for } j = 1, \dots, m$$

Στην ενότητα 2.2 αναλύονται διάφορες τεχνικές υπολογισμού της απόστασης. Στην ενότητα 2.3 αναφέρονται επεκτάσεις του προβλήματος τοποθέτησης μονάδων. Τέλος, στην ενότητα 2.4 αναλύονται τεχνικές επίλυσης του προβλήματος.

2.2 Συναρτήσεις Απόστασης

Βασικό χαρακτηριστικό των προβλημάτων χωροθέτησης μονάδων αποτελεί η συνάρτηση απόστασης μεταξύ των μονάδων και των πελατών. Ο Ogryczak [55] παρατήρησε ότι η λύση των προβλημάτων χωροθέτησης μονάδων σχετίζεται με την ελαχιστοποίηση ή τη μεγιστοποίηση της μέσης απόστασης. Στη συνέχεια δίνονται οι πιο συνήθεις μοντελοποιήσεις της συνάρτησης απόστασης, όπου το $d_{k,p}(x, y)$ ορίζεται ως η συνάρτηση απόστασης μεταξύ των σημείων x και y .

Απόσταση Minkowski βαθμού p :

$$d_{k,p}(x, y) = \left(\sum_{i=1}^n k_i |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (2.1)$$

Νόρμα $d_{k,p}$ εξαρτώμενη από το βάρος:

$$d_{k,p}(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (2.2)$$

Νόρμα $d_{k,p}$ ανεξάρτητη του βάρους:

$$d_{k,p}(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (2.3)$$

Οι επεκτάσεις της Εξίσωσης 2.3 δίνονται παρακάτω.

Γραμμική απόσταση:

$$d_{k,1}(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (2.4)$$

Ευκλείδεια απόσταση:

$$d_{k,2}(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}} \quad (2.5)$$

Απόσταση Chebyshev:

$$d_{k,\infty}(x, y) = \lim_{p \rightarrow \infty} \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} = \max(|x_1 - y_1|, \dots, |x_n - y_n|) \quad (2.6)$$

2.3 Επεκτάσεις του Προβλήματος Χωροθέτησης Μονάδων

Τα προβλήματα χωροθέτησης μονάδων αποτελούν μία γενική κατηγορία, η οποία έχει μελετηθεί εκτενώς τα τελευταία χρόνια. Για αυτόν τον λόγο έχουν προταθεί διάφορες επεκτάσεις του προβλήματος ανάλογα με τους περιορισμούς που τίθενται. Στη συνέχεια αναφέρονται κάποιες από τις επεκτάσεις του προβλήματος.

Πρόβλημα Χωροθέτησης μιας Μονάδας (Single Facility Location Problem - SFLP): Το πρόβλημα έχει ως στόχο την τοποθέτηση μιας μονάδας η οποία ελαχιστοποιεί το άθροισμα των αποστάσεων μεταξύ των μονάδων και των σημείων ζήτησης. Στη συνέχεια ακολουθεί η μαθηματική της έκφραση:

$$Z_1 = \min \left\{ f(\bar{x}) : f(\bar{x}) = \sum_{i=1}^n w_i d(\bar{x}, p_i) \right\} \quad (2.7)$$

όπου το w δηλώνει το βάρος των ακμών ενώ το $d(\bar{x}, p_i)$ την απόσταση της μονάδας από τα σημεία i .

Πρόβλημα Χωροθέτησης Πολλαπλών Μονάδων (Multi-Facility Location Problem - MFLP): Η συγκεκριμένη κατηγορία προβλημάτων έχει ως στόχο την τοποθέτηση πολλαπλών μονάδων για την ελαχιστοποίηση του αθροίσματος των αποστάσεων. Αξίζει να σημειωθεί ότι στο σύνολο των μονάδων μπορεί να περιέχονται και υπάρχουσες δομές. Η μαθηματική μορφή του προβλήματος δίνεται από τη σχέση:

$$Z_2 = \min \left(\sum_{1 \leq j < k \leq n} v_{jk} d(X_j, X_k) + \sum_{j=1}^n \sum_{i=1}^m w_{ji} d(X_j, p_i) \right) \quad (2.8)$$

όπου το n δηλώνει το πλήθος των νέων μονάδων ενώ το m το πλήθος των υπαρχόντων. Το w δηλώνει το βάρος των ακμών, ενώ το $d(X_j, p_i)$ την απόσταση της μονάδας j από τα σημεία i . Τέλος, το v δηλώνει το βάρος της κάθε μονάδας.

Πρόβλημα Χωροθέτησης Μονάδων σταθερού κόστους (Fixed Costs Capacitated Facility Location Problem - FC-CFLP): Στόχος της συγκεκριμένης κατηγορίας προβλημάτων είναι η ελαχιστοποίηση του σταθερού κόστους που σχετίζεται με τη μεταφορά προϊόντων. Αυτό επιτυγχάνεται μέσω του ελάχιστου αθροίσματος των

αποστάσεων μεταξύ των μονάδων και των καταναλωτών. Η μαθηματική έκφραση του προβλήματος δίνεται από τη σχέση:

$$Z_3 = \min \left(\sum_{i \in C} \sum_{j \in F} (t_{ij}y_{ij} + c_jx_j) \right) \quad (2.9)$$

όπου το c_j δηλώνει το κόστος κατασκευής της μονάδας j ενώ το t_{ij} το κόστος μεταφοράς από το σημείο j στο i . Το x_j είναι ένας όρος που παίρνει τιμή 1 αν η μονάδα είναι εγκατεστημένη στη θέση j και 0 σε οποιαδήποτε άλλη περίπτωση.

Πρόβλημα Χωροθέτησης Μονάδων με Χωρητικότητες μέσω του Προβλήματος p-μέσων (Capacitated p-Median Facility Location Problem - CpMFLP): Σε αυτήν την κατηγορία, το σύνολο των πιθανών θέσεων τοποθέτησης ταυτίζεται με αυτή των σημείων ζήτησης. Στο πρόβλημα p-μέσων ορίζεται ένας συνδεδεμένος γράφος, ο οποίος περιέχει τα σημεία ζήτησης και τις πιθανές τοποθεσίες των μονάδων. Στόχος του προβλήματος αποτελεί ο εντοπισμός των σημείων που ελαχιστοποιούν το άθροισμα των αποστάσεων μεταξύ των πελατών και των πλησιέστερων σε αυτούς μονάδων [5]. Οι τιμές τους δίνονται μέσα από το δίκτυο του γράφου. Τα προβλήματα p-μέσων έχουν μελετηθεί σε μεγάλο βαθμό λόγω της εφαρμογής τους σε πραγματικά προβλήματα. Το CpMFLP αποτελεί μια παραλλαγή του CFLP. Σε αυτήν δε λαμβάνεται υπόψιν το σταθερό κόστος των μονάδων. Η μαθηματική έκφραση του προβλήματος δίνεται από τη σχέση:

$$Z_4 = \min \left(\sum_{i \in N} \sum_{j \in N} c_{ij}y_{ij} \right) \quad (2.10)$$

όπου το c_{ij} δηλώνει το κόστος μεταφοράς από τη θέση j στην i . Το y_{ij} είναι ένας όρος που παίρνει τιμή 1 αν ο πελάτης i έχει ανατεθεί στη μονάδα j και 0 σε οποιαδήποτε άλλη περίπτωση.

Πρόβλημα Κάλυψης Περιοχής (Covering Location Problems - CLP): Σε αυτή την κατηγορία προβλημάτων, ο πελάτης χρησιμοποιεί μια υπηρεσία η οποία παρέχεται από ένα κατάστημα. Η εξυπηρέτησή του πραγματοποιείται εφόσον η απόσταση είναι μικρότερη ή ίση μίας προκαθορισμένης τιμής. Το πρόβλημα χωρίζεται σε δύο περιπτώσεις βάση της έκτασης κάλυψης. Εφόσον καλύπτονται όλα τα σημεία ζη-

τησης, το πρόβλημα χαρακτηρίζεται ως πρόβλημα ολικής κάλυψης περιοχής. Στην περίπτωση που υπάρχουν σημεία που δεν εξυπηρετούνται από μονάδες, το πρόβλημα χαρακτηρίζεται ως μερικής κάλυψης.

Πρόβλημα Χωροθέτησης μη Επιθυμητών Μονάδων (Undesirable Facility Location Problem - Undesirable FLP): Το πρόβλημα ανήκει στην κατηγορία των μοντέλων μέγιστου αθροίσματος. Αντίθετα με το πρόβλημα p -μέσων, στόχος του προβλήματος αποτελεί η τοποθέτηση μονάδων στη μέγιστη δυνατή απόσταση από τον πληθυσμό. Η συγκεκριμένη κατηγορία χρησιμοποιείται για την τοποθέτηση μονάδων που έχουν επιπτώσεις στο περιβάλλον ή την υγεία του πληθυσμού. Η μαθηματική μορφή του προβλήματος δίνεται από τη σχέση:

$$Z_5 = \min \left(\sum_i \sum_j \sigma_i d_{ij} y_{ij} \right) \quad (2.11)$$

όπου το σ_i δηλώνει τη ζήτηση του πελάτη i ενώ το d_{ij} την απόσταση της μονάδας j από το σημείο i . Το y_{ij} είναι ένας όρος που παίρνει τιμή 1 αν καλύπτονται οι προϋποθέσεις του πελάτη i από τη μονάδα j και 0 σε οποιαδήποτε άλλη περίπτωση.

2.4 Μέθοδοι Επίλυσης του Προβλήματος Χωροθέτησης Μονάδων

Το πρόβλημα χωροθέτησης μονάδων είναι ένα πρόβλημα βελτιστοποίησης το οποίο ανήκει στην κατηγορία προβλημάτων NP-hard. Σε αυτή, δεν υπάρχει γνωστός αλγόριθμος με πολυωνυμικό χρόνο για την εύρεση της βέλτιστης λύσης. Για αυτόν τον λόγο, έχουν προταθεί στη βιβλιογραφία διάφορες τεχνικές, όπως ευρετικοί αλγόριθμοι, προσεγγιστικοί αλγόριθμοι, αλλά και ακριβείς αλγόριθμοι.

Οι ευρετικοί αλγόριθμοι παράγουν λύση σε σχετικά μικρό χρονικό διάστημα χωρίς να προσφέρουν κάποια εγγύηση. Αντίθετα, οι ακριβείς αλγόριθμοι παράγουν τη βέλτιστη εγγυημένη λύση σε μεγάλο χρόνο. Τέλος, οι προσεγγιστικοί αλγόριθμοι βρίσκουν μια λύση που προσεγγίζει τη βέλτιστη εντός πολυωνυμικού χρόνου [19]. Στη συνέχεια παρατίθενται μερικές μέθοδοι επίλυσης του προβλήματος χωροθέτησης μονάδων.

Branch-and-Bound (B&B): Αποτελεί την πιο διαδεδομένη τεχνική επίλυσης προβλημάτων NP-hard μεγάλης κλίμακας [12]. Ο αλγόριθμος B&B ερευνά το χώρο των εφικτών λύσεων ενός προβλήματος για την αναζήτηση της βέλτιστης λύσης.

Lagrangian Relaxation Heuristic (LR Heuristic): Σε αυτήν τη μέθοδο, αρχικά πραγματοποιείται αναζήτηση ενός ή περισσότερων περίπλοκων περιορισμών [26]. Στη συνέχεια υλοποιείται μια τεχνική χαλάρωσης των υπόλοιπων περιορισμών μέσω της διπλής συνάρτησης Lagrange. Η συγκεκριμένη τεχνική μειώνει τη δυσκολία του προβλήματος. [8].

Constructive and Local Search: Οι αλγόριθμοι δημιουργικής αναζήτησης παράγουν την τοπική βέλτιστη λύση ενός προβλήματος επεκτείνοντας τη λύση της τρέχουσας κατάστασης έως την εύρεση της τελικής λύσης. Συχνά χρησιμοποιείται για την αρχικοποίηση μεθευρετικών αλγορίθμων για την εύρεση πιθανών τοπικών βέλτιστων λύσεων. Οι αλγόριθμοι τοπικής αναζήτησης παίρνουν μια ολοκληρωμένη λύση και τη βελτιστοποιούν σε τοπικό επίπεδο.

Tabu Search (TS): Ανήκει στην κατηγορία των μεθευρετικών αλγορίθμων. Χρησιμοποιείται στην αναζήτηση hill-climbing για την υπέρβαση του τοπικού βέλτιστου.

Large Neighborhood Search (LNS): Ανήκει στην κατηγορία των μεθευρετικών αλγορίθμων. Χρησιμοποιεί μια τεχνική καταστροφής-διόρθωσης της λύσης. Επέκταση της τεχνικής αποτελεί η μέθοδος Adaptive broad Neighborhood Search (ALNS) [60].

Particle Swarm Optimization (PSO): Αποτελεί μία στοχαστική μέθοδο βελτιστοποίησης, η οποία δε χρειάζεται χειρισμό για την εξαγωγή ενός συνόλου υποψηφίων λύσεων.

Ant Colony Optimization and Variants (ACO): Η μέθοδος ανήκει στην κατηγορία των μεθευρετικών αλγορίθμων. Η μοντελοποίηση της βασίζεται στον τρόπο με τον οποίο τα μυρμηγκία βρίσκουν το συντομότερο μονοπάτι προς την τροφή.

Simulated Annealing (SA): Η μέθοδος βασίζεται στην διαδικασία ανόπτησης των μετάλλων [61]. Είναι μια μέθοδος που βασίζεται στις πιθανότητες και έχει ως στόχο

την εύρεση του ελαχίστου της συνάρτησης κόστους.

Genetic Algorithm (GA): Οι γενετικοί αλγόριθμοι βασίζονται στη μίμηση διαδικασιών που σχετίζονται με τη φυσική εξέλιξη [51]. Για τη μοντελοποίησή τους χρησιμοποιούμε ένα διάνυσμα από bit, το οποίο δηλώνει το κάθε χρωμόσωμα. Για την αναζήτηση της λύσης ορίζεται ένα αποτέλεσμα σε κάθε χρωμόσωμα ανάλογο με τον βαθμό επίλυσής του.

Reinforcement Learning (RL): Ανήκει στον κλάδο της μηχανικής μάθησης. Ο μηχανισμός με τον οποίο ανακαλύπτει τη λύση την καθιστά ιδανική για το πρόβλημα χωροθέτησης μονάδων. Στο επόμενο κεφάλαιο αναλύεται εκτενέστερα η αρχιτεκτονική του μοντέλου RL.

Κεφάλαιο 3

Βιβλιογραφική Ανασκόπηση

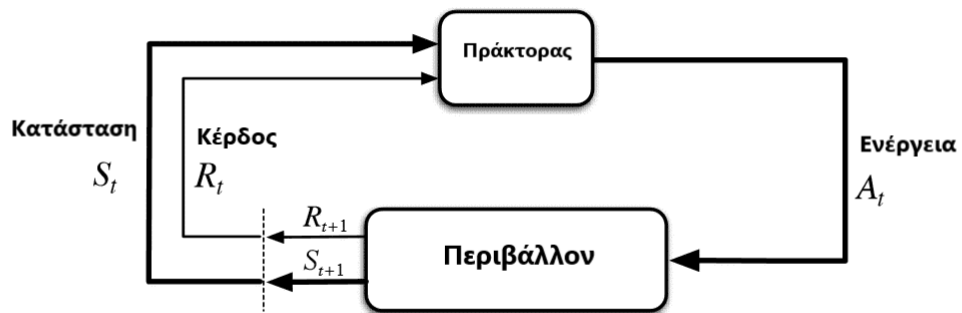
Σε αυτό το κεφάλαιο παρουσιάζονται τεχνικές ενισχυτικής μάθησης που έχουν χρησιμοποιηθεί για την επίλυση του προβλήματος χωροθέτησης μονάδων. Για την καλύτερη κατανόηση των μοντέλων, στην Ενότητα 3.1 περιγράφονται τα βασικά στοιχεία της ενισχυτικής μάθησης.

3.1 Μοντέλο Ενισχυτικής Μάθησης

Η ενισχυτική μάθηση είναι ένας εκ των τριών κλάδων της μηχανικής μάθησης. Βασικό στοιχείο της αποτελεί η μαρκοβιανή διαδικασία κέρδους. Σε αυτήν, ένας πράκτορας λαμβάνει μια απόφαση που τον οδηγεί σε μια μελλοντική κατάσταση. Τα στοιχεία από τα οποία αποτελείται ένα μοντέλο ενισχυτικής μάθησης δίνονται στη συνέχεια.

- Η αίσθηση του πράκτορα είναι ο μηχανισμός που αντιλαμβάνεται τις αλλαγές του περιβάλλοντος που προκύπτουν από τις ενέργειες του πράκτορα.
- Οι ενέργειες είναι οι κινήσεις που πραγματοποιεί ο πράκτορας για να οδηγηθεί στην επόμενη κατάσταση.
- Ο στόχος είναι η επιθυμητή κατάσταση που θέλει να φτάσει ο πράκτορας. Συχνά σχετίζεται με τη μεγιστοποίηση του κέρδους.

Η ενισχυτική μάθηση διαφέρει από τις υπόλοιπες τεχνικές μηχανικής μάθησης λόγω της δυνατότητας ταυτόχρονης εξερεύνησης και εκμετάλλευσης του χώρου για την εύρεση της λύσης. Τα προβλήματα της έχουν δομή μοντέλου αναζήτησης. Σε αυτά, ο πράκτορας αναζητά τον στόχο σε ένα αβέβαιο περιβάλλον. Σε ένα μοντέλο



Σχήμα 3.1: Απεικόνιση μοντέλου ενισχυτικής μάθησης

ενισχυτικής μάθησης ο πράκτορας έχει έναν στόχο, αισθάνεται τα χαρακτηριστικά του περιβάλλοντος και λαμβάνει αποφάσεις που επηρεάζουν το περιβάλλον.

Πέρα από τον πράκτορα και το περιβάλλον, ένα μοντέλο ενισχυτικής μάθησης περιέχει κάποια επιπλέον στοιχεία τα οποία παρατίθενται στη συνέχεια. Η πολιτική ορίζει τη συμπεριφορά του πράκτορα την προκειμένη στιγμή. Πιο συγκεκριμένα παρουσιάζει μια χαρτογράφηση των πιθανών ενεργειών για μία δεδομένη κατάσταση. Σε κάποιες περιπτώσεις, η πολιτική ορίζεται ως μία συνάρτηση ή ένας πίνακας αναζήτησης. Γενικά μπορούμε να πούμε ότι η πολιτική ορίζει την πιθανότητα κάθε ενέργειας. Το κέρδος συνδέεται άμεσα με τον στόχο του προβλήματος. Σε κάθε επανάληψη, το περιβάλλον στέλνει στον πράκτορα μία τιμή η οποία ορίζει το κέρδος. Ο πράκτορας έχει ως στόχο τη μεγιστοποίηση του κέρδους σε βάθος χρόνου. Γενικά, το κέρδος μπορεί να υπολογιστεί μέσω μιας συνάρτησης η οποία εξαρτάται από την παρούσα κατάσταση και τις ενέργειες που έχουν ληφθεί. Η συνάρτηση αξίας ορίζει το κέρδος σε βάθος χρόνου. Μέσω αυτής, ο πράκτορας επιλέγει το μονοπάτι που του προσφέρει μακροπρόθεσμα το μεγαλύτερο κέρδος ανεξάρτητα με το κέρδος κάθε κατάστασης. Τέλος, ένα σημαντικό στοιχείο της ενισχυτικής μάθησης αποτελεί η μοντελοποίησή του περιβάλλοντος. Σκοπός της είναι η περιγραφή της συμπεριφοράς του περιβάλλοντος. Στη συνέχεια δίνεται η απεικόνιση του μοντέλου ενισχυτικής μάθησης (Σχήμα 3.1) για την καλύτερη κατανόηση του προβλήματος.

3.2 Εφαρμογή Μοντέλου Βαθιάς Ενισχυτικής Μάθησης στο Πρόβλημα K-Server

Το Πρόβλημα των K-Server ανήκει στα προβλήματα χωροθέτησης μονάδων και στόχος του είναι η τοποθέτηση μονάδων στο χώρο έχοντας ως κριτήριο την ελαχιστοποίηση του κόστους. Η χρήση αλγορίθμου ενισχυτικής μάθησης είναι εφικτή λόγω της δυνατότητας μοντελοποίησης του προβλήματος σε Markov decision process. Το άρθρο προτείνει ένα συνδυασμό του αλγορίθμου Q-learning και ενός πολυεπίπεδου νευρώνα Perceptron. Η υλοποίηση του επιτυγχάνεται με την ενημέρωση των βαρών του MLP μέσω της συνάρτησης Q. Για τη μοντελοποίηση του προβλήματος ο χώρος παρουσιάζεται με τη μορφή ενός γράφου. Επιπλέον, χρησιμοποιούνται δύο ομογενείς πίνακες για την κατανομή των servers της τρέχουσας και την επόμενης κατάστασης, καθώς και ένας πίνακας για την ακολουθία των αιτήσεων.

Στην κλασική προσέγγιση των προβλημάτων ενισχυτικής μάθησης, ο πράκτορας παρατηρεί το περιβάλλον για να πάρει μια απόφαση και ένα κέρδος. Σε αυτήν την προσέγγιση, ο πράκτορας αντιλαμβάνεται το περιβάλλον από μια εικόνα μέσω κατάλληλης κωδικοποίησης για την εξαγωγή των τιμών που θα αρχικοποιήσουν τους πίνακες. Για κάθε κατανομή servers-requests λαμβάνεται μια απόφαση που οδηγεί στον πίνακα κατανομής της επόμενης κατάστασης. Για το πρόβλημα χρησιμοποιείται ένα πολυεπίπεδο νευρωνικό δίκτυο με μια σιγμοειδή συνάρτηση ενεργοποίησης και αριθμό εξόδων ίσο με τον αριθμό των servers που ζητείται να τοποθετηθούν. Στο κέρδος κάθε απόφασης εισάγουμε τη μικρότερη απόσταση μεταξύ της πηγής και του τελικού προορισμού ώστε να υπολογιστεί το μικρότερο δυνατό κόστος.

Το νευρωνικό δίκτυο προσεγγίζει τη συνάρτηση Q [43, 59, 67]. Η απόφαση υπολογίζεται από τον αλγόριθμο ϵ -greedy ο οποίος οδηγεί στο κέρδος της επόμενης κατάστασης. Στον αλγόριθμο χρησιμοποιείται η αντικειμενική συνάρτηση του Huber [33] η οποία βασίζεται στην απόκλιση της συνάρτησης Q του στόχου με αυτή της παρούσας κατάστασης. Επίσης, χρησιμοποιείται η εμπειρία του πράκτορα η οποία αποθηκεύεται στη μνήμη και δημιουργείται κατά τη διάρκεια της εκπαίδευσης.

