



ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΠΑΤΡΩΝ  
UNIVERSITY OF PATRAS

Πολυτεχνική Σχολή  
Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Διπλωματική Εργασία

---

# Αποδοτική ανάθεση πόρων σε Δίκτυα 5G με τεχνικές Machine Learning

---

Μιχαήλ Κουρής

ΑΜ: 1041085

Επιβλέπων: Χρήστος Μπούρας, Καθηγητής

Μέλη Επιτροπής Αξιολόγησης

Χρήστος Μπούρας, Καθηγητής

Ιωάννης Γαροφαλάκης, Καθηγητής

Κυριακός Βλάχος, Αναπληρωτής Καθηγητής

Πάτρα, Μάρτιος 2022

© Copyright συγγραφέας Μιχαήλ Κουρής, 2022

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας για εμπορικό σκοπό. Για μη κερδοσκοπικό σκοπό ή εκπαιδευτικής φύσεως επιτρέπεται η ανατύπωση, η αποθήκευση και η διανομή με την προϋπόθεση να γίνεται αναφορά στην πηγή προελεύσεως και τη διατήρηση του παρόντος μηνύματος

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών & Πληροφορικής του Πανεπιστημίου Πατρών, τον Επιβλέποντα ή της επιτροπής που την ενέκρινε, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

### **Υπεύθυνη Δήλωση**

Δηλώνω ότι είμαι συγγραφέας αυτής της πτυχιακής εργασίας, και ότι κάθε βοήθεια είναι πλήρως αναγνωρισμένη, όπως και κάθε πηγή από όπου αντλήθηκε κάποια ιδέα και δεδομένα, και γίνεται αναφορά των παραπάνω στην πτυχιακή εργασία

**Μιχαήλ Κουρής**

It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so.

-Mark Twain

## Ευχαριστίες

Για την εκπόνηση της διπλωματικής μου εργασίας θα ήθελα πρωτίστως να ευχαριστήσω τον επιβλέποντα Καθηγητή μου κ. Χρήστο Μπούρα που με εμπιστεύτηκε να την αναλάβω και την ευκαιρία που αυτή αντιπροσωπεύει. Θα ήθελα επίσης να ευχαριστήσω τον Δρ. Βασίλη Κόκκινο για όλη την καθοδήγηση και υπομονή σε όλη την διάρκεια της εργασίας. Ευχαριστώ επίσης τον κ. Φοίβο Αλλαγιώτη και την κ. Ασπασία Κουκουβέλα για την συνεργασία μας και για τις επεξηγήσεις που μου παρείχαν σχετικά με την ερευνητική δουλειά τους, πάνω στην οποία στηρίχτηκε η παρούσα εργασία. Τέλος ευχαριστώ τον κ. Σωτήρη Παπαθεοδώρου για την διαρκή του στήριξη.

## Περίληψη

Τα τελευταία χρόνια οι ανθρώπινες τηλεπικοινωνιακές ανάγκες αλλά και τεχνικές δυνατότητες έχουν αυξηθεί ραγδαία. Πουθενά δεν είναι αυτό πιο προφανές από τον τομέα των ασύρματων επικοινωνιών. Τα ασύρματα δίκτυα πλέον είναι υπεύθυνα για μεγαλύτερη διακίνηση δεδομένων από τους σταθερούς υπολογιστές και ως εκ τούτου γίνονται προσπάθειες για την βελτιστοποίηση των πρωτοκόλλων και τεχνικών που διέπουν και πραγματοποιούν αυτά τα ασύρματα δίκτυα. Οι ανάγκες ασύρματου ίντερνετ είναι τόσο μεγάλες, και η χρήση του 4G και του 5G τόσο διαδεδομένη, που προκύπτουν προκλήσεις στην διάθεση και κατανομή των πόρων. Ένας άλλος ραγδαία εξελισσόμενος τομέας είναι αυτός της μηχανικής μάθησης και των νευρωνικών δικτύων. Αυτός ο τομέας μέσω της πιο εύκολα διαθέσιμης υπολογιστικής ισχύς διευκολύνεται σημαντικά και επιτρέπει την χρήση αυτών των εργαλείων σε προβλήματα που στο παρελθόν θα ήταν ασύμφορη η χρήση τους. Σε αυτή την εργασία μελετάται η χρήση τεχνικών μηχανικής εκμάθησης για την κατανομή πόρων σε ασύρματα δίκτυα σε αντίθεση με την παραδοσιακή αλγοριθμική μέθοδο. Τέλος διαπραγματεύεται περιπτώσεις στις οποίες μπορεί να συμφέρει μια τέτοιου είδους προσέγγιση.