Τα πειράματα υλοποιήθηκαν σε ένα παρόμοιο πρόβλημα στο οποίο έχει ομαδοποιηθεί ο πληθυσμός σε σημεία του χάρτη με σκοπό την τοποθέτηση νοσοκομείων. Στόχος του προβλήματος είναι η ελαχιστοποίηση του χρόνου μετακίνησης των ασθε-

νοφόρων από το νοσοκομείο στους ασθενείς. Τα πειράματα πραγματοποιήθηκαν δημιουργώντας 10 θέσεις και 2 ασθενοφόρα τα οποία αρχικοποιήθηκαν με τυχαίο τρόπο. Στη συνέχεια, αυξήθηκαν οι θέσεις σε 15 και 20 ενώ τα ασθενοφόρα σε 5. Οι παράμετροι που χρησιμοποιήθηκαν έχουν αρχικοποιηθεί μετά από πειράματα.

Ο αλγόριθμος που έχει προταθεί αποτελείται από τη φάση της εκπαίδευσης και της εκτέλεσης και συγκρίνεται η απόδοση του με τον Q-learning και έναν άπληστο αλγόριθμο. Για τη σύγκριση των αλγορίθμων πραγματοποιήθηκαν 100 πειράματα για 6 διαφορετικές αρχικοποιήσεις των θέσεων και των ασθενοφόρων. Κατά την υλοποίηση του αλγορίθμου καταγράφηκε ο συνολικός χρόνος του πράκτορα στο ασθενοφόρο και απεικονίζεται σε ένα πίνακα ο μικρότερος, ο μεγαλύτερος και ο μέσος χρόνος. Τέλος, καταγράφηκε ο αριθμός των συμβάντων που τελέστηκαν σε μικρότερο από το μέσο χρόνο για διαφορετική διάρκεια.

Από τον Πίνακα 3.1 παρατηρήθηκε ότι ο αλγόριθμος Q-learning έχει συνολικά μεγαλύτερο ποσοστό επιτυχίας. Λόγω του μεγάλου αριθμού των καταστάσεων-αποφάσεων παρατηρήθηκε ότι για πολλούς κόμβους ο αλγόριθμος που προτείνεται έχει μικρότερο αριθμό ζευγών καταστάσεων-αποφάσεων. Συμπεραίνουμε επομένως ότι ο αλγόριθμος που προτείνεται έχει καλύτερη απόδοση σε μεγάλης κλίμακας προβλήματα.

Πίνακας 3.1: Παρουσίαση απόδοσης για διαφορετικό αριθμό διαστάσεων [44]

Algorithms	n, k	T = 100				T = 1,000				T = 10,000			
		Min	Max	Mean	Wins	Min	Max	Mean	Wins	Min	Max	Mean	Wins
Q-learning		291	432	355.59	20	3,273	3,745	3,509.69	42	34,227	35,778	35,047.11	94
Q-learning with MLP	10, 2	284	432	356.14	12	3,257	3,738	3,513.74	40	34,258	35,821	35,091.05	6
Greedy		294	417	362.52	11	3,358	3,777	3,568.77	10	35,039	36,360	35,704.00	0
Q-learning		334	510	411.30	27	3,836	4,324	4,058.46	13	40,008	41,319	40,681.76	0
Q-learning with MLP	15,2	319	503	408.64	46	3,761	4,282	4,019.94	86	39,716	40,913	40,303.76	100
Greedy		324	522	418.96	15	3,913	4,356	4,144.74	1	40,892	42,020	41,485.92	0
Q-learning		348	497	419.29	20	3,973	4,452	4,230.58	3	41,731	42,952	42,343.22	0
Q-learning with MLP	20,2	343	501	411.78	67	3,887	4,384	4,136.27	97	40,649	42,126	41,368.35	100
Greedy		373	496	428.10	8	4,099	4,575	4,318.00	0	42,510	43,993	43,210.14	0
Q-learning		277	393	344.75	33	3,263	3,761	3,509.33	19	34,417	35,767	35,081.79	9
Q-learning with MLP	20,3	285	415	344.31	46	3,248	3,708	3,469.00	78	34,156	35,663	34,790.00	91
Greedy		301	433	353.51	17	3,373	3,773	3,597.96	0	35,325	36,646	36,003.36	0
Q-learning		258	384	317.74	13	2,963	3,386	3,206.60	0	31,272	32,507	31,943.31	0
Q-learning with MLP	20,4	226	381	294.94	50	2,814	3,143	2,980.62	89	29,250	30,566	29,819.78	100
Greedy		250	369	302.90	34	2,797	3,214	3,055.83	11	30,114	31,131	30,568.75	0
Q-learning		253	381	314.51	0	2,961	3,436	3,205.79	0	31,500	32,843	32,058.52	0
Q-learning with MLP	20,5	199	320	252.25	58	2,351	2,724	2,526.08	93	24,682	26,041	25,318.13	100
Greedy		199	321	258.26	42	2,444	2,738	2,600.74	7	25,267	26,652	26,072.42	0

Keywords: Min - Minimum Time Spent
 Max - Maximum Time Spent
 Mean - Average Time Spent
 Wins - Victories

3.3 Αύξηση Γράφου για Ίση Πρόσβαση με τη Χρήση Ενισχυτικής Μάθησης

Στο άρθρο [57] παρουσιάζεται η κατηγορία προβλημάτων Graph Augmentation for Equitable Access (GAEA) στόχος της οποίας είναι η επεξεργασία των ακμών των γράφων υπό ένα προκαθορισμένο προϋπολογισμό. Μια μέθοδος επίλυσης των GAEA είναι η άπληστη βελτιστοποίηση για τη δίκαιη επιλογή βάσης της κεντρικότητας (Greedy Equitable Centrality Improvement, GECI [18]). Στόχος της είναι η εύρεση των ακμών με τη μεγαλύτερη μεταβολή που οδηγεί στους κόμβους-κέρδους.

Η προτεινόμενη μέθοδος ονομάζεται Equitable Mechanism Design in MRP (EMD-MRP), μοντελοποιείται βάσει της Μαρκοβιανής διαδικασίας κέρδους (Markov Reward Process, MRP) και έχει ως στόχο τη μεγιστοποίηση της συνάρτησης κέρδους. Σε αυτή ορίζεται μια μεταβλητή γ για την παρουσίαση της εξάρτησης του συστήματος από τις επόμενες καταστάσεις. Λόγω της κατάταξης της στην κατηγορία των NP-hard προβλημάτων, δεν είναι δυνατή η προσέγγιση της από τον παράγοντα $1 - \frac{1}{3*\epsilon}$.

Το πρόβλημα χωροθέτησης μονάδων είναι πρόβλημα υποδιανομής. Για το κόστος των μονάδων υπάρχει η άπληστη λύση $1 - \frac{1}{3*\epsilon}$. Η χρονική πολυπλοκότητα κάθε επεισοδίου του MRP επηρεάζεται σε μεγάλο βαθμό από την εκτίμηση της συνάρτησης κέρδους του κάθε συνόλου.

Στο άρθρο αναφέρονται τέσσερα μοντέλα συνθετικών γράφων τα οποία παρατίθενται παρακάτω:

- Στο μοντέλο Erdos-Renyi [24] δημιουργείται ένας τυχαίος γράφος ο οποίος πραγματοποιείται από την πιθανότητα τοποθέτησης μιας ακμής ανάμεσα σε δύο κόμβους. Επίσης χρησιμοποιείται το γινόμενο της πιθανότητας και του αριθμού των κόμβων που υπάρχουν στον γράφο για τον υπολογισμό της πυκνότητας του γράφου. Η αύξηση της πυκνότητας μειώνει τη μεταβολή της τιμής του κέρδους των κόμβων.
- Το μοντέλο Preferential Attachment Cluster Graph [35] είναι μια επέκταση του Barabasi-Albert. Η παραμετροποίηση του γίνεται με την προσθήκη m ακμών σε κάθε κόμβο και την πιθανότητα δημιουργίας τριγώνου μεταξύ τριών κόμβων.
- Το μοντέλο Chung-Lu [11] βασίζεται σε έναν αναμενόμενο βαθμό ο οποίος

σχετίζεται με τη διανομή μιας αλληλουχίας εισόδων. Πραγματοποιείται από την πιθανότητα δειγματοληψίας ενός κόμβου και του συντελεστή μείωσης γ . Θέτοντας το γ σε 0 δημιουργείται ένας γράφος με τυχαίο βαθμό. Αυξάνοντας το γ μειώνονται οι κόμβοι με υψηλό και αυξάνονται οι κόμβοι με χαμηλό βαθμό. Αυτό έχει ως αποτέλεσμα τη δρομολόγηση των κόμβων με μεγάλο βαθμό. Το μοντέλο χρησιμοποιείται για τη μέτρηση της απόδοσης δίνοντας βαρύτητα στην κεντρικότητα κάθε κόμβου. Με αυτό τον τρόπο επιλέγεται η τοποθέτηση των κερδών.

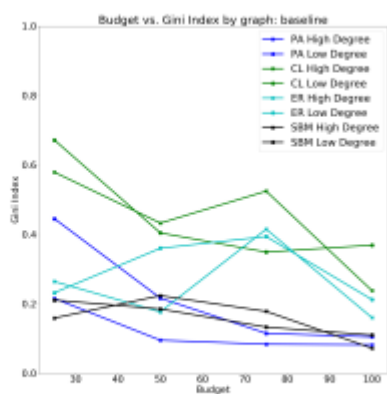
- Στο Stochastic Block Model [32] επιλέγονται ακμές μέσα από M clusters χρησιμοποιώντας ένα πίνακα που περιέχει την πιθανότητα της κάθε ακμής. Τυπικά, τα στοιχεία της διαγωνίου του πίνακα έχουν μεγαλύτερη πιθανότητα από τα υπόλοιπα στοιχεία της κάθε γραμμής. Σκοπός του μοντέλου είναι η μέτρηση της απόδοσης για τη δρομολόγηση ανάμεσα σε clusters. Στο παράδειγμα ορίζονται δύο clusters ίσου μεγέθους στα οποία έχει αρχικοποιηθεί η πιθανότητα εντός και μεταξύ τους. Βασικός στόχος του παραδείγματος είναι η κατεύθυνση σωματιδίων σε μια αραιά συνδεδεμένη περιοχή του γράφου.

Για κάθε ένα από τα παραπάνω μοντέλα χρησιμοποιήθηκαν δυο τεχνικές εκ των οποίων η μια βασίστηκε σε κόμβους με υψηλό βαθμό ενώ η άλλη με χαμηλό. Η αξιολόγηση τους έγινε χρησιμοποιώντας την τεχνική Monte Carlo για το αναμενόμενο κέρδος κάθε ομάδας. Οι μέθοδοι με τις οποίες έγινε η αξιολόγηση των γράφων είναι το μέσο κέρδος και ο δείκτης Gini.

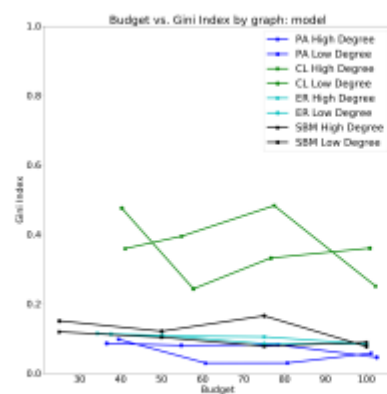
Στις γραφικές παραστάσεις του Σχήματος 3.2 παρατηρείται ότι η προτεινόμενη μέθοδος υπερτερεί της βασικής γραμμής στα περισσότερα μοντέλα. Ακόμα, στο σενάριο με το χαμηλότερο κόστος το μοντέλο υπερτερεί στο Gini Index. Ως προς τη χρησιμότητα, για τον ίδιο προϋπολογισμό το μοντέλο υπερτερεί της βασικής γραμμής κατά 0.5.

Τα αποτελέσματα οκτώ παραδειγμάτων τεσσάρων συνθετικών γράφων απεικονίστηκαν σε μια γραφική παράσταση ως προς τη χρησιμότητα και το δείκτη Gini. Παρατηρήθηκε ότι ο PA High Degree είχε την καλύτερη απόδοση ενώ το μοντέλο του Chung-Lu τη χειρότερη.

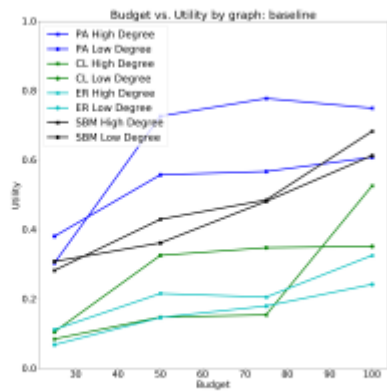
Το μοντέλο EMD-MRP υλοποιήθηκε για το πρόβλημα χωροθέτησης μονάδων σε έναν γράφο PA μεγέθους $N = 200$. Παρατηρήθηκε ότι ο δείκτης Gini μειώνεται όσο



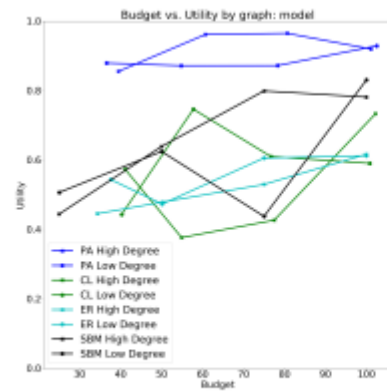
(a) GECI baseline



(b) Proposed method



(c) GECI baseline



(d) Proposed model

Σχήμα 3.2: Γραφική παράσταση του προϋπολογισμού ως προς το δείκτη Gini [57]

το budget αυξάνεται ενώ η μέση χρηστικότητα αυξάνεται όσο το budget μεγαλώνει.

Στο άρθρο παρουσιάζεται επίσης μια εφαρμογή που βασίζεται στη δίκαιη πρόσβαση σε σχολεία στην πόλη του Σικάγο. Για τη δημιουργία του γράφου χρησιμοποιείται το δίκτυο μεταφοράς των δημόσιων λεωφορείων. Οι κόμβοι επιλέγονται βάσει της διασταύρωσης των διαδρομών ενώ οι ακμές από τις διαδρομές. Ο γράφος που δημιουργήθηκε αποτελείται από 2011 κόμβους και 7984 ακμές. Η τοποθεσία των σχολείων και ο τρόπος αξιολόγησης τους επιλέχθηκαν από μια βάση δεδομένων. Για το πρόβλημα δημιουργήθηκαν τρία σωματίδια τα οποία αντιπροσωπεύουν λευκούς, μαύρους και ισπανικής καταγωγής βάσει του δημογραφικού του 2010. Στη συνέχεια επιλέχθηκαν τυχαία οι κόμβοι που αποτελούν τις αρχικές θέσεις των σωματιδίων ενώ τα αρχικά βάρη αρχικοποιήθηκαν αντιστρόφως ανάλογα της απόστασης. Για budget 400 ακμών παρατηρήθηκε ότι για αρχικές τιμές 0.20 και 0.62 το προτεινόμενο μοντέλο έχει μέση χρηστικότητα 0.90 και δείκτη Gini 0.07 ενώ το μοντέλο GECI 0.21 και 0.65 αντίστοιχα.

Ένα δεύτερο παράδειγμα υλοποιήθηκε στα μέσα κοινωνικής δικτύωσης φοιτητών. Για την υλοποίηση του προβλήματος χρησιμοποιήθηκε το σύνολο δεδομένων Facebook100 [68] στο οποίο ερευνήθηκαν οι κοινωνικές επαφές ατόμων από 100 πανεπιστήμια. Ως κόμβοι κέρδους ορίστηκαν οι δημοφιλείς τελειόφοιτοι ενώ σκοπός της ανάλυσης ήταν η ίση πρόσβαση νέων φοιτητών στους κόμβους με επιρροή. Από τον Πίνακα 3.2, παρατηρήθηκε ότι το μοντέλο ενισχυτικής μάθησης υπερτερεί του αρχικού και του μοντέλου GECI. Συγκρίνοντας τη διαφορά ανάμεσα στο συντομότερο μονοπάτι των ομάδων, παρατηρήθηκε ότι το GECI διατηρεί μικρότερο περιθώριο. Αυτό οφείλεται στη χαλαρότητα του μοντέλου EMD-MRP.

Πίνακας 3.2: Δείκτης Gini [57]

	Original	EMD-MRP	GECI
Reed	0.214	0.093	0.153
Caltech	0.092	0.065	0.812
Michigan State	0.115	0.086	0.157

3.4 Επίλυση του Προβλήματος Συνδυαστικής Βελτιστοποίησης με τη Χρήση Ενισχυτικής Μάθησης Πολλαπλών Πρακτόρων

Στο προτεινόμενο άρθρο [48] παρουσιάζονται αλγόριθμοι οι οποίοι βασίζονται στη Μαρκοβιανή διαδικασία αποφάσεων (Markov Decision Process, MDP) και ορίζονται από μια ομάδα καταστάσεων, τις διαθέσιμες κινήσεις τους, την πιθανότητα μετακίνησης και το κέρδος. Στόχος τους είναι η εύρεση κατάλληλης πολιτικής για τη μεγιστοποίηση του μέσου κέρδους η οποία ορίζεται από τη συνάρτηση state-value και βασίζεται στην πολιτική της τρέχουσας κατάστασης. Ως ελεύθερα μοντέλα ορίζονται οι αλγόριθμοι στους οποίους ο πράκτορας δε γνωρίζει την πιθανότητα μετακίνησης. Σε αυτά η συνάρτηση state-value ορίζει το αναμενόμενο μέσο κέρδος το οποίο λαμβάνεται από την πολιτική της τρέχουσας κατάστασης. Η σύγκλιση βασίζεται στη βέλτιστη συνάρτηση κέρδους.

Για τη δημιουργία ενός προβλήματος βελτιστοποίησης χρησιμοποιήθηκε το πρόβλημα του πλανόδιου πωλητή (Traveling Salesman Problem, TSP). Ως γενικό κέρδος ορίστηκε το αντίστροφο μήκος ενώ ως άμεσο κέρδος η αντίστροφη απόσταση μεταξύ δύο πόλεων. Στη συνέχεια αναλύθηκε η τεχνική των γενετικών αλγορίθμων, σκοπός των οποίων είναι ο έλεγχος την καταλληλότητα κάθε λύσης. Σε αυτούς κάθε σχήμα έχει μια έμμεση πιθανότητα εμφάνισης της λύσης. Λόγω της μεγάλης αύξησης των σχημάτων, οι τιμές γίνονται απαγορευτικά μεγάλες. Οι αλγόριθμοι ενισχυτικής μάθησης συγκλίνουν στο μέγιστο, δεδομένων των συνθηκών και εφαρμόζονται σε προβλήματα με μεγάλο αριθμό καταστάσεων. Μπορούν να συνδυαστούν με προσεγγιστικές συναρτήσεις για την επίτευξη γενικεύσεων και τη μείωση των απαιτήσεων ως προς τον χώρο. Μπορούν επίσης να βελτιωθούν ως προς την ακρίβεια και τη μείωση του υπολογιστικού κόστους. Μειονέκτημα της παραπάνω τεχνικής είναι η εξάρτηση της από το ιστορικό των καταστάσεων η οποία εμφανίζεται σε προβλήματα όπως του TSP και επηρεάζει τις μετέπειτα καταστάσεις. Λόγω αυτού του γεγονότος, το σύστημα χάνει τις ιδιότητες που έχει αποκτήσει ως MDP.

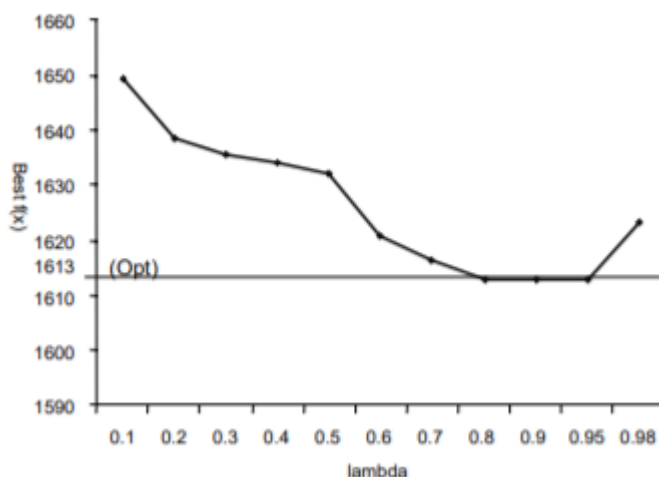
Στη συνέχεια υλοποιήθηκαν δυο εφαρμογές του αλγορίθμου, στο ασύμμετρο πρόβλημα του TSP και της τετραγωνικής ανάθεσης στο οποίο ανήκει το πρόβλημα χωροθέτησης μονάδων. Το ασύμμετρο πρόβλημα του πλανόδιου πωλητή χαρακτηρίζεται ως NP-hard για την εύρεση του συντομότερου κύκλου Hamilton ενός πλήρους

γράφου. Η λύση βασίζεται στη μέθοδο Q-learning ενός βήματος. Το κέρδος αποτελείται από τον συνδυασμό του άμεσου και του γενικού κέρδους. Η ανανέωση πραγματοποιείται στο τέλος κάθε επεισοδίου ενώ η πολυπλοκότητα του είναι $O(n^2)$. Το γενικό κέρδος υπολογίζεται βάση της μέσης καταλληλότητας των γονέων. Ωστόσο υπάρχουν και άλλοι τρόποι υπολογισμού του γενικού κέρδους οι οποίοι επηρεάζουν την υπολογιστική προσπάθεια. Για τη στρατηγική επιλογής ο συγγραφέας χρησιμοποίησε την πολιτική ε-greedy. Στη μέθοδο δημιουργήθηκε η μεταβλητή λ η οποία ελέγχει το κλάσμα των αντιγραμμένων τιμών ως προς το σύνολο των αναθέσεων. Για $\lambda=0$, ο Q-learning χρησιμοποιείται για όλες τις αναθέσεις ενώ για την τιμή 1 τα αποτελέσματα είναι παρόμοια με αυτά του προβλήματος τετραγωνικής ανάθεσης.

Το πρόβλημα τετραγωνικής ανάθεσης (Quadratic Assignment Problem, QAP) έχει ως στόχο την εύρεση των ελαχίστων μεταθέσεων. Οι παράμετροι του είναι ο αριθμός των μονάδων, το κόστος τοποθέτησης, το κόστος μεταφοράς και η ροή από την αφετηρία στον προορισμό. Η μοντελοποίηση του το καθιστά μη γραμμικό. Η επιλογή τοποθέτησης των μονάδων βασίζεται στην action-values ενώ εξαρτάται από τη συχνότητα τοποθέτησης των μονάδων. Για τη μείωση των δυσκολιών που εμφανίζονται στο συγκεκριμένο πρόβλημα χρησιμοποιείται η μέθοδος Monte-Carlo η οποία εισάγει μια μικρή καθυστέρηση στη σύγκληση. Η σειρά επιλογής των μονάδων γίνεται με τυχαίο τρόπο. Ο ε-greedy χρησιμοποιείται για την επιλογή των μονάδων από το σύνολο των μη ανατεθειμένων ενώ βελτιστοποιείται με τον απλό 1-Opt optimizer.

Ένα ενδιαφέρον συμπέρασμα που εξάγεται είναι η επιρροή που ασκεί η παράμετρος λ ως προς την ποιότητα της αναζήτησης, αλλάζοντας το ποσό της αναπαραγωγής. Για τη υλοποίηση των παραδειγμάτων στο ασύμμετρο πρόβλημα του πλανόδιου πωλητή (Asymmetric Traveling Salesman Problem, ATSP) χρησιμοποιήθηκε το `ftv 44`. Στη συνέχεια δίνεται η γραφική παράσταση (Σχήμα 3.3) της καλύτερης μέσης τιμής ως προς το λ για δέκα επαναλήψεις.

Παρατηρείται ότι όσο η τιμή του λ αυξάνεται τόσο ο αλγόριθμος βρίσκει καλύτερη λύση. Η μέγιστη απόδοση δίνεται για λ από 0.75 έως 0.95. Τα πειράματα για το QAP και ATSP έδωσαν παρόμοια αποτελέσματα. Για την παραμετροποίηση χρησιμοποιήθηκαν 50 πράκτορες, η επιλογή αυτών έγινε βάση της μεθόδου της ρουλέτας ενώ οι παράμετροι λ , α , β , γ , ϵ αρχικοποιήθηκαν για κάθε εφαρμογή με τυχαίο τρόπο από προκαθορισμένο διάστημα.



Σχήμα 3.3: Απεικόνιση καλύτερης μέσης τιμής για το ftv44 ως προς το λ [48]

Ο μέσος όρος των αποτελεσμάτων δέκα υλοποιήσεων των προβλημάτων καταγράφηκε σε δύο πίνακες. Τα παραδείγματα προήλθαν από τις βιβλιοθήκες QAPLIB [10] και TSPLIB [58]. Η προτεινόμενη υλοποίηση έχει αρκετά κοινά με τα Ant Systems [20, 21] και τους γενετικούς αλγορίθμους [47], για αυτόν τον λόγο συγκρίνεται με τις δύο μεθόδους. Παρατηρείται ότι ο αλγόριθμος ενισχυτικής μάθησης έχει καλύτερη απόδοση σε όλα τα προβλήματα σε σχέση με τον AS, ενώ ως προς τον GA+LS [46] ήταν καλύτερος σε 8 από τα 15 προβλήματα. Επίσης ο αλγόριθμος βρήκε τη βέλτιστη λύση και στις 10 εκτελέσεις σε περιπτώσεις μικρού και μεσαίου μεγέθους. Τέλος, στην περίπτωση του GA+LS, η τυπική απόκλιση από την καλύτερη λύση ήταν μη μηδενική ακόμα και σε απλά προβλήματα QAP.

3.5 Μοντέλο Βελτιστοποίησης της Εγγύησης Προσβασιμότητας για την Τοποθέτηση Μονάδων Έκτακτης Ανάγκης με Ενίσχυση Συνδέσμων

Σκοπός του άρθρου [74] είναι η τοποθέτηση κύριων και βοηθητικών μονάδων επειγόντων περιστατικών για την αντιμετώπιση ακραίων καταστροφών χρησιμοποιώντας την τεχνική της ενίσχυσης των ακμών. Στο μοντέλο, η επιλογή των τοποθεσιών έγινε βάση των εξόδων κατασκευής, λειτουργίας και της δυνατότητας πρόσβασης στα σημεία ζήτησης. Για το πρόβλημα δημιουργήθηκε ένας μη κατευθυνόμενος γράφος. Σε αυτόν ορίστηκε η εγγύηση προσβασιμότητας (reachability guarantee) ως η ικανότητα σύνδεσης δύο σημείων με σκοπό την αντίσταση σε κατα-

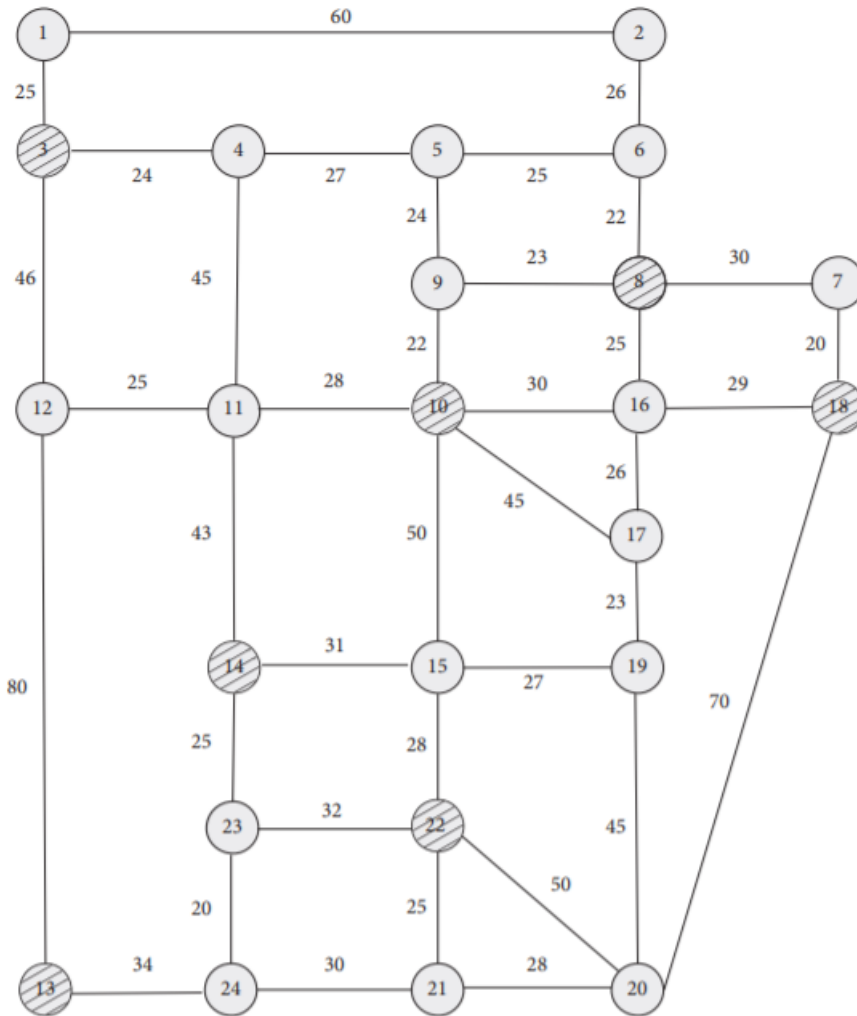
στροφές. Επιπλέον για κάθε ακμή ορίστηκε η ανοχή σε καταστροφές. Η τοποθέτηση των μονάδων έγινε μετά από τον υπολογισμό της εγγύησης προσβασιμότητας για κάθε ένα από τα σημεία ζήτησης. Ο γράφος τροποποιήθηκε βάση του προβλήματος ροής ενώ το λειτουργικό κόστος υπολογίστηκε βάση της ελάχιστης απόστασης ανάμεσα στη μονάδα και τα σημεία ζήτησης. Η ενίσχυση των συνδέσμων (reinforce link) έγινε βάση προκαθορισμένων τιμών με σκοπό την αύξηση της εγγύησης προσβασιμότητας.

Για τη μοντελοποίηση του προβλήματος χρησιμοποιήθηκε το σύνολο των υποψηφίων θέσεων των μονάδων και το σύνολο των σημείων ζήτησης. Επίσης ορίστηκε η ανοχή κάθε ακμής, η πιθανή αύξηση της προσβασιμότητας κατά τη διαδικασία της ενίσχυσης, το λειτουργικό κόστος και ο προϋπολογισμός. Σκοπός του μοντέλου είναι η εύρεση της τοποθεσίας της κύριας και βοηθητικής μονάδας και ο καθορισμός του συνόλου των ακμών που θα ενισχυθούν. Στόχος του είναι η μεγιστοποίηση της εγγύησης της προσβασιμότητας και η μείωση της ελάχιστης απόστασης.

Στο άρθρο παρουσιάζεται ένα μοντέλο δυο επιπέδων [31]. Στο κάτω επίπεδο υπολογίζεται η απόσταση και η προσβασιμότητα ενώ στο πάνω εισάγονται τα προηγούμενα δεδομένα για την αξιολόγηση του υποσυνόλου των θέσεων και των ακμών. Λόγω της πολυπλοκότητας της μεθόδου, προτείνεται ένα νέο μοντέλο το οποίο βασίζεται στο πρόβλημα μέγιστης ροής. Η ενσωμάτωση του στο μοντέλο τοποθεσιών το καθιστά εύκολα επιλύσιμο από αρκετά εμπορικά λογισμικά βελτιστοποίησης. Τελικός σκοπός του μοντέλου είναι το κόστος λειτουργίας υπο φυσιολογικές συνθήκες και η εγγύηση προσβασιμότητας ανάμεσα στις κύριες και τις βοηθητικές μονάδες. Λόγω της ύπαρξης δύο στόχων, το τελικό αποτέλεσμα δίνεται από μια σταθμισμένη αντικειμενική συνάρτηση.

Το μοντέλο δοκιμάστηκε σε μια απλοποιημένη μορφή του δικτύου μεταφοράς του Sioux Falls (Σχήμα 3.4) στην οποία ορίστηκαν 24 κόμβοι και 38 ακμές. Στο άρθρο δίνεται ένας πίνακας ο οποίος περιέχει την ανεκτικότητα και το κόστος ενίσχυσης των ακμών. Επίσης χρησιμοποιείται ένας δεύτερος πίνακας ο οποίος περιέχει το κόστος εγκατάστασης των μονάδων σε κάθε μια από τις πιθανές θέσεις. Στη συνέχεια δίνεται ο Πίνακας 3.3 ο οποίος περιέχει τα αποτελέσματα του μοντέλου για διαφορετικές τιμές του προϋπολογισμού G και του συντελεστή ω .

Στον Πίνακα 3.3 το κόστος λειτουργίας συμβολίζεται ως g_1 , η εγγύηση προσβα-



Σχήμα 3.4: Δικτύου μεταφοράς του Sioux Falls [74]

σιμότητας ως g_2 , το συνολικό κόστος δημιουργίας ως C_F και το συνολικό κόστος ενίσχυσης ως C_L . Παρατηρείται ότι για δεδομένο προϋπολογισμό, όσο το ω αυξάνεται τόσο τα g_1 , g_2 και C_L μειώνονται. Αντίθετα, το C_F αυξάνεται. Λόγω της ύπαρξης δύο τελικών στόχων τα αποτελέσματα αξιολογούνται βάσει της ποσοστιαίας μεταβολής δύο διαδοχικών προϋπολογισμών. Όσον αφορά την ενίσχυση των ακμών, παρατηρείται ότι η νέα προσβασιμότητα δε μειώνεται.

3.6 Εφαρμογή Μεθόδου Βελτιστοποίησης Πολλαπλών Πρακτόρων για το Πρόβλημα Τετραγωνικής Ανάθεσης

Στο άρθρο [64] παρουσιάζεται μια μέθοδος βελτιστοποίησης του προβλήματος Quadratic Assignment Problem (QAP). Επιπλέον αναλύεται ένα πλήθος ευρετικών αλγορίθμων για το QAP οι οποίοι παρουσιάζονται στη συνέχεια και αποσκοπούν στη σύγκριση με το προτεινόμενο μοντέλο. Οι ευρευτικοί αλγόριθμοι είναι οι εξής:

- Ο αλγόριθμος Robust Tabu Search (Ro-TS) [66] πραγματοποιεί ανταλλαγή μεταξύ δύο στοιχείων της λύσης. Η αποτροπή αντιστρόφων ανταλλαγών επιτυγχάνεται με τη χρήση λίστας.
- Ο Improved Hybrid Genetic Algorithm (IHGA) [49] ενσωματώνει μια τοπικά βελτιστοποιημένη διαδικασία και μια νέα βέλτιστη διασταύρωση.
- Ο αλγόριθμος Population-based Iterated Local Search (PILS) [65] αποτελεί μια επέκταση του Iterated Local Search (ILS).
- Ο αλγόριθμος Cooperative Parallel Tabu Search (CPTS) [34] πραγματοποιεί παράλληλη εκτέλεση της αναζήτησης tabu (Tabu Search, TS) χρησιμοποιώντας διαφορετικές συνθήκες τερματισμού και αριθμού επαναλήψεων.
- Ο αλγόριθμος Breakout Local Search (BLS) [9] αποτελείται από τη φάση της τοπικής αναζήτησης και της διαταραχής. Σκοπός της τοπικής αναζήτησης είναι η εύρεση νέων τοπικών βέλτιστων τιμών. Η διαταραχή χρησιμοποιείται για την εύρεση νέων υποσχόμενων περιοχών.

Η προτεινόμενη μέθοδος βασίζεται στη χρήση πολλαπλών πρακτόρων για την επίλυση του QAP (Multi-Agent Optimization Method for QAP, MAOM-QAP). Τα

συστήματα πολλαπλών πρακτόρων προσφέρουν κατανεμημένη πληροφορία, συνεργασία μεταξύ των πρακτόρων και δυναμική λήψη αποφάσεων. Το MAOM-QAP αποτελείται από έναν πράκτορα λήψης αποφάσεων (decision maker agent), πράκτορες τοπικής αναζήτησης (local search agents) και πράκτορες διασταύρωσης (crossover agents). Η επιλογή ενεργοποίησης των πρακτόρων γίνεται βάση ενός πίνακα βαρών. Σε αυτόν, για την αύξηση των τιμών πραγματοποιείται μια διαδικασία ενισχυτικής μάθησης.

Για τη μοντελοποίηση του πίνακα χρησιμοποιείται ένα ζεύγος καταστάσεων-ενεργειών. Για κάθε ζεύγος δίνεται ένα βάρος W_{ij} το οποίο ορίζει τη βελτίωση της κατάστασης κάθε εποχής της αναζήτησης. Επίσης, ορίζεται η πιθανότητα εφαρμογής μιας ενέργειας για μια κατάσταση. Σε κάθε εποχή, αρχικά η βελτιωμένη κατάσταση σχετίζεται με την προκαθορισμένη κατάσταση. Στη συνέχεια επιλέγεται η καλύτερη ενέργεια βάση της πιθανότητας εφαρμογής. Στο τέλος κάθε εποχής η τιμή του βάρους αυξάνεται εφόσον η λύση του μοντέλου έχει βελτιωθεί.

Ο πράκτορας που λαμβάνει τις αποφάσεις (decision maker agent) επιλέγει ποια κατηγορία πρακτόρων θα πυροδοτηθεί. Επίσης, είναι ο μόνος που έχει στη διάθεση του το αρχείο με τις καλύτερες λύσεις. Ο πίνακας βαρών αποτελείται από τέσσερις περιπτώσεις σημαντικών καταστάσεων που μπορούν να πραγματοποιηθούν:

- Ο αλγόριθμος να μην εκτείνεται για g_0 γενιές.
- Η λύση που προέρχεται από τους πράκτορες να βελτιώνει ελάχιστα την αντικειμενική συνάρτηση για g_1 γενιές.
- Η λύση που προέρχεται από τους πράκτορες να βελτιώνει αρκετά την αντικειμενική συνάρτηση για g_1 γενιές.
- Η λύση που προέρχεται από τους πράκτορες να μη βελτιώνει την αντικειμενική συνάρτηση για g_2 γενιές.

Τα g_0 , g_1 , g_2 είναι παράμετροι του προβλήματος και σχετίζονται με τον μέγιστο αριθμό επαναλήψεων.

Η επιλογή ενεργοποίησης των πρακτόρων γίνεται βάση των καταστάσεων που δόθηκαν παραπάνω. Στην περίπτωση της πρώτης και της τρίτης συνθήκης πυροδοτείται ο πράκτορας τοπικής αναζήτησης. Αντίστοιχα, στην περίπτωση της δεύτερης και της τέταρτης συνθήκης πυροδοτείται ο πράκτορας διασταύρωσης.

Οι πράκτορες τοπικής αναζήτησης (local search agents) πραγματοποιούν αναζήτηση μεμονωμένα στις περιοχές που έχουν τοποθετηθεί. Κάθε πράκτορας αποφασίζει αν θα ανταλλάξει πληροφορίες με τους υπόλοιπους ενεργούς πράκτορες. Η βέλτιστη λύση αποστέλλεται στον decision maker agent. Στο QAP, οι local search agents πραγματοποιούν την αναζήτηση τους βάση του tabu search χρησιμοποιώντας δύο διαφορετικές τεχνικές τοπικής αναζήτησης. Στην πρώτη, ο πράκτορας εξερευνά τη γειτονιά έχοντας ως στόχο την εύρεση της βέλτιστης τοπικής λύσης. Στη δεύτερη, ο πράκτορας επιλέγει τυχαία μια μονάδα. Στη συνέχεια, αναζητά τη μονάδα που θα του δώσει το μεγαλύτερο κέρδος ανταλλαγής. Με αυτό τον τρόπο επιλέγει την καλύτερη λύση. Η τυχαία επιλογή μιας εκ των δυο τεχνικών προσφέρει στον πράκτορα ένα βαθμό διαφοροποίησης της λύσης. Επίσης, χρησιμοποιεί μια λίστα tabu για την αποτροπή επανεξέτασης προηγούμενων λύσεων.

Ο πίνακας βαρών αποτελείται από τρεις περιπτώσεις σημαντικών καταστάσεων που μπορούν να πραγματοποιηθούν και δίνονται παρακάτω.

- Η βέλτιστη τοπική λύση που προέρχεται από τους πράκτορες βελτιώνεται ελάχιστα για q_3 γενιές.
- Η βέλτιστη τοπική λύση που προέρχεται από τους πράκτορες δε βελτιώνεται για q_4 γενιές.
- Η βέλτιστη τοπική λύση που προέρχεται από τους πράκτορες δε βελτιώνεται για q_5 γενιές.

Τα q_3 , q_4 , q_5 είναι παράμετροι του προβλήματος και σχετίζονται με τον μέγιστο αριθμό επαναλήψεων.

Η επιλογή ενεργοποίησης των πρακτόρων γίνεται βάση των καταστάσεων που δόθηκαν παραπάνω. Στην περίπτωση της πρώτης συνθήκης ενεργοποιείται ένας νέος πράκτορας τοπικής αναζήτησης. Αντίστοιχα, στην περίπτωση της δεύτερης και της τρίτης συνθήκης ενεργοποιείται ο πράκτορας διαταραχής (perturbation agent) ο οποίος παραμορφώνει τη λύση. Στη συνέχεια τη στέλνει στον πράκτορα τοπικής αναζήτησης.

Στην προτεινόμενη τεχνική, οι πράκτορες διασταύρωσης (crossover agents) χρησιμοποιούν δύο διαφορετικές τεχνικές εύρεσης των λύσεων. Στην πρώτη, επιλέγεται

η τοποθεσία της μονάδας του γονέα με τη μικρότερη αντικειμενική τιμή. Στη συνέχεια, από τον δεύτερο γονέα επιλέγεται η τοποθεσία της επόμενης μονάδας. Η παραπάνω διαδικασία τελειώνει με το πέρας των μονάδων και έχει στόχο τη δημιουργία μιας καινούργιας λύσης. Η δεύτερη τεχνική διαφέρει στον αριθμό των συνεχόμενων τοποθεσιών που αντιγράφονται στη νέα λύση. Ενώ η πρώτη τεχνική αντιγράφει με βήμα ένα, η δεύτερη αντιγράφει με ρυθμό ζ ο οποίος ορίζεται ως παράμετρος.

Η υλοποίηση του μοντέλου πραγματοποιήθηκε χρησιμοποιώντας τη γλώσσα Java και την πλατφόρμα Jade. Τα πειράματα βασίστηκαν σε κάποια προβλήματα της βιβλιοθήκης QAPLIB (<https://www.opt.math.tugraz.at/qaplib/>). Η παραμετροποίηση των μεταβλητών καθώς και ο αριθμός επαναλήψεων των πρακτόρων έγινε βάση πειραματικής μελέτης.

Το MAOM-QAP συγκρίθηκε με τους παρακάτω αλγορίθμους:

- Improved hybrid genetic algorithm (IHGA) [49]
- Iterated tabu search (ITS) [50]
- Population-based iterated local search (PILS) [65]
- A hybrid genetic tabu search algorithm (MRT60) [23]
- Cooperative parallel tabu search (CPTS) [34]
- The Breakout local search (BLS) [9]
- The population-based Memetic Algorithm (BMA) [9]

Αξίζει να σημειωθεί ότι η υλοποίηση των παραπάνω αλγορίθμων έγινε βάση των παραμέτρων που αποσκοπούσαν στην καλύτερη λύση του κάθε αλγορίθμου, δοκιμάστηκαν δηλαδή διάφορες παράμετροι και κρατήθηκαν οι καλύτερες. Η σύγκριση των αλγορίθμων έγινε ως προς την απόκλιση από τη μέση καλύτερη λύση.

Παρατηρήθηκε ότι το MAOM-QAP βρήκε την καλύτερη λύση σε 7 από τα 9 τυχαία δημιουργημένα προβλήματα της βιβλιοθήκης όπως και άλλοι αλγόριθμοι. Στην επίλυση πραγματικών προβλημάτων, η καλύτερη λύση βρέθηκε σε 4 από τα 5 προβλήματα, ενώ το πρόβλημα για το οποίο δε βρέθηκε η καλύτερη λύση ήταν το μεγαλύτερο. Για τα προβλήματα που μοντελοποιήθηκαν βάση της απόστασης

Manhattan ενός δικτύου, το MAOM-QAP βρήκε τη καλύτερη λύση σε 14 από τα 15 προβλήματα. Λόγω της απόκλισης από τη βέλτιστη λύση, το MAOM-QAP βρέθηκε στα τρία καλύτερα μοντέλα μαζί με το BMA και το CPTS.

Στο άρθρο πραγματοποιήθηκε η επίλυση των παραπάνω μοντέλων χρησιμοποιώντας το MAOM-QAP με και χωρίς τον perturbation agent. Τα περάματα έδειξαν ότι το μοντέλο που δεν είχε τον perturbation agent, δεν κατάφερε να λύσει κανένα από τα 21 προβλήματα. Στη συνέχεια, πραγματοποιήθηκε η παραπάνω σύγκριση ως προς τον crossover agent. Το μοντέλο που δεν είχε τον crossover agent κατάφερε να επιλύσει τα 4 από τα 21 προβλήματα. Τα δυο περάματα απέδειξαν την ανάγκη ύπαρξης των παραπάνω πρακτόρων στο MAOM-QAP.

3.7 Βελτίωση της Εκπαίδευσης των Ακμών με τη Χρήση Ενισχυτικής Μάθησης

Στο άρθρο [38] προτείνεται ένα μοντέλο ενισχυτικής μάθησης (Reinforcement Learning, RL) το οποίο βασίζεται στον αλγόριθμο Branch and Bound (B&B) [2] και μοντελοποιείται βάση της Μαρκοβιανής διαδικασίας αποφάσεων (Markov Decision Process, MDP). Για την επίλυση του προβλήματος χρησιμοποιήθηκε η μείωση της χαλάρωσης του γραμμικού προγραμματισμού (Reduced Linear Programming Relaxation), τα βήματα της οποίας δίνονται παρακάτω.

- Αφαίρεση των τιμών που δεν μεταβάλλονται και ένταξη αυτών στο σύνολο των περιορισμών.
- Αφαίρεση των μικρών τιμών από το σύνολο των περιορισμών.