## **Abstract**

In the last few years humans' telecommunications needs and capabilities have rapidly increased. Nowhere is that more obvious than in the wireless communications sector. Wireless networks are, at this point, responsible for a bigger data traffic than hardwired devices and as such, attempts are being made to optimize the protocols and techniques that comprise and realize said wireless networks. The data needs of the wireless internet are so great and the use of 4G and 5G so ubiquitous that challenges arise in the availability and distribution of resources. Another rapidly advancing sector is that of Machine Learning and subsequently neural networks, the, now more than ever, profound availability of processing power greatly empowers and makes the deployment of these tools easier than ever, even in problems where in the past their application wouldn't be feasible. This thesis investigates the use of machine learning techniques to distribute resources in wireless networks and contrasts it to a traditional algorithmic method. It also entertains scenarios in which such approaches might be applicable.

## ΠΕΡΙΕΧΟΜΕΝΑ

Κεφάλαιο 1 <sup>ο</sup> : Εισαγωγή .....	10
Κεφάλαιο 2 <sup>ο</sup> : Η Τεχνολογία MIMO στα Δίκτυα 5G και το Πρόβλημα Κατανομής Πόρων .....	14
2.1. Η τεχνολογία MASSIVE MIMO .....	16
2.2. Σενάρια HomNets .....	19
2.2.1. Η μέθοδος του Multi-layer sectorization .....	19
2.2.2. Η τεχνική Adaptive Beamforming .....	21
2.2.3. Συνεργασία μεγάλης κλίμακας (Large-Scale Cooperation) .....	22
2.3. Σενάρια HetNets .....	24
2.3.1. Ασύρματα backhaul δίκτυα .....	24
2.3.2. Κάλυψη Hotspot .....	25
Κεφάλαιο 3 <sup>ο</sup> : Εισαγωγή στη Μηχανική Μάθηση .....	28
3.1. Γραμμική Παραγωγή .....	29
3.2. Μηχανές Διανύσματος Υποστήριξης .....	30
3.3. Νευρωνικά Δίκτυα Τροφοδότησης .....	31
Κεφάλαιο 4 <sup>ο</sup> : Περιγραφή Πειραματικής Διαδικασίας .....	36
Κεφάλαιο 5 <sup>ο</sup> : Βελτιστοποίηση Ανάθεσης Πόρων με Τεχνικές Μηχανικής Μάθησης - Πειραματικά Αποτελέσματα .....	44
Κεφάλαιο 6 <sup>ο</sup> : Συμπεράσματα και Μελλοντική Εργασία .....	59
Παράρτημα Α: Κώδικας .....	61