Επίσης ορίστηκε το πρωταρχικό-διπλό δίκτυο πολιτικής (Primal Dual Policy Net) ως ένα δίκτυο πολιτικής βασισμένο στη χαλάρωση του γραμμικού προγραμματισμού. Στην τεχνική χρησιμοποιήθηκαν νευρωνικά δίκτυα δύο επιπέδων, τα οποία μοντελοποιήθηκαν από τις παρακάτω εξισώσεις:

- Μοντελοποίηση περιορισμών: $Y_i \leftarrow f_c \left(Y_i, \sum_j A_{ij} m_c(X_j) \right)$
- Μοντελοποίηση μεταβλητών: $X_j \leftarrow f_v \left(X_j, \sum_j A_{ij} m_v(Y_i) \right)$

Ο πράκτορας εκπαιδεύεται από μια εξελισσόμενη στρατηγική, παρόμοια του NSR-ES [16]. Στον αλγόριθμο B&B, η διαδικασία επίλυσης παρουσιάζεται με τη

μορφή δέντρου αναζήτησης κάθε φύλλο του οποίου δίνει ένα λυμένο υποπρόβλημα. Κάθε υποπρόβλημα παρουσιάζεται ως μια περιοχή η οποία ονομάζεται πολύτοπο (polytope). Για κάθε ένα από αυτά ορίζεται η συνάρτηση βάρους και η συνάρτηση απόστασης μεταξύ δύο πολυτόπων. Επίσης ορίζεται η απόσταση Wasserstein ως η βέλτιστη απόσταση μεταφοράς δυο απεικονίσεων [69, 56] και δίνεται από τον τύπο $D(b_1, b_2) = \min_{\Gamma} \sum_{i,j} \Gamma_{ij} W_{ij}(b_1, b_2)$. Σε αυτόν, το Γ_i δηλώνει το κανονικοποιημένο βάρος του πολυτόπου b_i .

Στο άρθρο παρουσιάζεται μια νέα εξελικτική στρατηγική αναζήτησης (Novelty Search Evolutionary Strategy) το αποτέλεσμα της οποίας υπολογίζεται από τη σχέση $N(\theta, Q, M) = \frac{1}{k} \sum_{\pi_j \in kNN(M, \theta)} D(b(\pi_\theta, Q), b(\pi_j, Q))$. Σε αυτήν, το $kNN(M, \theta)$ δηλώνει τον k κοντινότερο γείτονα της πολιτικής π_θ στο M . Οι τιμές των D, N προσδιορίζονται από τη συμπεριφορά του κλαδέματος του δέντρου κατά τη διάρκεια εύρεσης της λύσης.

Για τα πειράματα χρησιμοποιήθηκαν τρία σύνολα δεδομένων τα οποία είναι το σύνολο κάλυψης (set covering [6]), το σύνολο μέγιστης ανεξαρτησίας (maximum independent set [21]) και το σύνολο χωροθέτησης μονάδων χρησιμοποιώντας τη χωρητικότητα (capacitated facility location [4]). Οι πράκτορες που συγκρίθηκαν με το προτεινόμενο μοντέλο είναι οι παρακάτω:

- Reliability Pseudocost Branch (RPB) [1]
- Full Strong Branching (FSB) [27]
- Vanilla Full Strong Branching (VFS) [27]
- Support Vector Machine (SVM) rank [39]
- GCN approach [28]

Για τον υπολογισμό της ποιότητας των αποφάσεων των αλγορίθμων, πραγματοποιήθηκαν εκατό πειράματα. Σε αυτά ορίστηκε η έννοια της νίκης ως ο ελάχιστος αριθμός λυμένων κόμβων μεταξύ των πρακτόρων. Ο αριθμός των νικών δε συμπεριλήφθηκε για τους αλγόριθμους RPB, FSB. Αυτό οφείλεται στο πλεονέκτημα που τους δίνει η πληροφορία διακλάδωσης του γραμμικού προγραμματισμού (branching LP information).

Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα των μοντέλων. Παρατηρείται στον Πίνακα 3.4 ότι η προτεινόμενη τεχνική επιλύει τα προβλήματα με τον καλύτερο μέσο χρόνο. Επιπλέον, χρειάζεται τον μικρότερο αριθμό κόμβων για να επιλύσει το Set Covering και το Independent Set.

Στο μοντέλο ενισχυτικής μάθησης, οι πράκτορες μαθαίνουν να βρίσκουν μικρότερο πρωταρχικό όριο και μεγαλύτερη διπλή τιμή. Αναλύοντας το προτεινόμενο μοντέλο με τα VFS, SVM και VFC για τις παραπάνω τιμές, παρατηρήθηκε ότι το RL βρίσκει μικρότερο πρωταρχικό όριο στο αρχικό στάδιο επίλυσης. Αυτό του δίνει μεγαλύτερη δυνατότητα κλαδέματος και το καθιστά ταχύτερο κατά την επίλυση. Λόγω της επιρροής του στη διπλή τιμή, η αρχικοποίηση του ορίου γίνεται βάση της βέλτιστης τιμής [39]. Τα αποτελέσματα της σύγκρισης με τις παραπάνω τεχνικές έδειξαν ότι ενώ αρχικά το προτεινόμενο μοντέλο έχει το χειρότερο πλάτος, μεγιστοποιείται μακροπρόθεσμα η διπλή τιμή.

Στο άρθρο πραγματοποιήθηκε μια ακόμα σύγκριση η οποία αφορά διαφορετικούς τύπους πρακτόρων του μοντέλου RL. Οι πράκτορες αναφέρονται παρακάτω:

- PD policy + ES
- PD policy + NS-ES
- GCN + ES
- GCN + NS-ES

Από το πείραμα παρατηρήθηκε ότι ο πρώτος και ο δεύτερος πράκτορας έχουν το μεγαλύτερο κέρδος. Οι πολιτικές PD έχουν μεγαλύτερο περιθώριο βελτίωσης από τους πράκτορες GCN. Η παραπάνω μελέτη υποδηλώνει ότι η PD policy + NS-ES είναι απαραίτητη για το μοντέλο RL.

3.8 Μάθηση από Άκρο σε Άκρο και Βελτιστοποίηση Γράφων

Στο άρθρο [71] εξετάζεται μια απλοποιημένη έκδοση του k-means clustering. Για την επίλυση του χρησιμοποιήθηκαν νευρωνικά δίκτυα γράφων τα οποία ενσωματώνουν κόμβους του συνεχούς χώρου. Στόχος τους είναι η βελτιστοποίηση διακριτών γράφων. Μια δεύτερη μελέτη αφορά την ανάθεση του cluster ως λύση ενός διακριτού προβλήματος. Τα πειράματα των μοντέλων πραγματοποιήθηκαν για δύο κατηγορίες

προβλημάτων βελτιστοποίησης, αυτών που ο γράφος διαχωρίζεται και εκείνων που απαιτείται η επιλογή ενός υποσυνόλου.

Για την επίλυση του προβλήματος χρησιμοποιήθηκαν ρυθμίσεις που συνδυάζουν τη βελτιστοποίηση με τη μάθηση. Ως είσοδος ορίστηκε ένας γράφος, ενώ η μοντελοποίηση του έγινε με σκοπό την πρόβλεψη συνδέσμων. Για το πρόβλημα ορίστηκε ένας πίνακας A ο οποίος προβλέπεται από τον πίνακα A^{train} . Το πρόβλημα βελτιστοποίησης λύνεται από τον τύπο: $\max_{x \in X} f(x, A)$. Σε αυτόν, το x ορίστηκε ως η μεταβλητή απόφασης, η f ως η αντικειμενική συνάρτηση, το X ως το εφικτό σύνολο. Για την εκπαίδευση χρησιμοποιήθηκε ένα μοντέλο με σκοπό την εκτίμηση ενός πίνακα \bar{A} μέσω μιας συνάρτησης απώλειας.

Στο άρθρο παρουσιάζεται το CLUSTERNET ως ένα μοντέλο που αποτελείται από δύο διαφοροποιημένα συστήματα. Στο πρώτο στρώμα χρησιμοποιείται ο πίνακας εκπαίδευσης και τα χαρακτηριστικά των κόμβων που τους ενσωματώνουν στον γράφο. Στα πειράματα χρησιμοποιείται το GCN [40]. Επίσης, χρησιμοποιείται ένα δεύτερο στρώμα το οποίο πραγματοποιεί μια διαφοροποιημένη βελτιστοποίηση.

Στο CLUSTERNET χρησιμοποιούνται τα παρακάτω επίπεδα βελτιστοποίησης:

- Forward pass
- Backward pass
- Exact backward pass
- Approximate backward pass

Στο άρθρο παρουσιάζονται οι παρακάτω τομείς για την πρόβλεψη συνδέσμων με βελτιστοποίηση:

- Learning problem: Για τον ορισμό των παραμέτρων χρησιμοποιήθηκε ένα σύνολο δεδομένων το οποίο δεν περιλαμβάνεται στα αποτελέσματα. Σκοπός της μάθησης είναι η πρόβλεψη ακμών και η λύση του προβλήματος βελτιστοποίησης του προβλεπόμενου γράφου.
- Optimization problems: Στόχος του προβλήματος είναι η εύρεση μιας θέσης η οποία μεγιστοποιεί την αρθρωτή μορφή (modularity) [53].

-
- **Baseline learning methods:** Για το CLUSTERNET χρησιμοποιήθηκε ένα GCN δύο επιπέδων, ακολουθούμενο από ένα επίπεδο clustering. Το μοντέλο συγκρίθηκε με ένα GCN δύο σταδίων, ένα GCN βασισμένο στο [62] και ένα GCN-e2e. Από τα παραπάνω μοντέλα, το GCN δύο επιπέδων είχε την καλύτερη απόδοση.
 - **Baseline optimization approaches:** Για τις κατηγορίες προβλημάτων κοινωνικής ανίχνευσης (community detection) και χωροθέτησης μονάδων (facility location), η σύγκριση αυτών έγινε με ειδικούς αλγόριθμους που βρέθηκαν στη βιβλιογραφία. Η πρώτη κατηγορία συγκρίθηκε με τους αλγόριθμους βελτιστοποίησης CNM [13], Newman [52], SC [70]. Η δεύτερη κατηγορία συγκρίθηκε με το greedy heuristic ως προς το σημείο με τη μεγαλύτερη οριακή βελτίωση της αντικειμενικής τιμής και τον αλγόριθμο του Gonzalez [29].
 - **Datasets:** Τα σύνολα δεδομένων που χρησιμοποιήθηκαν είναι τα ακόλουθα: cora [63], citeseer [63], protein [15], adol [14], fb [41]. Στο πρόβλημα χωροθέτησης μονάδων, οι κόμβοι των cora και citeseer έχουν κάποια χαρακτηριστικά βασισμένα στη βιβλιογραφία. Για τα υπόλοιπα σύνολα δεδομένων δημιουργήθηκαν χαρακτηριστικά node2vec χωρίς επίβλεψη [30] χρησιμοποιώντας τις εκπαιδευμένες ακμές.

Στη συνέχεια, δίνονται οι Πίνακες 3.5, 3.6 με τα αποτελέσματα των προβλημάτων community detection και facility location για $K = 5$ clusters σε μεμονωμένους γράφους.

Το CLUSTERNET έχει την καλύτερη απόδοση στα περισσότερα πειράματα. Στη διαδικασία της βελτιστοποίησης παρουσιάζει αρκετά ανταγωνιστική απόδοση. Για το πρόβλημα κοινωνικής ανίχνευσης παρατηρείται μεγάλη βελτίωση σε σχέση με τα υπόλοιπα μοντέλα. Στο πρόβλημα χωροθέτησης μονάδων, η απόδοση είναι όμοια με αυτή των πιο γνωστών μοντέλων.

Για τα παραπάνω μοντέλα πραγματοποιήθηκε μια δεύτερη σειρά πειραμάτων η οποία είχε στόχο τη διερεύνηση της ικανότητας μάθησης γενικευμένων στρατηγικών βελτιστοποίησης. Σε αυτή τη διαδικασία, αρχικά πραγματοποιήθηκε η εκπαίδευση των μοντέλων μέσω ενός συνόλου γράφων το οποίο χωρίστηκε σε δύο κατηγορίες. Στην πρώτη παρουσιάστηκε μία συνθετική γεννήτρια [72] βασισμένη στο μοντέλο χωρικής προτίμησης [7]. Χρησιμοποιήθηκαν 20 γράφοι για την εκπαίδευση, 10 για

την επικύρωση και 30 για τη δοκιμή. Στη δεύτερη, το σύνολο δεδομένων επιλέχθηκε από τον διαχωρισμό ενός γράφου σε 20 αντικείμενα χρησιμοποιώντας το metis [37]. Από την προηγούμενη διαδικασία δημιουργήθηκαν 10 γράφοι εκπαίδευσης, 2 για επικύρωση και 8 για δοκιμή. Κατά τη διαδικασία της δοκιμής εμφανίστηκε το 40% των ακμών. Η διαδικασία πραγματοποιήθηκε για το Learning + optimization όπως και στα προηγούμενα παραδείγματα.

Στον Πίνακα 3.7 απεικονίζονται τα αποτελέσματα των δύο προηγούμενων τεχνικών. Στόχος του είναι η αξιολόγηση των γενικευμένων αναπαραστάσεων της εκπαίδευσης.

Το CLUSTERNET έχει την καλύτερη απόδοση σε όλα τα μοντέλα εκτός από ένα. Στο πρόβλημα χωροθέτησης μονάδων, το σύνολο δεδομένων pubmed έχει τη δεύτερη καλύτερη απόδοση. Στο τέλος του πίνακα δίνονται τα αποτελέσματα του προτεινόμενου μοντέλου σε συνδυασμό με την τεχνική finetune. Από τα αποτελέσματα παρατηρείται ότι η απόδοση του CLUSTERNET συνήθως βελτιώνεται. Το κέρδος μπορεί να αυξηθεί εάν υπάρχει επιπλέον χρόνος εκπαίδευσης.

Πίνακας 3.3: Λύσεις για διαφορετικές τιμές των G , ω [74]

G	ω	$\beta\chi$	η	cF	Q	Set of reinforced links RL	J							
							3	8	10	13	14	18	22	
300	[0.1, 0.2]	507	10.54	106	193	(1-2), (4-5), (6-8), (10-11), (3-12), (7-18)	P	9	9	9	9	17	9	17
							B	17	17	17	17	9	17	9
								P	9	9	9	21	21	9
	[0.3, 0.5]	357	9.10	108	191	(1-2), (4-5), (6-8), (3-12), (7-18), (20-21)	B	21	21	21	9	9	21	9
							P	9	9	9	23	23	9	23
	[0.6, 0.9]	304	6.89	115	180	(1-2), (4-5), (6-8), (3-12), (7-18), (23-24)	B	23	23	23	9	9	23	9
							P	16	16	15	15	15	16	15
350	0.1	396	10.85	131	215	(1-2), (4-5), (6-8), (3-12), (7-18), (16-18), (21-22)	B	15	15	16	16	16	15	16
							P	9	9	9	23	23	9	23
	0.2	304	10.13	165	182	(1-2), (4-5), (6-8), (3-12), (7-18), (16-18), (18-20)	B	17	17	17	17	9	17	9
							P	9	9	9	23	23	16	23
	[0.3, 0.5]	260	9.57	176	168	(1-2), (4-5), (6-8), (9-10), (3-12), (7-18)	B	16	16	16	9	16	9	9
							P	4	16	16	23	23	16	23
	[0.6, 0.9]	219	6.69	191	151	(1-2), (4-5), (6-8), (3-12), (7-18)	B	16	4	4	16	16	4	16
							P	16	16	15	15	15	16	15
400	0.1	396	11.73	131	269	(1-2), (4-5), (6-8), (10-11), (3-12), (12-13), (7-18), (16-18), (21-22)	B	15	15	16	16	16	15	16
							P	9	9	9	23	23	16	23
	[0.2, 0.3]	260	10.82	176	222	(1-2), (3-4), (6-8), (5-9), (12-13), (7-18)	B	16	16	16	16	16	9	9
							P	4	16	16	23	23	16	23
	0.4	219	9.83	241	151	(1-2), (4-5), (6-8), (3-12), (7-18)	B	17	17	17	16	16	17	17
							P	4	9	9	23	23	16	23
	[0.5, 0.9]	209	9.43	247	151	(1-2), (4-5), (6-8), (3-12), (7-18)	B	16	16	16	16	16	21	16
							P	9	9	9	21	21	16	21
450	0.1	313	12.03	169	275	(1-2), (4-5), (6-8), (3-12), (11-14), (7-18), (21-22)	B	15	15	16	16	16	15	16
							P	9	9	9	23	23	16	23
	0.2	260	11.60	176	274	(1-2), (3-4), (6-8), (5-9), (7-18), (16-18), (21-22)	B	16	16	16	9	16	9	16
							P	4	16	16	23	23	16	23
	[0.3, 0.5]	219	10.91	191	258	(1-2), (3-4), (6-8), (5-9), (7-18), (21-22)	B	16	4	4	4	16	4	16
							P	4	9	9	23	23	16	23
	0.6	209	10.27	247	199	(1-2), (4-5), (6-8), (3-12), (7-18), (21-22)	B	16	16	16	16	4	9	16
							P	4	9	9	23	23	7	23
	[0.7, 0.9]	200	9.46	258	187	(1-2), (4-5), (6-8), (3-12), (7-18), (22-23)	B	7	7	7	7	4	9	7
							P	9	9	9	23	23	16	23
500	[0.1, 0.2]	260	12.47	176	324	(1-2), (4-5), (6-8), (8-9), (10-11), (3-12), (8-16), (7-18)	B	16	16	16	9	16	9	9
							P	4	16	16	23	23	16	23
	[0.3, 0.4]	219	11.74	191	308	(1-2), (3-4), (6-8), (5-9), (12-13), (7-18), (18-20), (21-22), (13-24)	B	16	4	4	16	16	4	16
							P	4	9	9	23	23	16	23
	0.5	209	11.3	247	246	(1-2), (3-4), (6-8), (5-9), (7-18), (22-23)	B	16	16	16	16	4	4	16
							P	4	9	9	23	23	16	21
	0.6	202	10.76	299	199	(1-2), (4-5), (6-8), (3-12), (7-18), (21-22)	B	21	16	16	16	4	21	16
							P	4	9	9	23	23	7	21
	[0.7, 0.9]	193	9.46	310	187	(1-2), (4-5), (6-8), (3-12), (7-18), (22-23)	B	7	7	7	7	21	21	7

Πίνακας 3.4: Εκτίμηση της πολιτικής ανά πρόβλημα [38]

Method	T_{avg}	N_{avg}	Wins	T_{avg}	N_{avg}	Wins	T_{avg}	N_{avg}	Wins
FSB	99.73	146	na/100	19.19	140	na/100	27.16	964	na/100
RPB	12.64	1292	na/100	3.06	250	na/100	21.39	1449	na/100
VFS	1935.35	1518	5/75	244.14	1304	0/100	173.50	1848	31/100
SVM	21.19	1581	1/100	10.83	498	1/100	29.64	2096	17/100
GCN	10.37	1104	28/100	1.56	418	2/100	26.31	1752	13/100
RL	7.91	820	66/100	1.26	200	97/100	20.85	1640	39/100

Set Covering

Independent Set

Facility Location

Πίνακας 3.5: Επίδοση στο πρόβλημα ανίχνευσης κοινότητας [71]

	Learning + Optimization					Optimization				
	cora	cite	prot	adol	fb	cora	cite	prot	adol	fb
ClusterNct	0.54	0.55	0.29	0.49	0.30	0.72	0.73	0.52	0.58	0.76
GCN-c2c	0.16	0.02	0.13	0.12	0.13	0.19	0.03	0.16	0.20	0.23
Train-CNM	0.20	0.42	0.09	0.01	0.14	0.08	0.34	0.05	0.57	0.77
Train-Newman	0.09	0.15	0.15	0.15	0.08	0.20	0.23	0.29	0.30	0.55
Train-SC	0.03	0.02	0.03	0.23	0.19	0.09	0.05	0.06	0.49	0.61
GCN-2stage-CNM	0.17	0.21	0.18	0.28	0.13	-	-	-	-	-
GCN-2stage-Newman	0.00	0.00	0.00	0.14	0.02	-	-	-	-	-
GCN-2stage-SC	0.14	0.16	0.04	0.31	0.25	-	-	-	-	-

Πίνακας 3.6: Επίδοση στο πρόβλημα χωροθέτησης μονάδων [71]

	Learning + Optimization					Optimization				
	cora	cite	prot	adol	fb	cora	cite	prot	adol	fb
ClusterNct	10	14	6	6	4	9	14	6	5	3
GCN-c2c	12	15	8	6	5	11	14	7	6	5
Train-greedy	14	16	8	8	6	9	14	7	6	5
Train-gonzalez	12	17	8	6	6	10	15	7	7	3
GCN-2Stage-greedy	14	17	8	7	6	-	-	-	-	-
GCN-2Stage-gonzalez	13	17	8	6	6	-	-	-	-	-

Πίνακας 3.7: Σύγκριση επιδόσεων ως προς την τεχνική Finetune [71]

	Community Detection					Facility Location				
	synthetic		pubmed		No finetune	synthetic		pubmed		
	Avg.	%	Avg.	%		Avg.	%	Avg.	%	
No finetune										
ClusterNet	0.57	26/30	0.30	7/8	ClusterNet	7.90	25/30	7.88	3/8	
GCN-c2c	0.26	0/30	0.01	0/8	GCN-c2c	8.63	11/30	8.62	1/8	
Train-CNM	0.14	0/30	0.16	1/8	Train-greedy	14.00	0/30	9.50	1/8	
Train-Newman	0.24	0/30	0.17	0/8	Train-gonzalez	10.30	2/30	9.38	1/8	
Train-SC	0.16	0/30	0.04	0/8	2Stage-grccdy	9.60	3/30	10.00	0/8	
2Stage-CNM	0.51	0/30	0.24	0/8	2Stage-gonz.	10.00	2/30	6.88	5/8	
2Stage-Newman	0.01	0/30	0.01	0/8	ClstrNct-1 train	7.93	12/30	7.88	2/8	
2Stage-SC	0.52	4/30	0.15	0/8						
ClstrNet-1 train	0.55	0/30	0.25	0/8						
Finetune					Finetune					
ClstrNct-ft	0.60	20/30	0.40	2/8	ClstrNct-ft	8.08	12/30	8.01	3/8	
ClstrNct-ft-only	0.60	10/30	0.42	6/8	ClstrNet-ft-only	7.84	16/30	7.76	4/8	

Κεφάλαιο 4

Ανάλυση Μοντέλων

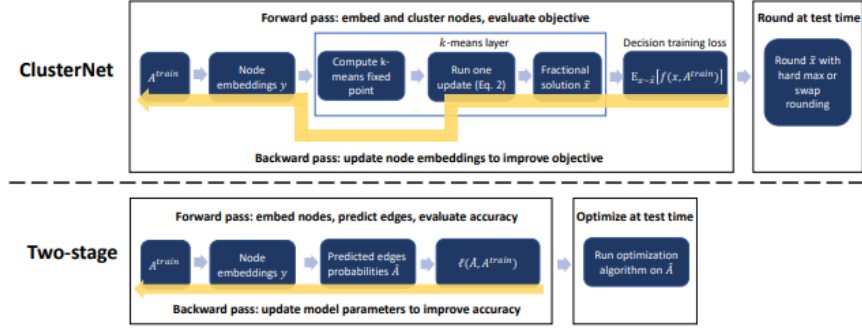
4.1 CLUSTERNET System

Το προτεινόμενο μοντέλο συνδυάζει δύο διαφορεικά στοιχεία σε ένα σύστημα εκπαίδευσης από άκρο σε άκρο. Σε αυτό χρησιμοποιείται ένα στρώμα ενσωμάτωσης γράφων. Τα χαρακτηριστικά του δίνονται μέσω του πίνακα γειτνίασης A^{train} . Σε αυτόν περιέχονται οι εκπαιδευμένες ακμές του γράφου. Στα πειράματα χρησιμοποιούνται Graph Convolutional Networks (GCNs) [40] για την ενσωμάτωση των γράφων.

Στη συνέχεια χρησιμοποιείται ένα επίπεδο για βελτιστοποίηση μέσω διαφορικών εξισώσεων. Το επίπεδο παίρνει ως είσοδο τις ενσωματώσεις του συνεχούς χώρου και τις χρησιμοποιεί για την παραγωγή μιας λύσης x για το πρόβλημα βελτιστοποίησης γράφου. Συγκεκριμένα, προτείνεται η χρήση ενός επιπέδου για την υλοποίηση μιας διαφορικής εκδοχής της ομαδοποίησης k-means. Το στρώμα παράγει μία χαλαρή ανάθεση των κόμβων σε ομάδες, μαζί με τα κέντρα του ενσωματωμένου χώρου. Ακολουθεί η γραφική απεικόνιση του CLUSTERNET ως προς το μοντέλο Two-stage στο Σχήμα 4.1.

Στη συνέχεια αναλύεται το επίπεδο βελτιστοποίησης. Αρχικά αναλύονται οι μέθοδοι forward και backward pass οι οποίες χρησιμοποιούνται στη διαδικασία της ομαδοποίησης. Τέλος εξηγείται ο τρόπος με τον οποίο ερμηνεύονται οι αναθέσεις ομάδων ως λύσεις του προβλήματος βελτιστοποίησης γράφων.

Forward Pass Για την υλοποίηση της μεθόδου αρχικά ορίζεται το x_j ως η ενσωμάτωση του κόμβου j και το μ_k ως το κέντρο της k ομάδας. Μία τρίτη έννοια



Σχήμα 4.1: Απεικόνιση μοντέλου CLUSTERNET και Two-Stage

που πρέπει να οριστεί είναι αυτή του r_{jk} ως ο βαθμός για τον οποίο αποφασίζεται ποιός κόμβος j ανατίθεται την ομάδα k . Στην κλασική προσέγγιση του k-means, ο βαθμός ανάθεσης δηλώνει ένα δυαδικό μέγεθος. Στο προτεινόμενο μοντέλο χρησιμοποιείται μία τεχνική χαλάρωσης μέσω μιας κλασματικής τιμής, για την οποία ισχύει $\sum_k r_{jk} = 1$ για κάθε j .

Πιο συγκεκριμένα, για την ανάθεση κάθε σημείου σε ένα κέντρο χρησιμοποιείται μια συνάρτηση soft-max η οποία βασίζεται στην απόσταση. Η τιμή της δίνεται από τη σχέση:

$$r_{jk} = \frac{\exp(-\beta \|x_j - \mu_k\|)}{\sum_i \exp(-\beta \|x_j - \mu_i\|)} \quad (4.1)$$

Στην παραπάνω σχέση, για τον υπολογισμό της νόρμας χρησιμοποιείται το αρνητικό συνημίτονο. Ωστόσο μπορεί να χρησιμοποιηθεί οποιαδήποτε νόρμα. Η μεταβλητή β είναι μία υπερ παράμετρος του μοντέλου η οποία τείνει στο άπειρο για την κλασική προσέγγιση του k-means.

Η βελτιστοποίηση των κέντρων των ομάδων μπορεί να πραγματοποιηθεί μέσω μιας επαναληπτικής διαδικασίας ανάλογης της ενημέρωσης του k-means. Σε αυτή πραγματοποιείται μια εναλλακτική ρύθμιση στα κέντρα των ομάδων η οποία δίνεται από τις σχέσεις:

$$\mu_k = \frac{\sum_j r_{jk} x_j}{\sum_j r_{jk}} \quad \forall k = 1 \dots K \quad (4.2)$$

$$r_{jk} = \frac{\exp(-\beta \|x_j - \mu_k\|)}{\sum_i \exp(-\beta \|x_j - \mu_i\|)} \quad k = 1 \dots K, j = 1 \dots n \quad (4.3)$$

Η επαναλαμβανόμενη σύγκλιση οδηγεί σε ένα σταθερό σημείο στο οποίο το μ παραμένει σταθερό [45]. Η έξοδος του forward pass δίνεται από το ζεύγος (μ, r) .

Backward Pass Βασίζεται στο θεώρημα έμμεσης συνάρτησης (Implicit Function Theorem). Στόχος της μεθόδου είναι η διαφοροποίηση του σταθερού σημείου του forward pass η οποία υπολογίζεται μέσω των παραγώγων $\frac{\partial \mu}{\partial x}$, $\frac{\partial r}{\partial x}$. Η χρήση του θεωρήματος επιτρέπει την οπισθοπροβολή (backpropagate) της κλήσης της απώλειας απόφασης. Ορίζεται από τη συνάρτηση:

$$f_{i,l}(\mu, x) = \mu_i^j - \frac{\sum_j r_{jk} x_j^l}{\sum_j r_{jk}} \quad (4.4)$$

Η έξοδος του forward pass δίνει το σταθερό σημείο εάν ισχύει επαναληπτικά η σχέση $f_{i,l}(\mu, x) = 0$. Η εφαρμογή του παραπάνω θεωρήματος οδηγεί στη σχέση $\frac{\partial \mu}{\partial x} = \left[\frac{\partial f(\mu, x)}{\partial \mu} \right]^{-1} \frac{\partial f(\mu, x)}{\partial x}$. Το $\frac{\partial r}{\partial x}$ υπολογίζεται από την προηγούμενη σχέση με τη χρήση του αλυσιδωτού κανόνα (chain rule).

Exact backward pass Σε αυτό το σημείο αναλύεται ο υπολογισμός του $\frac{\partial \mu}{\partial x}$. Για τον υπολογισμό των τιμών $\frac{\partial f(\mu, x)}{\partial \mu}$, $\frac{\partial f(\mu, x)}{\partial x}$ χρειάζεται χρονική πολυπλοκότητα $O(nKp^2)$ και $O(npK^2)$. Για τη δεύτερη σχέση απαιτείται η αναστροφή της παραγώγου η οποία εισάγει επιπλέον πολυπλοκότητα $O(K^3p^3)$. Λόγω των παραπάνω, απαιτείται η χρήση μιας νέας τεχνικής η οποία αναλύεται στη συνέχεια.

Approximate backward pass Στον πίνακα της παραγώγου $\frac{\partial f}{\partial \mu}$ υπερέχουν τα στοιχεία της διαγωνίου. Για μεγάλες τιμές, η παράμετρος β εισάγει σταθερότητα. Εμπειρικά έχει παρατηρηθεί ότι το β έχει μεγάλες τιμές. Συμπεραίνεται επομένως ότι η παράγωγος μπορεί να υπολογιστεί κατά προσέγγιση από τη διαγώνιο της. Αυτό μας δίνει τη σχέση $\frac{\partial \mu}{\partial x} \approx \frac{\partial f}{\partial x}$.

Η παραπάνω υπόθεση μπορεί να χρησιμοποιηθεί για καλά διαχωρίσιμες και ισορροπημένες ομάδες. Η ποιότητα της προσέγγισης εξαρτάται από τις παραμέτρους:

- β , η σκληρότητα ανάθεσης των ομάδων
- δ , ο δείκτης διαχωρισμού των ομάδων
- α , ο δείκτης ισορροπίας του μεγέθους των ομάδων

Η κλήση δίνεται προσεγγιστικά από τη σχέση $\frac{\partial f}{\partial \mu} = I$ ενώ υπολογίζεται αναπτύσσοντας μια επανάληψη των Εξισώσεων 4.2 και 4.3 του forward pass. Από την

Εξίσωση 4.4 παρατηρείται ότι η μεταβλητή μ_i^l σέβεται το x . Η συνάρτηση ενημέρωσης του μ_k δίνεται από τη σχέση:

$$\frac{\partial f_k}{\partial x} = \frac{\partial \sum_j r_{jk} x_j}{\partial x \sum_j r_{jk}} \quad (4.5)$$

Οι ενημερώσεις του forward pass είναι γραμμένες διαφορικές εξισώσεις. Λόγω αυτού μπορούν να χρησιμοποιηθούν διαφορικά εργαλεία (standard auto-differentiation tools) στην τελευταία ενημέρωση του για τον υπολογισμό της προσέγγισης του backward pass. Η τελευταία επανάληψη απαιτεί πολυπλοκότητα $O(npK)$.

Λήψη λύσεων του προβλήματος βελτιστοποίησης (Obtaining solutions to the optimization problem) Οι αναθέσεις ομάδων r πραγματοποιούν χαλαρή κατάτμηση του γράφου. Μια γενικευμένη συνεχής αντικειμενική συνάρτηση προκύπτει από την τυχαία διαδικασία ανάθεσης ενός κόμβου j σε ένα διαμέρισμα με πιθανότητες που δίνονται από το r_j . Η διαδικασία επαναλαμβάνεται για όλους τους κόμβους ανεξάρτητα. Η απώλεια απόφασης εκπαίδευσης δίνεται από την παραπάνω διαδικασία. Μοντελοποιείται από τη σχέση:

$$\ell = E_{r^{hard} \sim r} [f(r^{hard}, A^{train})] \quad (4.6)$$

Σε αυτήν, το $r^{hard} \sim r$ δηλώνει την τυχαία ανάθεση. Με αυτόν τον τρόπο το ℓ μπορεί να διαφοριστεί ως προς το r για να υπολογιστεί σε κλειστή μορφή με ένα διαφορικό εργαλείο.

Η προτεινόμενη προσέγγιση μπορεί να εφαρμοστεί και στην περίπτωση που η προσδοκία δεν είναι σε κλειστή μορφή εφαρμόζοντας επαναληπτική δειγματοληψία του $r^{hard} \sim r$ και χρησιμοποιώντας μια πολιτική για τον υπολογισμό της κλίσης του στόχου. Στη διαδικασία της δοκιμής εφαρμόζεται ένα σκληρό μέγιστο στο r για τη λήψη της εκχώρησης κάθε κόμβου.

Οι λύσεις δίνονται στο συνεχές διάστημα από ένα διάνυσμα x το οποίο ανήκει στο $[0, 1]$, για το οποίο ισχύει $\|x\|_1 = K$. Διαισθητικά, το x_j δίνει την πιθανότητα εισαγωγής του στη λύση. Η προτεινόμενη προσέγγιση τοποθετεί μεγαλύτερη πιθανότητα στους κόμβους που βρίσκονται κοντά στα κεντρα των ομάδων. Σε κάθε κέντρο μ_i δίνεται μια μονάδα πιθανότητας, η οποία κατανέμεται στο x μέσω της

σχέσης:

$$a_{ij} = \text{softmin}(\eta \|x - \mu_1\|)_j \quad (4.7)$$

Η συνολική πιθανότητα που κατανέμεται στον κόμβο j δίνεται από τη σχέση:

$$b_j = \sum_{i=1}^K a_{ij} \quad (4.8)$$

Λόγω της απαίτησης ισχύς της σχέσης $b_j < 1$, το b_j εισάγεται σε μια σιγμοειδή συνάρτηση. Το x δίνεται από την σχέση $x = 2\sigma(\gamma b) - 0.5$. Σε αυτήν, το γ αποτελεί μια παράμετρο του συστήματος. Στην περίπτωση που το x υπερβεί τον προϋπολογισμό, η έξοδος δίνεται από το κλάσμα $\frac{Kx}{\|x\|_1}$.

Ακολουθώντας τη διαδικασία που περιγράφεται παραπάνω, για διακριτές λύσεις του $x^{hard} \sim x$ προκύπτει η σχέση $E_{x^{hard} \sim x} [f(x^{hard}, A^{train})]$.

Κατά τη διαδικασία της δοκιμής, απαιτείται η εύρεση ενός εφικτού διακριτού διανύσματος. Αξίζει να σημειωθεί ότι η στρογγυλοποίηση των καταχωρήσεων μπορεί να παράξει ένα διάνυσμα με περισσότερους από K κόμβους. Σε αυτό το σημείο εφαρμόζεται μία γενική προσέγγιση που βασίζεται στη στρογγυλοποίηση του σωλήνα [3]. Η λύση επιλέγεται βάση της καλύτερης απώλειας απόφασης των παρατηρήσιμων ακμών.

4.2 Equitable Mechanism Design in Markov Reward Process (EMD-MRP)

Για τη μοντελοποίηση του προβλήματος ορίζεται ένας γράφος $\mathbb{G} = (V, E)$. Σε αυτόν, το V δηλώνει ένα σύνολο κόμβων, ενώ το E ένα σύνολο κατευθυνόμενων και χωρίς βάρος ακμών. Επιπλέον ορίζεται το G ως το σύνολο των περιοχών. Τέλος, οι κόμβοι κέρδους συμβολίζονται με R και ανήκουν στο V .

Μια έννοια που εισάγεται κατά τη μοντελοποίηση του προβλήματος είναι το σωματίδιο p_g . Σε αυτό, το g δηλώνει την περιοχή στην οποία ανήκει. Επιπλέον, η κατανομή μιας περιοχής συμβολίζεται ως $\mu_0(g)$. Η κοντινότερη απόσταση μεταξύ του σωματιδίου και του κόμβου κέρδους ορίζεται ως $d(p_g, r)$. Η ωφέλιμη συνάρτηση

κάθε περιοχής δίνεται από τη σχέση:

$$u_g(p_g; E) = E_{p_g \sim \mu(g)} \frac{1}{d(p_g, r)} \quad (4.9)$$

Για λόγους απλούστευσης, το $u_g(p_g; E)$ αναφέρεται ως u_g . Η ωφέλιμη συνάρτηση μιας περιοχής g δίνεται από τη σχέση:

$$U_g = E_{p_g \sim \mu(g)}[u_g] \quad (4.10)$$

Αντίστοιχα, η μέση τιμή της για όλες τις περιοχές περιγράφεται από τη σχέση:

$$\bar{U}_g = \sum_{g \in G} \frac{U_g}{|G|} \quad (4.11)$$

Από τα παραπάνω, η ανισότητα που προκύπτει από την ωφέλεια των περιοχών ως προς το μέσο όρο δίνεται από τη σχέση $\sum_{g \in G} |U_g - \bar{U}_g|$. Η συνάρτηση αύξησης του γράφου για δοθέντα προϋπολογισμό B δίνεται από τη σχέση $e : \mathbb{G}, B \rightarrow \mathbb{G}'$ όπου $\mathbb{G}' = (V, E \cup E^u)$. Το E^u δηλώνει τις προστιθέμενες ακμές υπό τον περιορισμό του προϋπολογισμού. Για τον υπολογισμό της χρησιμοποιείται η απόσταση Hamming $D(\mathbb{G} - \mathbb{G}') < B$.

Άπληστη Βελτίωση της Ισότητας των Κέντρων (Greedy Equitable Centrality Improvement) Ως βάση του EMD-MRP χρησιμοποιήθηκε το μοντέλο Greedy Equitable Centrality Improvement (GECI). Σε αυτό ορίζεται μια περιοχή $N_g(E)$ ως το σύνολο των κόμβων για τους οποίους ισχύει $u \in V$, $u \notin E$ και ανήκουν στο συντομότερο μονοπάτι που οδηγεί στον κόμβο κέρδους της περιοχής. Το GECI ορίζεται για προϋπολογισμό B από τον Αλγόριθμο 1.

Σε αυτή, για κάθε αύξηση της ακμής E^u επιλέγεται η περιοχή g_{min} με τη μικρότερη ωφέλεια U_g . Στη συνέχεια επιλέγεται ένα ζεύγος κόμβων (u_{max}, v_{max}) για το σχηματισμό της προστιθέμενης ακμής. Το σύνολο των αυξήσεων των ακμών E^u ενημερώνεται μέσω της σχέσης $E^u := E^u \cup \{(u_{max}, v_{max})\}$. Η αύξηση του γράφου επαναλαμβάνεται έως το σημείο όπου ο αριθμός των προστιθέμενων ακμών ξεπεράσει τον προϋπολογισμό. Στο GAEA, η βελτιστοποίηση μοντελοποιείται μέσω της

Input : Ο γράφος $\mathbb{G}(V, E)$, το σύνολο των περιοχών N_g , ο προϋπολογισμός B
 $E^u := \emptyset$
for $b = 1, 2, \dots, B$ **do**
 $E^u := E \cup E^u$
 $g_{min} := \operatorname{argmin}\{U_g(E^e) | g \in G\}$
 for $u \in V | u \in N_{g_{min}}(E^e)$ **do**
 for $u \in V | u \in E^e$ **do**
 | Υπολογισμός του $U_{g_{min}}(E^e \cup \{(u, v)\})$
 end
 end
 $u_{max}, v_{max} := \operatorname{argmax}\{U_{g_{min}}(E^e \cup \{(u, v)\})\}$
 $E^u := E^e \cup \{(u_{max}, v_{max})\}$
end
return E^u

Αλγόριθμος 1: Άπληστη βελτίωση της ισότητας των κέντρων

σχέσης:

$$J = \max \mathbb{E}_{g \in G}[U_g] \quad \alpha \nu \quad \sum_{g \in G} |U_g - \bar{U}_G| = 0, \quad \sum_{g \in G} \|E^u\|_0 < B \quad (4.12)$$

Η Σχέση 4.12 δηλώνει τον περιορισμό ως προς τη ισότητα και τον προϋπολογισμό. Οι περιορισμοί δεν είναι διαφορίσιμοι. Για αυτο τον λόγο δεν μπορούν να λυθούν από κάποιο αλγόριθμο.

Σχεδιασμός Μηχανισμού Ισότητας μέσω της Μαρκοβιανής Διαδικασίας Κέρδους (Equitable Mechanism Design in MRP) Ο γράφος πλαισιώνεται από ένα μηχανισμό, βασισμένο στη Μαρκοβιανή Διαδικασία Κέρδους (Markov Reward Process (MRP)). Το MRP αποτελείται από ένα πεπερασμένο σύνολο καταστάσεων S . Η δυναμική του ορίζεται από ένα σύνολο πιθανοτήτων μετάβασης P , μια συνάρτηση κέρδους R και έναν ορίζοντα ο οποίος δηλώνει το μέγιστο αριθμό βημάτων T μιας τυχαίας επανάληψης. Οι καταστάσεις S αντιστοιχούν στους κόμβους του γράφου. Η δυναμική του συστήματος ορίζεται από τη σχέση $P = D^{-1}E$ με διαγώνιο πίνακα $D(i, i) = \sum_j E_{(i,j)}$. Σε αντίθεση με τη Μαρκοβιανή Διαδικασία Απόφασης (Markov Decision Processes (MDP)), το MRP δε βελτιώνει την πολιτική. Αντ' αυτού βελτιώνει τη δυναμική μέσω της συνάρτησης αξίας. Η μοντελοποίηση της δίνεται από τη σχέση:

$$v_g(s_0) = \sum_{t=0}^{T-1} \gamma^t R P^t s_0 \quad (4.13)$$

Το γ αποτελεί ένα συντελεστή μείωσης. Η συνάρτηση αξίας κάθε περιοχής g ορίζεται μέσω της σχέσης:

$$V^g = \mathbb{E}_{s_0 \sim \mu(g)}[v^g(s_0)] \quad (4.14)$$

Το $E \in \{0, 1\}^{|S| \times |S|}$ δηλώνει τον πίνακα γειτνίασης των αυξήσεων των ακμών του μη επεξεργασμένου γραφού. Για τη μοντελοποίηση του προβλήματος δημιουργείται ένας πίνακας $A \in \{0, 1\}^{|S| \times |S|}$ ο οποίος χρησιμοποιείται σα μάσκα. Σκοπός του είναι η μοντελοποίηση μιας τοπολογίας για τον περιορισμό της επεξεργασίας των ακμών. Το σύνολο E^u περιέχει τις ακμές που χρειάζεται να προστεθούν. Γίνεται διαφορίσιμο μέσω συνεχούς χαλάρωσης που δίνεται από τη σιγμοειδή συνάρτηση του Gumbel [36]

$$E^u(i, j) = \frac{1}{1 + \exp(-\phi(\bar{0}) + g_i)/\tau)}, \forall i, j \in S \quad (4.15)$$

όπου $g_i = -\log(-\log(U))$ με το U να δηλώνει το θόρυβο Gumbel και το τ μια μεταβλητή εξασθένησης της εκπαίδευσης. Δίνεται από τη σχέση $\tau \leftarrow \tau * \nu$ ενώ αλλάζει με το πέρασ της εποχής. Το ν αποτελεί ένα συντελεστή μείωσης. Ο στόχος του προβλήματος δίνεται από τη σχέση:

$$E^u = \operatorname{argmax} \sum_{g \in G} V_g \text{ αν } \sum_{g \in G} |V_g - \bar{V}_G| = 0, \sum_{g \in G} \|E^u\|_0 < B \text{ όπου } \bar{V}_G = \sum_{g \in G} \frac{V_g}{|G|} \quad (4.16)$$

Στη συνέχεια αναδιατυπώνεται η βελτιστοποίηση με περιορισμούς. Στη νέα μοντελοποίηση, οι περιορισμοί αφαιρούνται μέσω της επαύξησης Lagrange [54]. Η σχέση δίνεται παρακάτω:

$$J = -\min_{E^u} \sum_{g \in G} V_g - \mu_1 \left(\sum_{g \in G} V_g - \bar{V}_G \right)^2 - \mu_2 \left(\min(0, \sum_{g \in G} \|E^u\|_0 - B) \right)^2 - \lambda_1 \left(\sum_{g \in G} |V_g - \bar{V}_G| \right) - \lambda_2 \left(\min(0, \sum_{g \in G} \|E^u\|_0 - B) \right) \quad (4.17)$$

Τα μ_1, μ_2 αποτελούν υπερπαραμέτρους του συστήματος ενώ τα λ_1, λ_2 τους συντελεστές Lagrange. Οι ενημερώσεις των λ_1, λ_2 μετά το πέρασ κάθε εποχής δίνονται παρακάτω:

$$\lambda_1^{new} \leftarrow \lambda_1^{old} + \mu_1 \left(\sum_{g \in G} |V_g - \bar{V}_G| \right) \quad (4.18)$$

$$\lambda_2^{new} \leftarrow \lambda_2^{old} + \mu_2 \left(\min(0, \sum_{g \in G} \|E^u\|_0 - B) \right) \quad (4.19)$$

Ο στόχος μαθαίνει τη δυναμική του MRP. Η Σχέση 4.17 εισάγει θόρυβο στην κλήση της στοχαστικής παρτίδας. Για την αποτροπή του φαινομένου, πραγματοποιείται εκπαίδευση στον κύριο ή σε έναν από τους περιορισμούς κάθε παρτίδας. Στην εκπαίδευση, ο στόχος J βελτιστοποιείται χωρίς περιορισμούς. Με την πάροδο του χρόνου εισάγεται ο περιορισμός της ισότητας βάση του προϋπολογισμού. Οι υπερπαραμέτροι είναι ρυθμισμένες για το καλύτερο δυνατό αποτέλεσμα. Η στρατηγική αναφέρεται λεπτομερώς στον Αλγόριθμο 2.