## **Κατάλογος Εικόνων**

Εικόνα 2.1: Απεικόνιση multi-layer sectorization

Εικόνα 2.2: Απεικόνιση adaptive beamforming

Εικόνα 2.3: Απεικόνιση large-scale cooperation

Εικόνα 2.4: Απεικόνιση wireless backhaul

Εικόνα 2.5: Απεικόνιση hotspot coverage

Εικόνα 3.1: Γραφική αναπαράσταση ενός νευρωνικού δικτύου

Εικόνα 5.1: Μέσος όγκος ροής δεδομένων διαθέσιμος στους χρήστες

Εικόνα 5.2: Ταχύτητα εκτέλεσης

Εικόνα 5.3: Architecture of Net 1

Εικόνα 5.4: Error Histogram Net 1

Εικόνα 5.5: Best Validation Performance Net 1

Εικόνα 5.6: Training State Plot Net 1

Εικόνα 5.7: Architecture of Net 2

Εικόνα 5.8: Error Histogram Net 2

Εικόνα 5.9: Best Validation Performance Net 2

Εικόνα 5.10: Training State Plot Net 2

Εικόνα 5.11: Architecture of Net 3

Εικόνα 5.12: Error Histogram Net 3

Εικόνα 5.13: Best Validation Performance Net 3

Εικόνα 5.14: Training State Plot Net 3

Εικόνα 5.15: Architecture of Net 4

Εικόνα 5.16: Error Histogram Net 4

Εικόνα 5.17: Best Validation Performance Net 4

Εικόνα 5.18: Training State Plot Net 4

Εικόνα 5.19: Architecture of Net 5

Εικόνα 5.20: Error Histogram Net 5

Εικόνα 5.21: Best Validation Performance Net 5

Εικόνα 5.22: Training State Plot Net 5

## Κεφάλαιο 1<sup>ο</sup>: Εισαγωγή

Με την έλευση των κινητών τηλεφώνων και γενικά των συνδεδεμένων στο διαδίκτυο ασύρματων συσκευών, καθώς και των όλο και πιο ετερογενών περιπτώσεων χρήσης, οδηγούμαστε σε μια ραγδαία αύξηση της ποσότητας δεδομένων που μεταφέρονται μέσω των δικτύων κινητής τηλεφωνίας. Συγκεκριμένα, μελέτες που προέρχονται από την βιομηχανία, δείχνουν την κίνηση των δικτύων κινητής τηλεφωνίας να υπερβαίνει αυτής των υπολογιστών [1].

Ωστόσο, είναι προφανές ότι η εξυπηρέτηση των χρηστών κινητής τηλεφωνίας αντιμετωπίζει προβλήματα λόγω της αύξησης της κίνησης δεδομένων. Μακροπρόθεσμα, οι πάροχοι internet (*ISPs*) επιχειρούν να εγκαταστήσουν πιο πυκνές εγκαταστάσεις ("*base stations*"), ωστόσο, υπάρχει περιθώριο και για την βελτίωση της αποτελεσματικότητας της χρήσης των ήδη υπάρχοντων πόρων. Αυτή η νέα στροφή προς τις ασύρματες συνδέσεις οδηγεί τους ISP σε νέες "έξυπνες" ετερογενείς αρχιτεκτονικές και εργαλεία ώστε να χειριστούν τον φόρτο εργασίας ικανοποιητικά. Βέβαια, η κατάσταση δυσχεραίνεται και από την συνεχώς εξελισσόμενη και όλο και πιο περίπλοκη δομή των δικτύων, η οποία καθιστά την εφαρμογή πιο παραδοσιακών αλγόριθμων και ευρετικών μεθόδων, όχι μεν αδύνατη, αλλά σαφώς πιο δυσχερή.

Για τους λόγους αυτούς έχει δημιουργηθεί ένα πολύ έντονο επιστημονικό ενδιαφέρον για την εφαρμογή στοιχείων μηχανικής εκμάθησης στα μελλοντικά δίκτυα 5ης γενιάς. Οι MIMO (**M**ultiple **I**nput **M**ultiple **O**utput) τεχνολογίες αποτελούν πλέον μια θεμελιώδη τεχνική στις σύγχρονες ασύρματες επικοινωνίες, είτε αυτές είναι για WiFi είτε για την κινητή τηλεφωνία. Τα δίκτυα 5ης γενιάς υποστηρίζουν πολύ μεγάλους ρυθμούς μετάδοσης δεδομένων, ενσωματώνοντας αλγόριθμους βελτιστοποίησης, μεγιστοποιώντας την επίδοση των τεχνικών MIMO και κατ' επέκταση του πρωτοκόλλου που χρησιμοποιείται για το 5G.