Input : Ο αρχικός πίνακας βαρών W_g^0
Initialize $\tau = 1, \lambda_1 = 0, \lambda_2 = 0, \mu_1, \mu_2, \alpha$
 ενημέρωση περιορισμού = True
for εως τη σύγκλιση **do**
 for για κάθε βελτιστοποιημένο παρτίδα του ADAM **do**
 το δείγμα $s_0 \sim \mu(g)$
 if ενημέρωση περιορισμού **then**
 $\theta \leftarrow \theta - \alpha \nabla_{\theta} (-\mu_1 (\sum_{g \in G} |V_g - \bar{V}_G|)^2 + \lambda_1 (\sum_{g \in G} |V_g - \bar{V}_G|))$
 $\theta \leftarrow \theta - \alpha \nabla_{\theta} (-\mu_2 (\min(0, \sum_{g \in G} E_g - B))^2 - \lambda_2 (\min(0, \sum_{g \in G} E_g - B)))$
 else
 $\theta \leftarrow \theta - \alpha \nabla (\sum_{g \in G} V^g)$
 end
 ενημέρωση περιορισμού \leftarrow μη ενημέρωση περιορισμού
 if πληρούνται οι προϋποθέσεις για ισότητα **then**
 | Ενημέρωση του λ_1 (Σχέση 4.18)
 if πληρούνται οι προϋποθέσεις για επεξεργασία **then**
 | Ενημέρωση του λ_2 (Σχέση 4.19)
 if πληρούνται οι προϋποθέσεις για το τ **then**
 | $\tau \leftarrow \tau * \nu$
 end
end

Αλγόριθμος 2: Σχεδιασμός μηχανισμού ισότητας μέσω της Μαρκοβιανής διαδικασίας κέρδους

Το πρόβλημα χωροθέτησης μονάδων αποτελεί ειδική περίπτωση του GAEA. Η επίλυση του επιτυγχάνεται μέσω της τοποθέτησης κόμβων κέρδους αντί της αύξησης των ακμών του γράφου. Σκοπός της διαδικασίας είναι η δίκαιη πρόσβαση των σωματιδίων p_g στους κόμβους κέρδους των περιοχών [42, 25]. Η δίκαιη χωροθέτηση μονάδων πραγματοποιείται μέσω μικρών αλλαγών στο παραπάνω μοντέλο. Η δυναμική του MRP παραμένει σταθερή. Ο στόχος δίνεται από το διάνυσμα κέρδους $R \in \{0, 1\}^{|S|}$. Προσαρμόζοντας τις Σχέσεις 4.15 και 4.16 στο πρόβλημα χωροθέτησης

μονάδων παράγονται οι σχέσεις:

$$R^u = \operatorname{argmax} \sum Vg \text{ αν } \sum_{g \in G} |V_g - \bar{V}_G| = 0, \quad \|R^u\|_0 < B \quad (4.20)$$

$$R^u(s) = \frac{1}{1 + \exp(-\phi(\bar{0}) + g_s)/\tau)}, \forall s \in S \quad (4.21)$$

Κεφάλαιο 5

Παρουσίαση Αποτελεσμάτων

Σε αυτό το κεφάλαιο αναλύεται η εκπαίδευση και η αξιολόγηση των μοντέλων ενισχυτικής μάθησης. Πιο συγκεκριμένα, η αξιολόγηση των μοντέλων έγινε με τη χρήση της τεχνικής ελαχίστου αθροίσματος (MiniSum). Στόχος των προβλημάτων MiniSum είναι η ελαχιστοποίηση της μέσης απόστασης μεταξύ των μονάδων και των σημείων ζήτησης. Κατά τη διαδικασία των πειραμάτων, πραγματοποιήθηκε η επίλυση του προβλήματος χωροθέτησης μιας μονάδας καθώς και του προβλήματος χωροθέτησης πολλαπλών μονάδων για αριθμό μονάδων 2, 3 και 4.

Για την αξιολόγηση των μοντέλων χρησιμοποιήθηκαν τα σύνολα δεδομένων: Cora [63] με 2,485 κόμβους και 12,623 ακμές, Citeseer [63] με 2,109 κόμβους και 9,446 ακμές, Protein Interaction Network [15] με 2,783 κόμβους και 14,797 ακμές. Επίσης, χρησιμοποιήθηκαν 20 γράφοι του Synthetic Distribution [73]. Οι γράφοι περιείχαν 500 κόμβους αλλά διέφεραν στον αριθμό των ακμών. Τέλος, χρησιμοποιήθηκε ένας γράφος 15 κόμβων και 59 ακμών για τη σύγκριση των μοντέλων σε μικρά προβλήματα. Πρέπει να σημειωθεί ότι οι παραπάνω τιμές δίνονται για πλήρεις γράφους.

Τα πειράματα υλοποιήθηκαν σε έναν υπολογιστή με επεξεργαστή Intel Core i7-6700HQ (2.60GHz), μνήμη RAM 16GB, λειτουργικό σύστημα Windows 10 (Έκδοση 20H2) και κάρτα γραφικών GeForce GTX 960M.

5.1 Συγκρινόμενα Μοντέλα

5.1.1 Μοντέλο Clusternet

Το μοντέλο Clusternet συνδυάζει την ενισχυτική μάθηση με ένα βαθύ νευρωνικό δίκτυο για την εκπαίδευση του μοντέλου. Για αυτό το λόγο η βέλτιστη λύση προ-

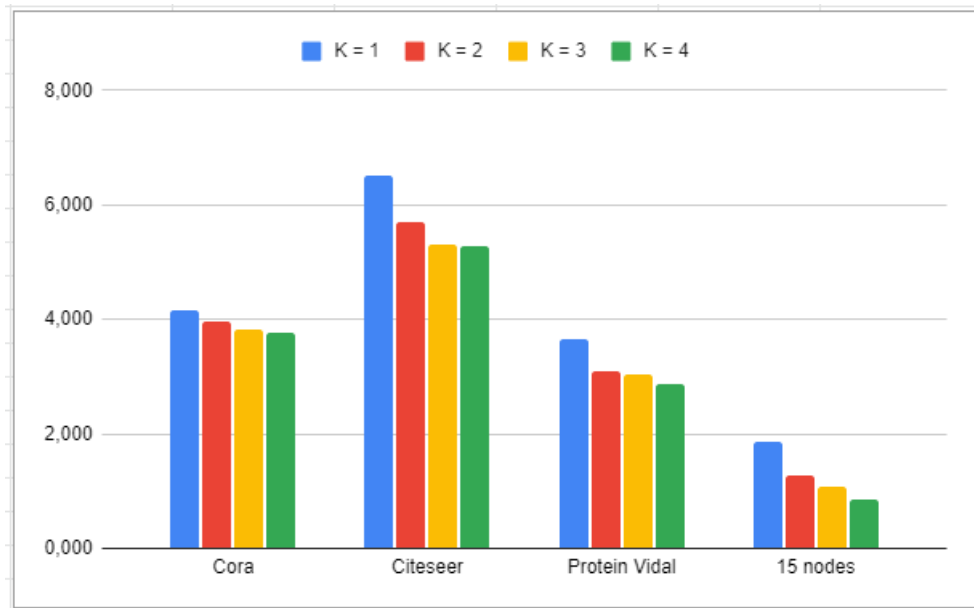
κύπτει από το εκπαιδευμένο μοντέλο που παράγει τη μικρότερη μέση απόσταση. Η εκπαίδευση του μοντέλου πραγματοποιήθηκε για 100, 200, ..., 2000 εποχές. Το μοντέλο έδωσε λύση στο πρόβλημα εγκατάστασης μιας έως τεσσάρων μονάδων. Ο ρυθμός μάθησης ορίστηκε για όλα τα προβλήματα στο 0.01 ενώ ο αριθμός των κριμένων νευρώνων πήρε την τιμή 5.

Στο Σχήμα 5.1 παρουσιάζονται οι γραφικές παραστάσεις των συνόλων δεδομένων Cora, Citeseer, Protein Vidal και ενός γράφου με 15 κόμβους. Για τον γράφο με τους 15 κόμβους παρατηρείται ότι για μία μονάδα, η μέση απόσταση μεταξύ των μονάδων και των σημείων ζήτησης ισούται με 1.866. Η προσθήκη μιας επιπλέον μονάδας μειώνει τη μέση απόσταση κατά 0.4. Η ίδια ελάττωση προκύπτει και για την προσθήκη της τρίτης μονάδας. Τέλος, η εγκατάσταση τεσσάρων μονάδων δίνει μέση απόσταση 0.8. Για το σύνολο δεδομένων Cora, η μέση απόσταση για την εγκατάσταση μιας μονάδας είναι 4.17. Για την εγκατάσταση δύο μονάδων, η μέση απόσταση που επιστρέφει το μοντέλο είναι 3.97. Στην περίπτωση των τριών και τεσσάρων μονάδων πραγματοποιείται μείωση κατά 0.1 ανά μονάδα. Για το σύνολο δεδομένων Citeseer η μέση απόσταση για την εγκατάσταση μιας μονάδας είναι 6.5. Η προσθήκη μιας δεύτερης μονάδας μειώνει τη μέση απόσταση κατά 0.8. Τέλος, για τρεις μονάδες, η μέση απόσταση ισούται με 5.3 ενώ για τέσσερις 5.28. Για το σύνολο δεδομένων Protein Vidal, μέση απόσταση για μια μονάδα ισούται με 3.6. Για την εγκατάσταση δύο μονάδων παρατηρείται ελάττωση της μέσης απόστασης κατά 0.5. Τέλος, η εγκατάσταση τριών και τεσσάρων μονάδων προσθέτει επιπλέον μείωση κατά 0.1.

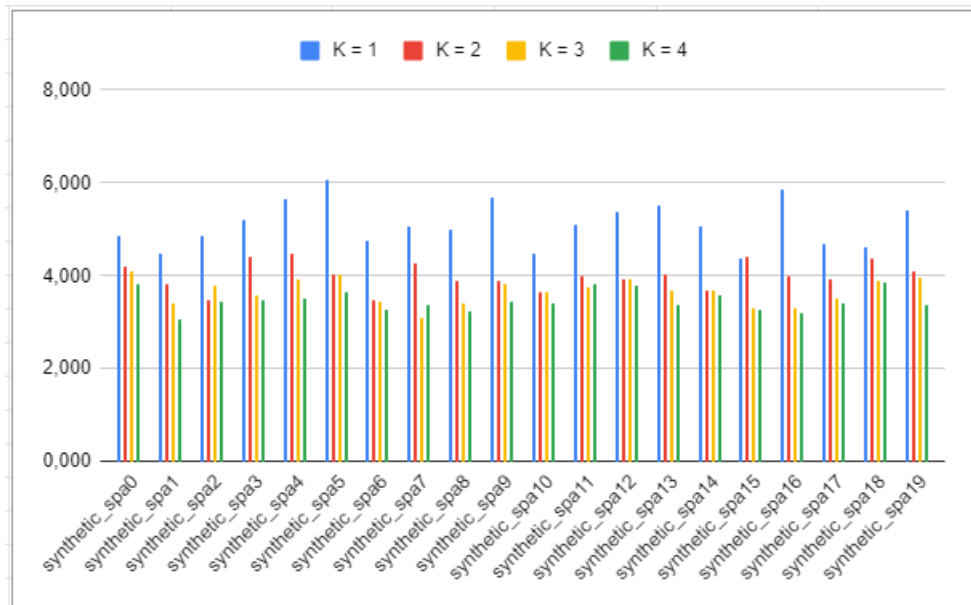
Στο Σχήμα 5.2 παρουσιάζονται γράφοι με αριθμό κόμβων ίσο με 500 για το σύνολο δεδομένων Synthetic SPA. Στην πλειοψηφία των γράφων παρατηρείται ότι η αύξηση του αριθμού των εγκατεστημένων μονάδων ελαττώνει τη μέση απόσταση. Στον Πίνακα 5.1 δίνεται η μέση απόσταση, ο χρόνος εκτέλεσης και ο χρόνος εκπαίδευσης του μοντέλου Clusternet για κάθε ένα από τα σύνολα δεδομένων.

5.1.2 Μοντέλο EMD-MRP

Το μοντέλο EMD-MRP συνδυάζει την ενισχυτική μάθηση με τον ευρετικό αλγόριθμο Lagrangian Relaxation (Ενότητα 2.4) και το μεθευρετικό αλγόριθμο Particle Swarm Optimization (Ενότητα 2.4). Για αυτόν τον λόγο, στα σχήματα που ακολου-



Σχήμα 5.1: Μέση απόσταση μοντέλου Clusternet για τα σύνολα δεδομένων Cora, Citeseer, Protein Vidal, 15 nodes



Σχήμα 5.2: Μέση απόσταση μοντέλου Clusternet για τους γράφους του συνόλου δεδομένων Synthetic SPA

θούν παρατηρείται εύρεση της βέλτιστης λύσης σε μικρό αριθμό εποχών. Η εκπαίδευση του μοντέλου πραγματοποιήθηκε για 1 έως 5 εποχές. Το μοντέλο έδωσε λύση στο πρόβλημα εγκατάστασης μιας έως τεσσάρων μονάδων. Ο ρυθμός μάθησης ορίστηκε για όλα τα προβλήματα στο 0.01 ενώ ο αριθμός των κρυμμένων νευρώνων πήρε την τιμή 5.

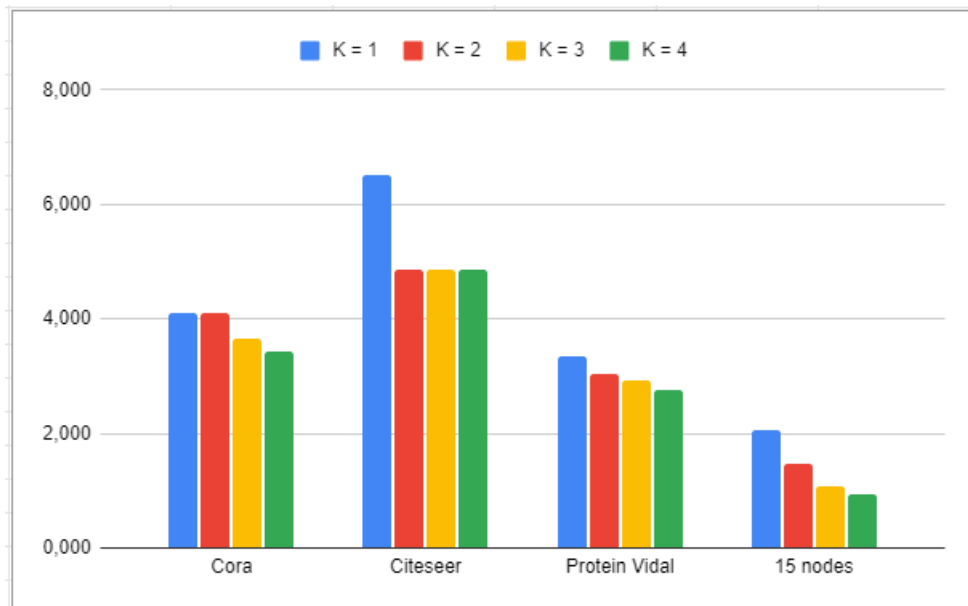
Στο Σχήμα 5.3 παρουσιάζονται οι γραφικές παραστάσεις των συνόλων δεδομένων Cora, Citeseer, Protein Vidal και ενός γράφου με 15 κόμβους. Για το γράφο με τους 15 κόμβων, η μέση απόσταση ισούται με 2.0 ενώ για δύο μονάδες με 1.5. Για το σύνολο δεδομένων Cora, η μέση απόσταση παραμένει σταθερή για την τοποθέτηση μιας και δύο μονάδων. Για την εγκατάσταση τριών μονάδων παρατηρείτε μείωση της μέσης απόστασης κατά 0.5. Τέλος, την εγκατάσταση τεσσάρων μονάδων παρατηρείται επιπλέον μείωση. Για το σύνολο δεδομένων Citeseer, η μέση απόσταση που προκύπτει από την εγκατάσταση μιας μονάδας είναι 6.5. Η τοποθέτηση περισσότερων μονάδων παράγει λύση που κυμαίνεται γύρω από την τιμή 4.8 με απόκλιση 0.01. Για το σύνολο δεδομένων Protein Vidal, η μέση απόσταση για την τοποθέτηση μιας μονάδας είναι 3.4. Για την προσθήκη μιας επιπλέον μονάδας η μέση απόσταση μείνεται κατά 0.4. Τέλος, για την εγκατάσταση τριών και τεσσάρων μονάδων παρατηρείτε επιπλέον μείωση 0.1 για κάθε μονάδα. Στα Σχήματα 5.4 παρουσιάζεται γράφοι με αριθμό κόμβων ίσο με 500 για το σύνολο δεδομένων Synthetic SPA. Στην πλειοψηφία των γράφων παρατηρείται ότι η αύξηση του αριθμού των εγκατεστημένων μονάδων ελαττώνει τη μέση απόσταση. Στον Πίνακα 5.1 δίνεται η μέση απόσταση, ο χρόνος εκτέλεσης και ο χρόνος εκπαίδευσης του μοντέλου EMD-MRP για κάθε ένα από τα σύνολα δεδομένων.

5.2 Σύγκριση Μοντέλων Ενισχυτικής Μάθησης

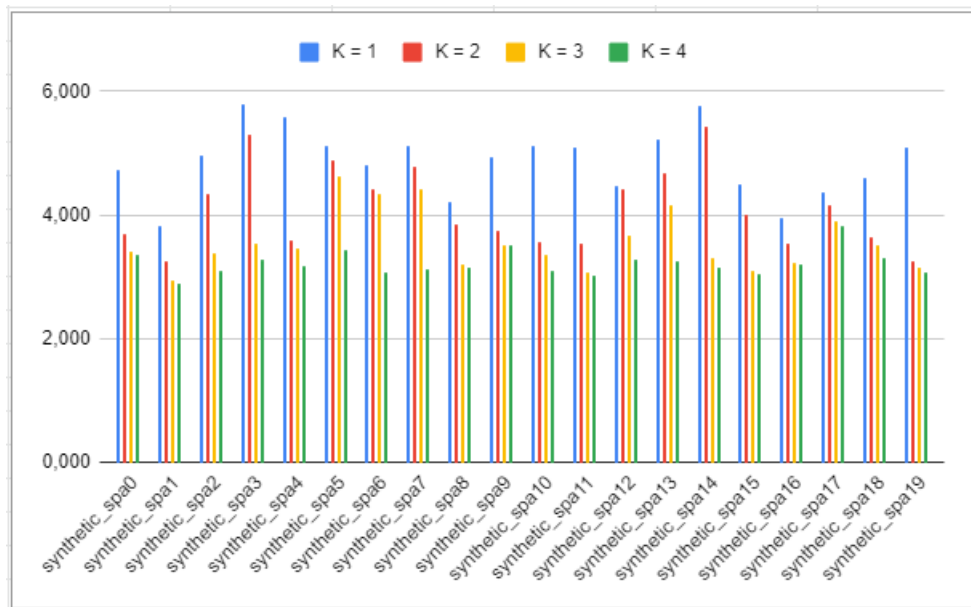
Σε αυτήν την ενότητα θα πραγματοποιηθεί σύγκριση των μοντέλων Clusternet και EMD-MDP ως προς το χρόνο εκπαίδευσης και τη μέση απόσταση μεταξύ των μονάδων και των σημείων ζήτησης.

5.2.1 Εκπαίδευση

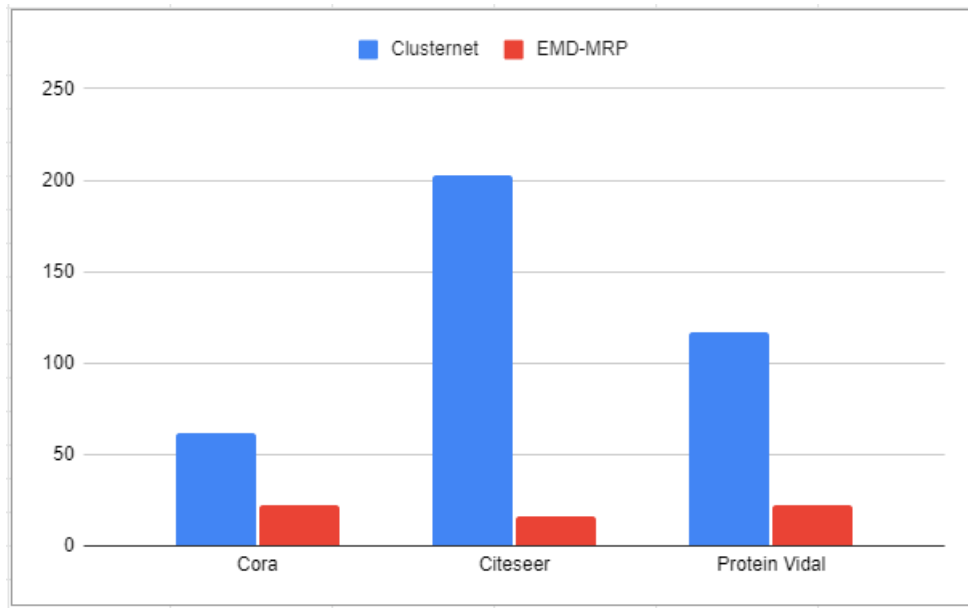
Για την επίλυση του προβλήματος χωροθέτησης μιας μονάδας παρατηρείται από τα Σχήματα 5.5, 5.6 ότι το EMD-MDP χρειάζεται λιγότερο χρόνο εκπαίδευσης στα



Σχήμα 5.3: Μέση απόσταση μοντέλου EMD-MRP για τα σύνολα δεδομένων Cora, Citeseer, Protein Vidal, 15 nodes



Σχήμα 5.4: Μέση απόσταση μοντέλου EMD-MRP για τους γράφους του συνόλου δεδομένων Synthetic SPA

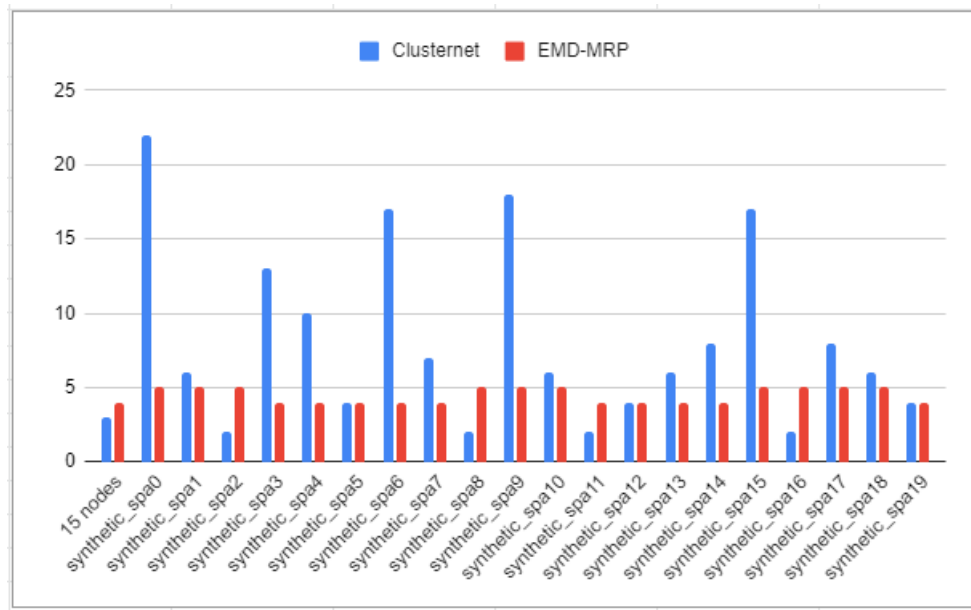


Σχήμα 5.5: Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στα σύνολα δεδομένων Cora, Citeseer, Protein Vidal για μια μονάδα

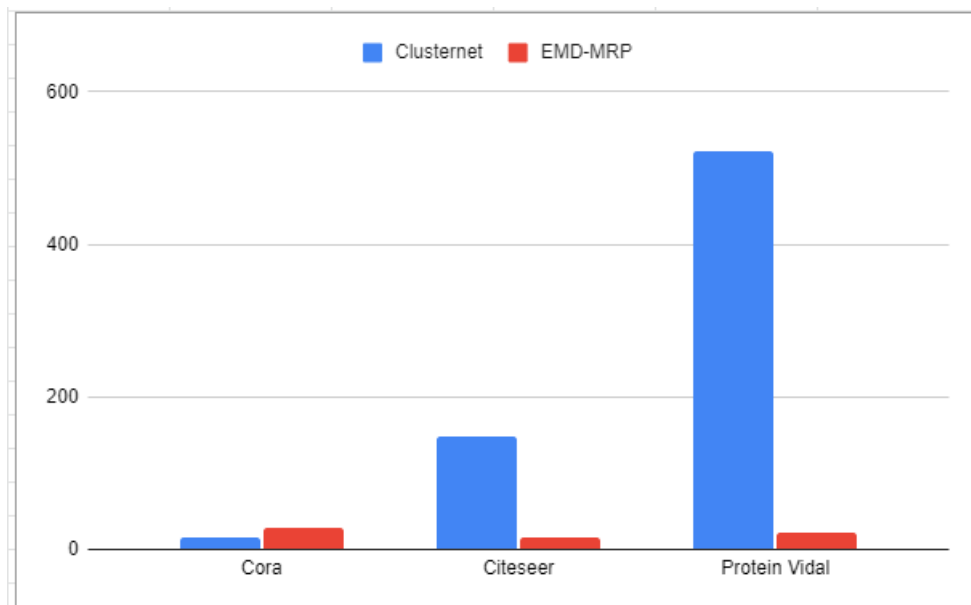
17 από τα 24 σύνολα δεδομένων. Για την τοποθέτηση δύο μονάδων παρατηρείται από τα Σχήματα 5.7, 5.8 ότι το EMD-MDP χρειάζεται λιγότερο χρόνο εκπαίδευσης στα 23 από τα 24. Για την τοποθέτηση τριών και τεσσάρων μονάδων (Σχήματα 5.9, 5.10 και 5.11, 5.12), το EMD-MDP χρειάζεται λιγότερο χρόνο εκπαίδευσης στην πλειοψηφία των συνόλων δεδομένων. Επιπλέον, από τους παραπάνω πίνακες συμπεραίνουμε ότι ο χρόνος εκπαίδευσης του μοντέλου Clusternet αυξάνεται ραγδαία με την αύξηση του πλήθους των κόμβων και των ακμών του γράφου.

5.2.2 Αξιολόγηση

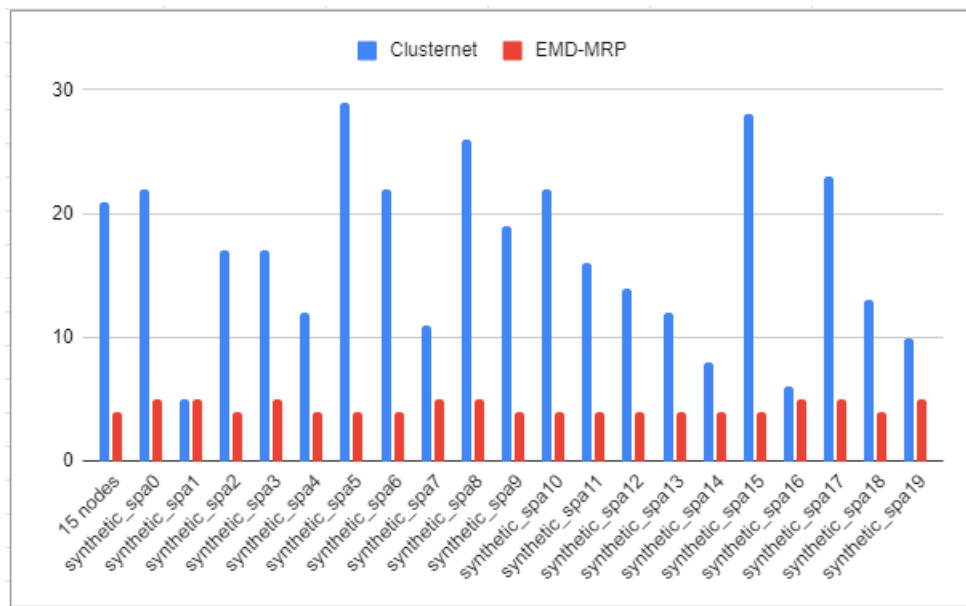
Στο Σχήμα 5.13 παρουσιάζεται η λύση των μοντέλων ανά σύνολο δεδομένων. Από αυτό, παρατηρείται ότι το μοντέλο EMD-MDP δίνει καλύτερη μέση απόσταση στα 16 από τα 24 σύνολα δεδομένων. Για την εγκατάσταση δύο μονάδων παρατηρούμε από το Σχήμα 5.14 ότι το μοντέλο EMD-MDP δίνει καλύτερη μέση απόσταση στα 13 από τα 24 σύνολα δεδομένων. Για την εγκατάσταση τριών μονάδων, παρατηρούμε από το Σχήμα 5.15 ότι το EMD-MDP δίνει καλύτερη μέση απόσταση στα 19 από τα 24 σύνολα δεδομένων. Τέλος, για την εγκατάσταση τεσσάρων μονάδων παρατηρούμε από το Σχήμα 5.16 ότι το μοντέλο EMD-MDP δίνει καλύτερη μέση απόσταση στα 21 από τα 24 σύνολα δεδομένων. Από τα παραπάνω συμπεραίνουμε ότι το μοντέλο EMD-MDP βρίσκει τη μικρότερη μέση απόσταση τόσο στο πρό-



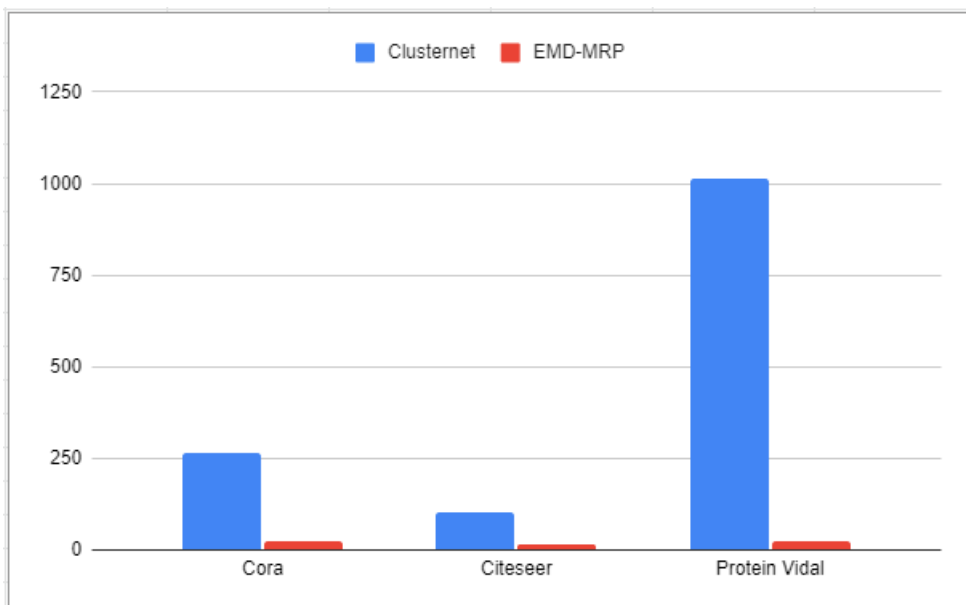
Σχήμα 5.6: Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στους γράφους του συνόλου δεδομένων Synthetic SPA, 15 nodes για μια μονάδα



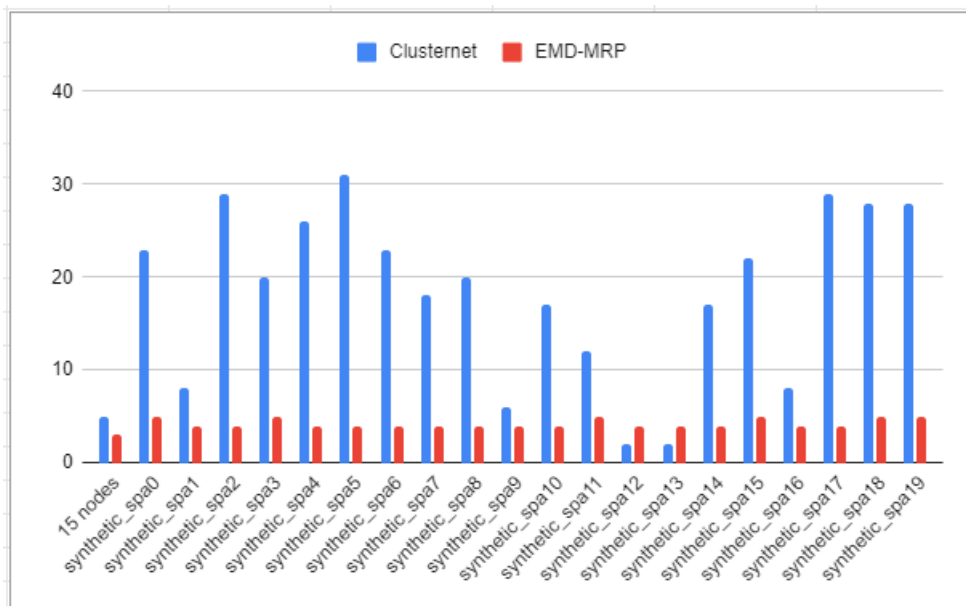
Σχήμα 5.7: Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στα σύνολα δεδομένων Cora, Citeseer, Protein Vidal για δύο μονάδες



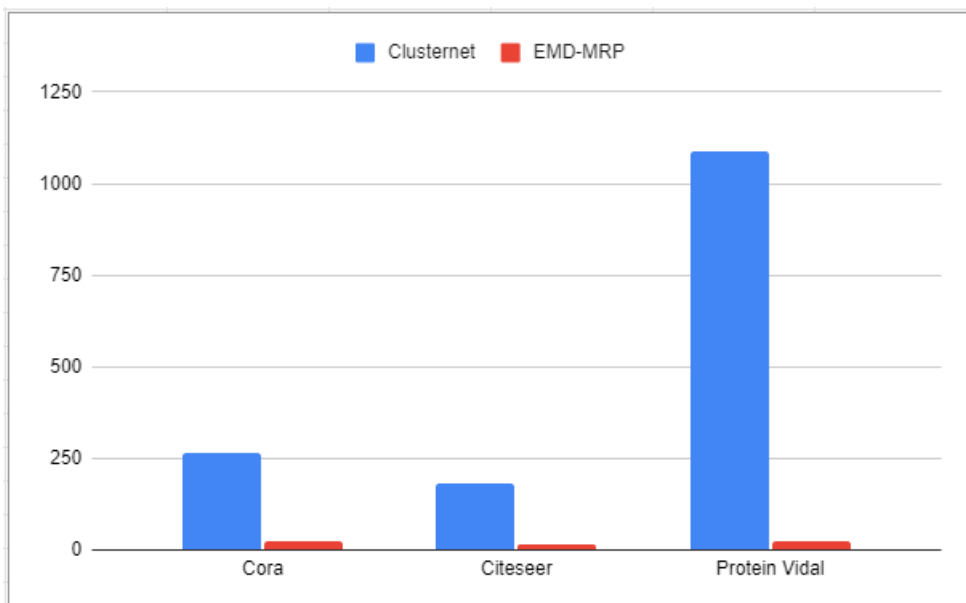
Σχήμα 5.8: Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στους γράφους του συνόλου δεδομένων Synthetic SPA, 15 nodes για δύο μονάδες



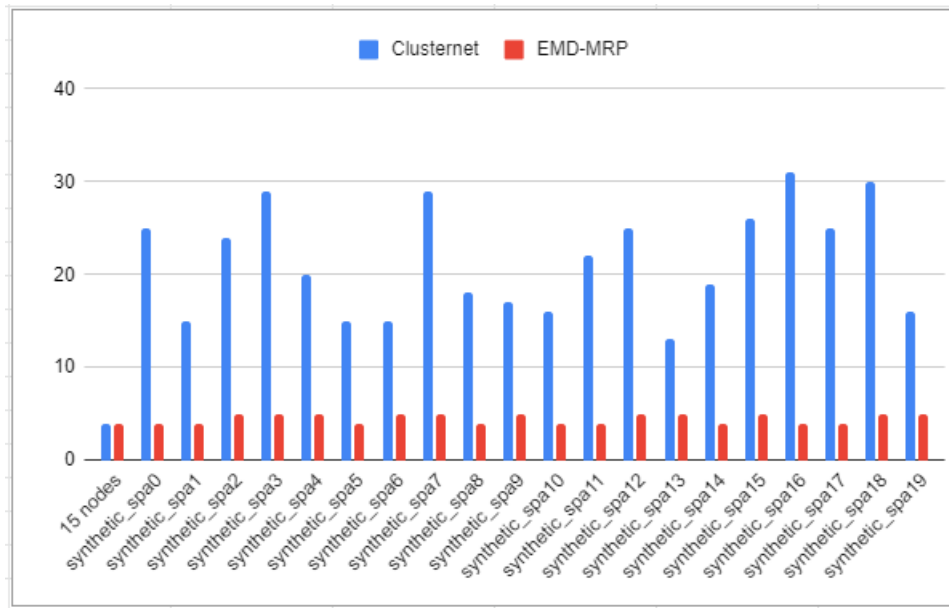
Σχήμα 5.9: Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στα σύνολα δεδομένων Cora, Citeseer, Protein Vidal για τρεις μονάδες



Σχήμα 5.10: Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στους γράφους του συνόλου δεδομένων Synthetic SPA, 15 nodes για τρεις μονάδες



Σχήμα 5.11: Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στα σύνολα δεδομένων Cora, Citeseer, Protein Vidal για τέσσερις μονάδες

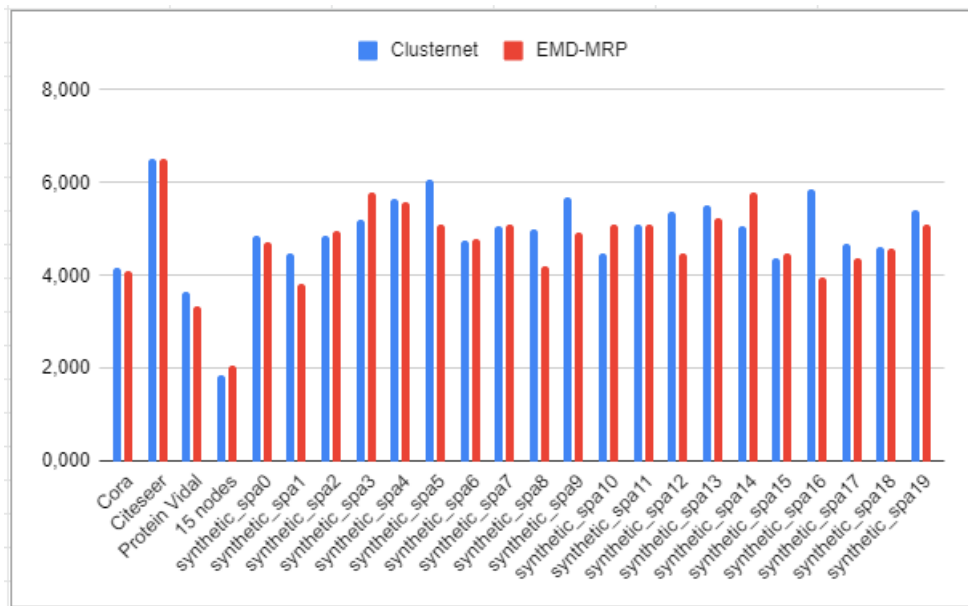


Σχήμα 5.12: Χρόνος εκπαίδευσης μοντέλου Clusternet και EMD-MDP στους γράφους του συνόλου δεδομένων Synthetic SPA, 15 nodes για τέσσερις μονάδες

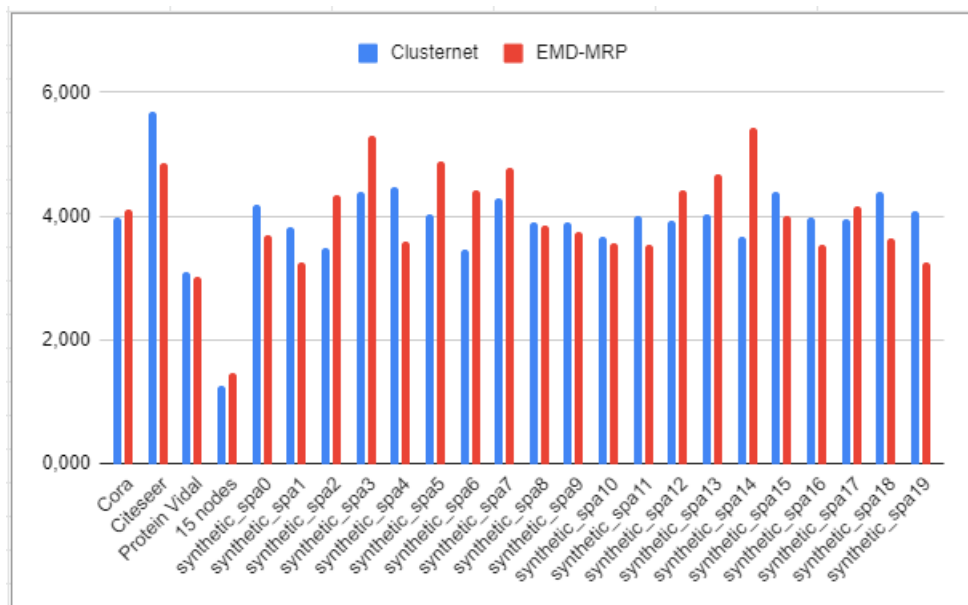
βλημα χωροθέτησης μιας μονάδας όσο και στο πρόβλημα χωροθέτησης πολλαπλών μονάδων.

5.2.3 Σύνοψη

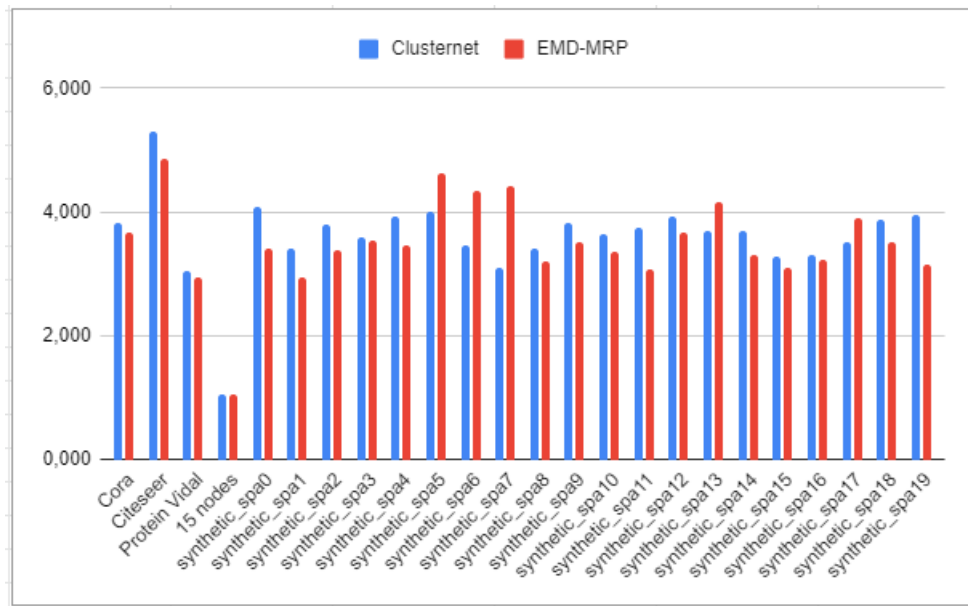
Στις προηγούμενες ενότητες του κεφαλαίου πραγματοποιήθηκε ανάλυση των μοντέλων Clusternet και EMD-MDP στον χρόνο εκπαίδευσης και στα αποτελέσματα του κάθε μοντέλου. Από τα παραπάνω συμπεραίνουμε ότι το EMD-MDP υπερτερεί του Clusternet τόσο στον χρόνο εκπαίδευσης όσο και στην ποιότητα των αποτελεσμάτων. Αυτό συμβαίνει επειδή το EMD-MDP χρησιμοποιεί εκτός από το μηχανισμό ενισχυτικής μάθησης και τη χαλάρωση Lagrange και τον μεθευρετικό αλγόριθμο Particle Swarm Optimization.