Ο σκοπός αυτής της διπλωματικής εργασίας είναι να εξεταστεί μια εναλλακτική λύση από αυτή που είχε προταθεί στην ερευνητική εργασία [*A mechanism for 5G MIMO performance optimization and evaluation Dimosthenis Tzanakos, Foivos Allayiotis, Vasileios Kokkinos, and Christos Bouras, Computer Engineering and Informatics Department, University of Patras*] και αφορούσε στην ανάθεση πόρων ως προς του χρήστες δικτύων 5G [1]. Στην δημοσίευση αυτή χρησιμοποιήθηκε μία παραλλαγή του *Hungarian algorithm*, με σκοπό να δημιουργήσει ζευγάρια χρηστών με τις κυψέλες κεραιών με τέτοιο τρόπο ώστε να μεγιστοποιεί την συνολική ροή δεδομένων. Αυτή η προσέγγιση εγγυάται βέλτιστη λύση πολυπλοκότητας  $O(n^3)$  στο πρόβλημα ανάθεσης. Στην παρούσα διπλωματική προτείνεται μια εναλλακτική λύση η οποία χρησιμοποιεί τα νευρωνικά δίκτυα με σκοπό την μείωση της πολυπλοκότητας της λύσης. Και οι δύο αυτές προσεγγίσεις κάνουν χρήση του *DEEPMIMO dataset generator* [2], το οποίο προσομοιώνει το περιβάλλον μίας αστικής περιοχής, με ένα επιλεγμένο από τον ερευνητή αριθμό κεραιών (*base stations*) και κυψελών (*antennae*) και χρηστών, τοποθετημένων στον χώρο. Παράγοντες που επηρεάζουν τον ρυθμό δεδομένων που μπορεί να επιτύχει κάθε χρήστης είναι η απόσταση, αλλά και εμπόδια που μπορεί να παρεμβάλλονται μεταξύ χρήστη και κυψελών. Πρακτικά αυτό που το νευρωνικό δίκτυο μαθαίνει είναι την τοπολογία του προσομοιωμένου περιβάλλοντος, όπως αυτή επηρεάζει την ισχύ του σήματος των κυψελών ως προς τους χρήστες.

Η συγκεκριμένη διπλωματική εργασία δομείται ως εξής:

- Το κεφάλαιο 2 διαπραγματεύεται την τεχνολογία MIMO σε δίκτυα 5ης γενιάς και τις προκλήσεις που την δίδουν. Εν συνεχεία ορίζονται και επεξηγούνται οι δύο βασικές μορφολογίες ασύρματων δικτύων σε συνδυασμό με τις βασικές τεχνικές που χρησιμοποιούνται για να αντιμετωπίσουν σύνθη προβλήματα κατανομής πόρων στους χρήστες.

- Το κεφάλαιο 3 εξηγεί συνοπτικά κάποιες βασικές έννοιες που αφορούν την Μηχανική Μάθηση και το μαθηματικό της υπόβαθρο. Επίσης επεξηγούνται τρεις βασικές τεχνικές οι οποίες χρησιμοποιούνται και στην λύση που προτείνεται απο αυτή την εργασία

- Το κεφάλαιο 4 περιγράφει την πειραματική διαδικασία και τα εργαλεία που χρησιμοποιούνται για την προσομοίωση πάνω στην οποία γίνεται η σύγκριση της παραδοσιακής αλγοριθμικής προσέγγισης με την λύση που εξετάζεται σε αυτή την εργασία

- Το κεφάλαιο 5 εξετάζει τα αποτελέσματα της πειραματικής διαδικασίας και αναλύει διαφορετικές παραλλαγές της λύσης. Παράλληλα κάποιες υποθέσεις γίνονται που εξηγούν τα αποτελέσματα

- Το κεφάλαιο 6 εξετάζει την ορθότητα η μη της προτεινόμενης λύσης και μερικές βελτιώσεις σε αυτή. Τέλος, δίνονται κάποιες ιδέες για μελλοντική έρευνα και κατευθύνσεις ενδιαφέροντος.

Στο Παράρτημα A παρατίθενται κάποια επεξηγηματικά κομμάτια του κώδικα που χρησιμοποιήθηκε σε περιβάλλον Matlab.

## **Βιβλιογραφία**

- [1] Cisco Visual Networking Index: Forecast and Methodology, 2016–2021, Cisco, San Jose, CA, USA, Jun. 2017.
- [2] Tzanakos, D., Allayiotis, F., Kokkinos, V. and Bouras, C. (2021) 'A mechanism for 5G MIMO performance optimization and evaluation', 13th IFIP Wireless and Mobile Networking Conference (WMNC), pp. 48-54, doi: 10.23919/WMNC53478.2021.9619012.
- [3] <http://www.DeepMIMO.net>

## Κεφαλαίο 2ο: Η Τεχνολογία MIMO στα Δίκτυα 5G και το Πρόβλημα Κατανομής Πόρων

Το *MIMO* (**M**ultiple-**I**nput and **M**ultiple-**O**utput) κρίνεται ένα σημαντικό και επίκαιρο θέμα σε μεγάλο βαθμό λόγω της έλευσης των δικτύων 5G και γενικότερα του μέλλοντος των ασύρματων επικοινωνιών. Προσφέροντας βελτιωμένους βαθμούς ελευθερίας τα δίκτυα 5ης γενιάς είναι ικανά να εξυπηρετούν ταυτόχρονα πολλαπλούς χρήστες με *high gains*, και ως εκ τούτου βελτιώνοντας την φασματική απόδοση, την ενεργειακή απόδοση αλλά και την αξιοπιστία.

Αν και το κινητό ίντερνετ (*mobile internet*) και το *internet of things* (*IoT*) είναι οι δυο κύριοι οδηγοί αυτής της εξέλιξης, εντούτοις δεν είναι οι μοναδικοί [1].

Τα δίκτυα 5ης γενιάς προσθέτουν πολλές τεχνολογίες σε σχέση με την προηγούμενη γενιά, όπως, **Software Defined Networking** (*SDN*), **Device-to-Device** (*D2D*) communications, **Unmanned Aerial Vehicles** (*UAV*), **Cloud Radio Access Networks** (*CRANs*), **Mobile Edge Computing** (*MEC*) και cloud computing με τα οποία πραγματοποιούμε το λεγόμενο '*internet of everything*'.

Για να υποστηριχθούν τα παραπάνω, κάποιοι στόχοι σχετικοί με την απόδοση σε σχέση με την προηγούμενη γενιά, δηλ. του 4G, έχουν τοποθετηθεί από την 3GPP (το συνονθύλευμα οργανισμών που παράγουν τα πρωτόκολλα για τις κινητές τηλεπικοινωνίες). Συγκεκριμένα:

- Αύξηση ταχύτητας διακίνησης κατά 3 τάξεις μεγέθους.
- Ενεργειακή απόδοση και χρόνος επικοινωνίας end-to-end κατά μία τάξη μεγέθους.
- Αύξηση του αριθμού του πιθανού συνδεδεμένου εξοπλισμού 10 με 100 φορές.
- Αύξηση της διάρκειας της μπαταρίας εξοπλισμού χαμηλής κατανάλωσης ενέργειας κατά 10 φορές.

Όσον αφορά την χωρητικότητα των συστημάτων 5G τρεις βασικές τεχνικές χρησιμοποιούνται. Στην πρώτη από αυτές, την *massive MIMO*, που χρησιμοποιήθηκε για να βελτιώσει την φασματική απόδοση, χρησιμοποιούνται φασματικοί πόροι *mm-wave* για να επεκταθεί το εύρος ζώνης του συστήματος και πολλαπλών επιπέδων (*multi-layer*) και υπερ-πυκνά (*ultra-dense*) δίκτυα χρησιμοποιούνται για να αυξήσουν την γεωγραφική επαναχρησιμοποίηση φασμάτων. Τα συστήματα που κάνουν χρήση μεγάλων πλεγμάτων κεραιών για να εξυπηρετήσουν πλήθη χρηστών ονομάζονται *massive MIMO* συστήματα επικοινωνίας. Τα *massive MIMO* είναι ικανά να αντιμετωπίσουν την πολύ έντονη απώλεια ισχύος των *mm-wave* σημάτων, δρώντας ως ασύρματος φορέας οπισθό-ζευξης και εξομαλύνοντάς τις παρεμβολές *multi-layer ultra-dense* δικτύων. [2]

Η αποδοτική χρήση του εύρους ζώνης, ή του φάσματος, συνήθως είναι ο πιο σημαντικός παράγοντας για την επιλογή των τεχνολογιών της επόμενης γενιάς στα συστήματα ασύρματων επικοινωνιών. Ωστόσο, με την δραματική αύξηση της χρήσης αυτών των συστημάτων, κρίνεται σημαντικός και ο παράγοντας εξοικονόμησης ενέργειας, και για μείωση του οικολογικού αποτυπώματος όσο και για το κόστος λειτουργίας. Συνεπώς, στα πλαίσια διατήρησης της αποδοτικής χρήσης των φασματικών πόρων γίνεται μια προσπάθεια προς την αποδοτικότερη χρήση ενέργειας αυτών των συστημάτων.