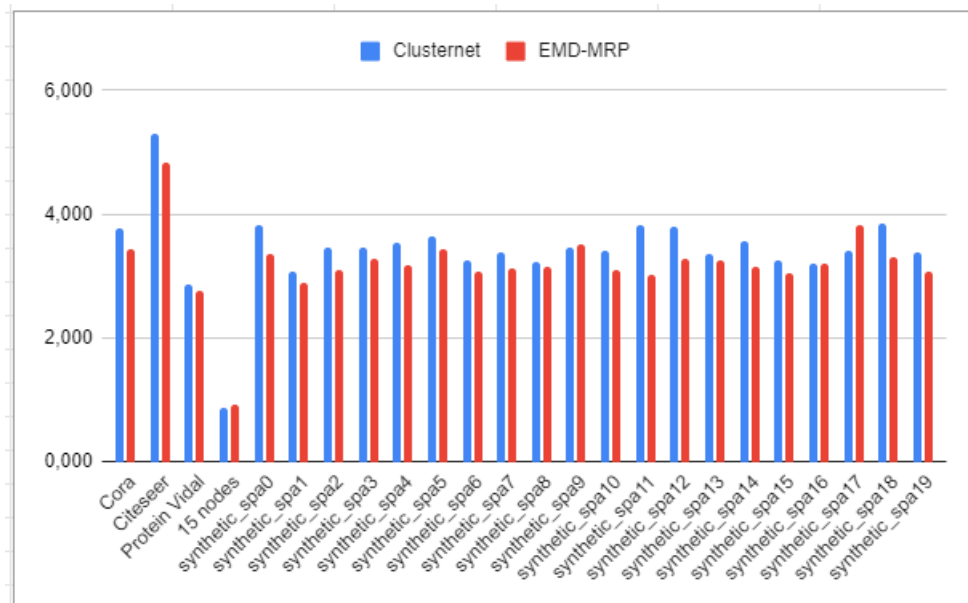
Σχήμα 5.13: Μέση απόσταση μοντέλου Clusternet και EMD-MDP ανά σύνολο δεδομένων για μια μονάδα



Σχήμα 5.14: Μέση απόσταση μοντέλου Clusternet και EMD-MDP ανά σύνολο δεδομένων για δύο μονάδες



Σχήμα 5.15: Μέση απόσταση μοντέλου ClusterNet και EMD-MDP ανά σύνολο δεδομένων για τρεις μονάδες



Σχήμα 5.16: Μέση απόσταση μοντέλου ClusterNet και EMD-MDP ανά σύνολο δεδομένων για τέσσερις μονάδες

Πίνακας 5.1: Πίνακας αποτελεσμάτων μέσης απόστασης, χρόνων εκτέλεσης και εκπαίδευσης των μοντέλων ανά σύνολο δεδομένων

		Clusternet			EMD-MRP		
		Mean Distance	Run Time	Training Time	Mean Distance	Run Time	Training Time
Cora	K = 1	4.172	85	62	4.117	47	22
	K = 2	3.972	42	16	4.117	53	28
	K = 3	3.836	321	293	3.662	63	22
	K = 4	3.771	294	263	3.428	84	24
Citeseer	K = 1	6.500	244	203	6.500	37	16
	K = 2	5.700	198	148	4.860	45	15
	K = 3	5.304	152	103	4.850	53	14
	K = 4	5.289	236	183	4.846	69	14
Protein Vidal	K = 1	3.646	129	117	3.346	50	22
	K = 2	3.090	542	522	3.029	59	22
	K = 3	3.051	1,036	1,013	2.934	85	24
	K = 4	2.878	1,116	1,088	2.772	111	24
15 nodes	K = 1	1.866	8	3	2.067	8	4
	K = 2	1.266	24	21	1.467	8	4
	K = 3	1.066	7	5	1.067	8	3
	K = 4	0.866	6	4	0.933	8	4
synthetic_spa0	K = 1	4.852	24	22	4.724	10	5
	K = 2	4.188	24	22	3.692	11	5
	K = 3	4.096	26	23	3.404	11	5
	K = 4	3.836	28	25	3.366	11	4
synthetic_spa1	K = 1	4.470	8	6	3.824	10	5
	K = 2	3.828	9	5	3.262	10	5
	K = 3	3.410	11	8	2.954	11	4
	K = 4	3.070	18	15	2.882	11	4
synthetic_spa2	K = 1	4.872	6	2	4.952	11	5
	K = 2	3.486	20	17	4.334	10	4
	K = 3	3.802	32	29	3.390	10	4
	K = 4	3.456	28	24	3.112	11	5
synthetic_spa3	K = 1	5.194	16	13	5.802	9	4
	K = 2	4.406	20	17	5.294	10	5
	K = 3	3.590	23	20	3.530	10	5
	K = 4	3.474	32	29	3.276	11	5
synthetic_spa4	K = 1	5.648	12	10	5.586	9	4
	K = 2	4.484	14	12	3.598	9	4
	K = 3	3.920	29	26	3.456	9	4
	K = 4	3.528	24	20	3.184	11	5
synthetic_spa5	K = 1	6.072	6	4	5.110	9	4
	K = 2	4.032	32	29	4.886	9	4
	K = 3	4.014	34	31	4.636	9	4
	K = 4	3.644	18	15	3.438	10	4
synthetic_spa6	K = 1	4.756	26	17	4.806	10	4
	K = 2	3.464	24	22	4.428	9	4
	K = 3	3.462	26	23	4.338	10	4
	K = 4	3.260	18	15	3.072	11	5
synthetic_spa7	K = 1	5.060	11	7	5.106	9	4
	K = 2	4.282	14	11	4.780	10	5
	K = 3	3.090	22	18	4.416	10	4
	K = 4	3.388	32	29	3.116	10	5
synthetic_spa8	K = 1	5.008	13	2	4.210	9	5
	K = 2	3.908	29	26	3.846	10	5
	K = 3	3.420	23	20	3.216	10	4
	K = 4	3.236	22	18	3.144	11	4

		Clusternet			EMD-MRP		
		Mean Distance	Run Time	Training Time	Mean Distance	Run Time	Training Time
synthetic_spa9	K = 1	5.680	21	18	4.944	10	5
	K = 2	3.908	21	19	3.740	10	4
	K = 3	3.818	9	6	3.516	10	4
	K = 4	3.458	20	17	3.514	11	5
synthetic_spa10	K = 1	4.488	9	6	5.108	10	5
	K = 2	3.666	24	22	3.566	9	4
	K = 3	3.656	20	17	3.370	10	4
	K = 4	3.406	20	16	3.108	10	4
synthetic_spa11	K = 1	5.086	11	2	5.086	9	4
	K = 2	4.012	18	16	3.540	9	4
	K = 3	3.748	14	12	3.070	10	5
	K = 4	3.818	24	22	3.026	10	4
synthetic_spa12	K = 1	5.392	7	4	4.472	9	4
	K = 2	3.938	16	14	4.418	9	4
	K = 3	3.922	5	2	3.658	10	4
	K = 4	3.792	27	25	3.284	10	5
synthetic_spa13	K = 1	5.508	8	6	5.222	9	4
	K = 2	4.044	14	12	4.680	9	4
	K = 3	3.698	5	2	4.168	10	4
	K = 4	3.368	16	13	3.252	10	5
synthetic_spa14	K = 1	5.082	10	8	5.776	9	4
	K = 2	3.674	10	8	5.426	9	4
	K = 3	3.688	20	17	3.302	10	4
	K = 4	3.576	22	19	3.144	11	4
synthetic_spa15	K = 1	4.364	26	17	4.490	10	5
	K = 2	4.398	30	28	3.998	10	4
	K = 3	3.290	24	22	3.088	11	5
	K = 4	3.256	29	26	3.044	11	5
synthetic_spa16	K = 1	5.860	11	2	3.966	10	5
	K = 2	3.986	9	6	3.534	10	5
	K = 3	3.300	10	8	3.218	10	4
	K = 4	3.196	34	31	3.216	11	4
synthetic_spa17	K = 1	4.674	10	8	4.362	10	5
	K = 2	3.944	26	23	4.152	10	5
	K = 3	3.524	32	29	3.900	10	4
	K = 4	3.416	27	25	3.826	10	4
synthetic_spa18	K = 1	4.624	8	6	4.594	10	5
	K = 2	4.384	16	13	3.634	9	4
	K = 3	3.878	31	28	3.504	10	5
	K = 4	3.858	33	30	3.312	11	5
synthetic_spa19	K = 1	5.422	6	4	5.090	9	4
	K = 2	4.084	12	10	3.268	10	5
	K = 3	3.954	31	28	3.140	9	5
	K = 4	3.390	20	16	3.084	11	5

Κεφάλαιο 6

Συμπεράσματα

Στην παρούσα διπλωματική εργασία πραγματοποιήθηκε η παρουσίαση και ανάλυση μεθόδων επίλυσης του προβλήματος χωροθέτησης μονάδων με τη χρήση μηχανισμών ενισχυτικής μάθησης. Αρχικά, παρουσιάστηκαν βασικές έννοιες του προβλήματος χωροθέτησης μονάδων, όπως η συνάρτηση απόστασης, η ζήτηση των πελατών και η αλληλεπίδραση των μονάδων με τους πελάτες. Επίσης, παρουσιάστηκε η μαθηματική μοντελοποίηση επεκτάσεων του προβλήματος καθώς και διάφορες τεχνικές εύρεσης της βέλτιστης λύσης του προβλήματος. Στη συνέχεια, παρουσιάστηκε η μοντελοποίηση των προβλημάτων ενισχυτικής μάθησης για την καλύτερη κατανόηση της μεθόδου. Επιπλέον, παρουσιάστηκαν επιστημονικά άρθρα που ανέλυαν τεχνικές που έχουν προταθεί στη βιβλιογραφία και χρησιμοποιούν την ενισχυτική μάθηση. Ποιο συγκεκριμένα αναλύθηκαν μοντέλα νευρωνικών δικτύων, συστήματα πολλαπλών πρακτόρων και τεχνικές ενίσχυσης των ακμών του γράφου που χρησιμοποιείται για τη μοντελοποίηση του προβλήματος.

Εκτός από τη βιβλιογραφική ανάλυση, στην παρούσα εργασία πραγματοποιήθηκε σύγκριση δύο μοντέλων βάσει των αποτελεσμάτων που προέκυψαν από την επίλυση διαφόρων προβλημάτων. Αρχικά παρουσιάστηκε το μοντέλο Clusternet το οποίο βασίζεται σε συνελικτικά δίκτυα γράφων (Graph Convolutional Networks) τα οποία ανήκουν στην κατηγορία νευρωνικών δικτύων. Το δεύτερο μοντέλο που αναλύθηκε είναι ο μηχανισμός ισότητας με τη χρήση της Μαρκοβιανής διαδικασίας κέρδους (EMD-MRP). Το συγκεκριμένο μοντέλο συνδυάζει την ενισχυτική μάθηση με τον ευρετικό αλγόριθμο της χαλάρωσης Lagrange και το μεθευρετικό αλγόριθμο Particle Swarm Optimization. Η σύγκριση των μοντέλων έγινε ως προς τον χρόνο εκπαίδευσης και την ελαχιστοποίηση της μέσης απόστασης μεταξύ των μονάδων και

των σημείων ζήτησης. Από τα αποτελέσματα παρατηρήθηκε ότι το μοντέλο EMD-MRP υπερτερεί του Clusternet τόσο στον χρόνο εκπαίδευσης όσο και στην ποιότητα των αποτελεσμάτων. Ως προς το χρόνο εκπαίδευσης, τα νευρωνικά δίκτυα απαιτούν σχετικά μεγάλο χρόνο εκπαίδευσης σε σχέση με άλλες τεχνικές. Αυτός είναι άλλωστε και ο λόγος που το μοντέλο Clusternet χρειάζεται περισσότερο χρόνο εκπαίδευσης. Ως προς την ποιότητα των αποτελεσμάτων, το EMD-MRP πετυχαίνει καλύτερη λύση επειδή συνδυάζει εκτός από την ενισχυτική μάθηση, ένα ευρετικό και ένα μεθευρετικό αλγόριθμο.

Από την ανάλυση των δύο μοντέλων συμπεραίνουμε αρχικά ότι τα μοντέλα ενισχυτικής μάθησης καταφέρνουν να βρουν μια λύση που συγκλίνει στη βέλτιστη λύση του προβλήματος. Επιπλέον, ο συνδυασμός της ενισχυτικής μάθησης με γνωστούς ευρετικούς αλγόριθμους φαίνεται να δίνει καλά αποτελέσματα όχι μόνο στην ακρίβεια των αποτελεσμάτων αλλά και στον χρόνο που απαιτείται για την εκπαίδευση των μοντέλων.

Βιβλιογραφία

- [1] Tobias Achterberg and Timo Berthold. Hybrid branching. In Willem-Jan van Hoeve and John N. Hooker, editors, *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, pages 309–311, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.
- [2] Tobias Achterberg and Roland Wunderling. *Mixed integer programming: Analyzing 12 years of progress*, pages 449–481. 01 2013.
- [3] Alexander Ageev and M.I. Sviridenko. Pipage rounding: A new method of constructing algorithms with proven performance guarantee. *Journal of Combinatorial Optimization*, 8:307–328, 01 2004.
- [4] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, Jan 2002.
- [5] Gunnar Andersson, Richard L. Francis, Tomas Normark, and M.Brenda Rayco. Aggregation method experimentation for large-scale network location problems. *Location Science*, 6(1):25–39, 1998.
- [6] Egon Balas and Andrew Ho. Set covering algorithms using cutting planes, heuristics, and subgradient optimization: a computational study. In *Combinatorial Optimization*, pages 37–60. Springer, 1980.
- [7] Marc Barthélemy. Spatial networks. *Physics Reports*, 499(1-3):1–101, Feb 2011.
- [8] J.E. Beasley. Lagrangean heuristics for location problems. *European Journal of Operational Research*, 65(3):383–399, 1993.
- [9] Una Benlic and Jin-Kao Hao. Memetic search for the quadratic assignment problem. *Expert Systems with Applications*, 42:584–595, 01 2015.
- [10] R. Burkard, S. Karisch, and F. Rendl. Qaplib – a quadratic assignment problem library. *Journal of Global Optimization*, 10:391–403, 1997.
- [11] Fan Chung and Linyuan Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6:125–145, 11 2002.
- [12] Jens Clausen. Branch and bound algorithms-principles and examples. *Department of Computer Science, University of Copenhagen*, pages 1–30, 1999.

-
- [13] Aaron Clauset, M. E. J. Newman, and Cristopher Moore. Finding community structure in very large networks. *Physical Review E*, 70(6), Dec 2004.
- [14] Koblenz Network Collection. Adolescent health. http://konect.uni-koblenz.de/networks/moreno_health. [Online; accessed 11-October-2024].
- [15] Koblenz Network Collection. Human protein (vidal). <http://konect.uni-koblenz.de/networks/maayan-vidal>. [Online; accessed 11-October-2024].
- [16] Edoardo Conti, Vashisht Madhavan, Felipe Petroski Such, Joel Lehman, Kenneth O Stanley, and Jeff Clune. Improving exploration in evolution strategies for deep reinforcement learning via a population of novelty-seeking agents. *arXiv preprint arXiv:1712.06560*, 2017.
- [17] Gérard Cornuéjols, George Nemhauser, and Laurence Wolsey. The uncapacitated facility location problem. Technical report, Cornell University Operations Research and Industrial Engineering, 1983.
- [18] Pierluigi Crescenzi, Gianlorenzo D’angelo, Lorenzo Severini, and Yllka Velaj. Greedily improving our own closeness centrality in a network. *ACM Transactions on Knowledge Discovery from Data*, 11:1–32, 07 2016.
- [19] Marek Cygan and Pinar Heggernes. *Parameterized and exact computation: 9th International Symposium, IPEC 2014, Wroclaw, Poland, September 10-12, 2014. Revised selected papers*, volume 8894. Springer, 2014.
- [20] M. Dorigo. *Optimization, Learning and Natural Algorithms*. PhD thesis, Politecnico di Milano, 1992.
- [21] Marco Dorigo, Vittorio Maniezzo, and Alberto Colorni. Ant system: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1):29–41, 1996.
- [22] Zvi Drezner. *Facility location: a survey of applications and methods*. New York, Springer, 1995.
- [23] Zvi Drezner. Extensive experiments with hybrid genetic algorithms for the solution of the quadratic assignment problem. *Computers & Operations Research*, 35(3):717–736, 2008.
- [24] Rick Durrett. *Erdős–Rényi random graphs*, page 27–69. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2006.
- [25] Benjamin Fish, Ashkan Bashardoust, Danah Boyd, Sorelle Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. Gaps in information access in social networks? *The World Wide Web Conference*, May 2019.
- [26] Marshall L. Fisher. The lagrangian relaxation method for solving integer programming problems. *Management Science*, 50(12):1861–1871, 2004.
- [27] Gerald Gamrath, Daniel Anderson, Ksenia Bestuzheva, Wei-Kun Chen, Leon Eifler, Maxime Gasse, Patrick Gemander, Ambros Gleixner, Leona Gottwald, Katrin Halbig, Gregor Hendel, Christopher Hojny, Thorsten Koch, Pierre Le Bodic, Stephen J. Maher, Frederic Matter, Matthias Miltenberger,

-
- Erik Mühmer, Benjamin Müller, Marc E. Pfetsch, Franziska Schlösser, Felipe Serrano, Yuji Shinano, Christine Tawfik, Stefan Vigerske, Fabian Wegscheider, Dieter Wening, and Jakob Witzig. *The SCIP Optimization Suite 7.0*. ZIB-Report. Zuse Institut Berlin, 2020.
- [28] Maxime Gasse, Didier Chetelat, Nicola Ferroni, Laurent Charlin, and Andrea Lodi. Exact combinatorial optimization with graph convolutional neural networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [29] Teofilo F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306, 1985.
- [30] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864, 2016.
- [31] Serigne Gueye and Mozart BC Menezes. General asymptotic and submodular results for the median problem with unreliable facilities. *Operations Research Letters*, 43(5):519–521, 2015.
- [32] Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137, 1983.
- [33] Peter J. Huber. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1):73 – 101, 1964.
- [34] T. James, C. Rego, and F. Glover. A cooperative parallel tabu search algorithm for the quadratic assignment problem. *Eur. J. Oper. Res.*, 195:810–826, 2009.
- [35] Tabitha James, Cesar Rego, and Fred Glover. A cooperative parallel tabu search algorithm for the quadratic assignment problem. *European Journal of Operational Research*, 195(3):810–826, 2009.
- [36] Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*, 2016.
- [37] George Karypis and Vipin Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM Journal on scientific Computing*, 20(1):359–392, 1998.
- [38] Elias Khalil, Pierre Le Bodic, Le Song, George Nemhauser, and Bistra Dilkina. Learning to branch in mixed integer programming. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30, 2016.
- [39] Elias B. Khalil, Pierre Le Bodic, Le Song, George Nemhauser, and Bistra Dilkina. Learning to branch in mixed integer programming. In Dale Schuurmans and Michael Wellman, editors, *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pages 724–731, United States of America, 2016. Association for the Advancement of Artificial Intelligence (AAAI). AAAI Conference on Artificial Intelligence 2016, AAAI 2016 ; Conference date: 12-02-2016 Through 17-02-2016.

-
- [40] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
- [41] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 420–429, 2007.
- [42] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '07*, page 420–429, New York, NY, USA, 2007. Association for Computing Machinery.
- [43] Long-Ji Lin. *Reinforcement learning for robots Using neural networks*. PhD thesis, USA, 1992. UMI Order No. GAX93-22750.
- [44] Ramon Augusto Sousa Lins, Adrião Duarte Neto Dória, and Jorge Dantas de Melo. Deep reinforcement learning applied to the k-server problem. *Expert Systems with Applications*, 135:212–218, 2019.
- [45] David MacKay. *Information theory, inference, and learning algorithms*, volume 50. 01 2003.
- [46] P. Merz and B. Freisleben. Genetic local search for the tsp: new results. *Proceedings of 1997 IEEE International Conference on Evolutionary Computation (ICEC '97)*, pages 159–164, 1997.
- [47] Peter Merz and Bernd Freisleben. A genetic local search approach to the quadratic assignment problem. In *in Proceedings of the 7th International Conference on Genetic Algorithms*, pages 465–472. Morgan Kaufmann, 1997.
- [48] Victor V Miagkikh and William F Punch III. An approach to solving combinatorial optimization problems using a population of reinforcement learning agents. In *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation-Volume 2*, pages 1358–1365, 1999.
- [49] Alfonsas Misevicius. An improved hybrid genetic algorithm: new results for the quadratic assignment problem. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 3–16. Springer, 2003.
- [50] Alfonsas Misevicius, Antanas Lenkevicius, and Dalius Rubliauskas. Iterated tabu search: An improvement to standard tabu search. *Information Technology And Control*, 35, 09 2006.
- [51] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.
- [52] M. E. J. Newman. Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, 74(3), Sep 2006.
- [53] M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, May 2006.
- [54] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

-
- [55] Włodzimierz Ogryczak. Inequality measures and equitable approaches to location problems. *European Journal of Operational Research*, 122(2):374–391, 2000.
- [56] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.
- [57] Govardana Sachithanandam Ramachandran, Ivan Brugere, Lav R Varshney, and Caiming Xiong. Gaea: Graph augmentation for equitable access via reinforcement learning. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pages 884–894, 2021.
- [58] Gerhard Reinelt. TspLib—a traveling salesman problem library. *INFORMS Journal on Computing*, 3(4):376–384, 1991.
- [59] Martin Riedmiller. Neural fitted q iteration – first experiences with a data efficient neural reinforcement learning method. In João Gama, Rui Camacho, Pavel B. Brazdil, Alípio Mário Jorge, and Luís Torgo, editors, *Machine Learning: ECML 2005*, pages 317–328, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.
- [60] Stefan Ropke and David Pisinger. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation Science*, 40:455–472, 11 2006.
- [61] Budi Santosa and I. Kresna. Simulated annealing to solve single stage capacitated warehouse location problem. *Procedia Manufacturing*, 4:62–70, 12 2015.
- [62] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *European semantic web conference*, pages 593–607. Springer, 2018.
- [63] Prithviraj Sen, Galileo Namata, Mustafa Bilgic, Lise Getoor, Brian Galligher, and Tina Eliassi-Rad. Collective classification in network data. *AI Magazine*, 29(3):93, Sep. 2008.
- [64] Ines Sghir, Jin-Kao Hao, Inès Jaafar, and Khaled Ghédira. A multi-agent based optimization method applied to the quadratic assignment problem. *Expert Systems with Applications*, 42:9252–9262, 12 2015.
- [65] Thomas Stützle. Iterated local search for the quadratic assignment problem. *European Journal of Operational Research*, 174(3):1519–1539, 2006.
- [66] É. Taillard. Robust taboo search for the quadratic assignment problem. *Parallel Computing*, 17:443–455, 1991.
- [67] Gerald Tesauro. Temporal difference learning and td-gammon. *Communications of the ACM*, 38(3):58–68, March 1995.
- [68] Amanda L. Traud, Peter J. Mucha, and Mason A. Porter. Social structure of facebook networks. *Physica A: Statistical Mechanics and its Applications*, 391(16):4165–4180, Aug 2012.
- [69] Cédric Villani. *Optimal transport – Old and new*, volume 338, pages xxii+973. 01 2008.

-
- [70] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007.
- [71] Bryan Wilder, Eric Ewing, Bistra Dilkina, and Milind Tambe. End to end learning and optimization on graphs. *Advances in Neural Information Processing Systems*, 32:4672–4683, 2019.
- [72] Bryan Wilder, Han Ching Ou, Kayla de la Haye, and Milind Tambe. Optimizing network structure for preventative health. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS-18)*, 2018.
- [73] Bryan Wilder, Han Ching Ou, Kayla de la Haye, and Milind Tambe. Optimizing network structure for preventative health. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS-18)*, 2018.
- [74] Wuyang Yu and Jijun Liu. Optimization model based on reachability guarantee for emergency facility location and link reinforcement. *Journal of Advanced Transportation*, 2020:1–12, 03 2020.