Οι τεχνικές *MIMO* παραμένουν στο επίκεντρο της προσοχής, όσον αφορά τις ασύρματες επικοινωνίες καθώς βελτιώνουν τον ρυθμό μετάδοσης δεδομένων αλλά και την εμβέλεια των ασύρματων συστημάτων, χωρίς να απαιτείται η αύξηση του εύρους ζώνης ή η ισχύος των κεραιών. Μάλιστα, κατατέθηκε και σχετική πατέντα η οποία κατοχυρώθηκε το 1993-1994, σύμφωνα με την οποία πολλαπλές κεραιές τοποθετούνται σε ένα πομπό με σκοπό να βελτιωθεί η ροή πληροφορίας μέσω χωρικής πολυπλεξίας. Τα συστήματα 4ης γενιάς (4G) κάνουν χρήση των *MIMO*, και συγκεκριμένα του *MU-MIMO*, δηλαδή πολλαπλών χρηστών οι οποίοι εξυπηρετούνται από τον ίδιο

πομπό. Υπάρχει η επιλογή μέσω του *MU-MIMO* είτε βελτίωσης της απόδοσης του εύρους ζώνης ή της αξιοπιστίας της μετάδοσης.

## 2.1. Η τεχνολογία MASSIVE MIMO

Όταν ένα τέτοιο σύστημα *MIMO* χρησιμοποιείται με εκατοντάδες κεραιές για να εξυπηρετήσει δεκάδες εξοπλισμούς χρηστών, τότε το σύστημα χαρακτηρίζεται ως *massive-MIMO*, γνωστό και ως *Large-scale MIMO (LS-MIMO)* και συχνά συσχετίζεται με ορολογίες που αφορούν μεγάλης κλίμακας συστήματα κεραιών, όπως για παράδειγμα *very large MIMO*, *very large MU-MIMO*, *MIMO* πλείστων διαστάσεων, κλπ.

Θεωρητικές μελέτες, αλλά και τα αποτελέσματα χρήσης του στον πραγματικό κόσμο δείχνουν ότι το *massive MIMO* είναι ικανό να βελτιώσει σημαντικά την αποδοτική χρήση του εύρους ζώνης, ενώ ταυτόχρονα μειώνει και την απαραίτητη ισχύ των κεραιών.

Η κλίση ως προς τον κάθετο άξονα σε μια συστοιχία κεραιών είναι σταθερή στην ‘παραδοσιακή’ *MIMO* τεχνολογία. Για να εκμεταλλευθούν τον κάθετο άξονα για την διάδοση του σήματος, συστοιχίες κεραιών με σφαιρικά, παραλληλόγραμμά και κυλινδρικά σχήματα δοκιμάστηκαν από το 3GPP. Ένα *MIMO* σύστημα με τέτοιου είδους συστοιχίες κεραιών μπορεί να ρυθμίσει και το αζιμούθιο και την κλίση, και ως εκ τούτου να διαδώσει σήματα σε τρισδιάστατο χώρο. Αυτό το σύστημα ονομάζεται *3D MIMO* και κάνει συχνά χρήση περισσότερων κεραιών για επέκταση της χωρητικότητας μέσω της αυξημένης δυνατότητας πολυπλεξίας. Πρακτικά μπορεί να θεωρηθεί ως μια εφαρμογή του *massive MIMO* ιδιαίτερα όταν έχουμε εξαιρετικά μεγάλες κεραιές.

Τα *massive-MIMO* μπορούν να βελτιώσουν την απόδοση του εύρους ζώνης καθώς μπορούν να επιτύχουν οφέλη στην πολυπλεξία, ενώ εξυπηρετούν δεκάδες χρήστες ταυτόχρονα. Σημαντική βελτίωση παρατηρείται και στην ενεργειακή απόδοση λόγω του γεγονότος ότι η χρήση πολλαπλών

κεραιών επιτρέπει την συγκέντρωση ενέργειας με μια εξαιρετικά λεπτή ακτίνα σε πολύ μικρές περιοχές όπου οι χρήστες είναι τοποθετημένοι. Αυτή η δυνατότητα συγκέντρωσης επίσης βοηθάει με την μείωση παρεμβολών από άλλους χρήστες (*Inter User Interference*). Λόγω των μεγάλων βαθμών ελευθερίας έχουμε βελτίωση και στην αντοχή του συστήματος. Έτσι, η απώλεια μιας κυψέλης, η και περισσότερων, δεν έχουν καταστροφικό αποτέλεσμα στην επίδοση ολόκληρου του συστήματος. Υπάρχουν μικρής πολυπλοκότητας αλγόριθμοι επεξεργασίας σημάτων ικανοί να προσεγγίσουν την επίδοση που επιτυγχάνουν από ιδεατές μεθόδους, όπως η **Maximum-Likelihood (ML)**, το **Multi-user detection and Dirty Paper Coding (DPC)** [3], [4].

Καθώς οι φασματικοί πόροι είναι περιορισμένοι, ο πλέον αποτελεσματικός τρόπος για την χρήση τους είναι ένα θέμα που ερευνάται. Τα δίκτυα 5ης γενιάς χρησιμοποιούν υψηλές συχνότητες για να επιτυγχάνουν υψηλό εύρος ζώνης. Ωστόσο το ερώτημα της ολικής κάλυψης και παρεμβάσεων είναι δυνητικά προβλήματα. Δεδομένων αυτών των προβλημάτων, σημαντικό μέρος της σχετικής έρευνας επικεντρώνεται στην μελέτη και αντιμετώπιση τους [5], λαμβάνοντας υπόψιν τις παρεμβολές, τον ρυθμό μεταφοράς δεδομένων και το σχετικό 'κόστος' του φάσματος. Αυτή η κατεύθυνση έρευνας ονομάζεται **Spectrum Resource Management (SRM)**.

Μία από αυτές τις προσεγγίσεις του *SRM* είναι τα λεγόμενα **Cognitive Radio Networks (CRNs)**. Στην μεθοδολογία αυτή, οι ερευνητές χωρίζουν το φάσμα σε ζώνες, στις κύριες ζώνες έχουν πρόσβαση οι πρωτεύοντες χρήστες (**Primary Users - PUs**) και στις εναλλακτικές οι δευτερεύοντες χρήστες (**Secondary Users - SUs**), εκτός και αν διαπιστωθεί αδρανής κύρια ζώνη, οπότε τότε μπορεί ένας δευτερεύων χρήστης να κάνει χρήση αυτής. Σημαντικός παράγοντας αυτής της προσέγγισης είναι ο συνωστισμός δικτύων σε συγκεκριμένες ζώνες, η κεντρική ιδέα είναι ότι ένας *SU* δεν παρεμποδίζει *PUs*. Αυτή είναι μία μέθοδος που επιτρέπει την αποτελεσματική χρήση του φάσματος.

Αν και δεν είναι ακόμα τελειοποιημένη, θεωρείται ότι θα έχει σημαντικό ρόλο στο μέλλον στην βελτίωση της επίδοσης του δικτύου.

Μία άλλη προσέγγιση, η οποία προτάθηκε το 2009 από την China Mobile είναι η Cloud Radio Access Network (*C-RAN*) η οποία ουσιαστικά θέτει τις κεραιές σε σύνδεση με μια κεντρική **BaseBand Unit (BBU)** η οποία έχει κεντρικό έλεγχο όλων των κεραιών των συνδεδεμένων σε αυτή, που επιτρέπει ένα συγκεντρωτικό έλεγχο χωρίς ανάγκη αποκεντρωμένων κανόνων για αποφάσεις σχετικές με λειτουργίες *routing*, *handshakes* κτλ. Ένα μειονέκτημα αυτού του κεντρικού ελέγχου είναι το υπολογιστικό κόστος, όταν έχουμε περιόδους μεγάλου φόρτου, και η γενικότερη ακαμψία του για επέκταση.

Στην αντίθετη κατεύθυνση από την προηγούμενη μεθοδολογία, βρίσκονται τα *Fog Radio-Access Networks (F-RAN)*, στα οποία ο υπολογιστικός φόρτος ανατίθεται με κατανομημένο τρόπο σε άλλο υλικό, εκτός του *BBU*, το οποίο συμπεριλαμβάνει σημεία πρόσβασης (*access points*), εξοπλισμό χρηστών κ.α. Αυτή η τελείως διαφορετική λογική του από τις προηγούμενες μεθόδους σημαίνει ότι υπάρχουν προβλήματα τα οποία απαιτούν επίλυση από την ερευνητική κοινότητα καθώς η διαχείριση του δικτύου και οι διάφοροι μηχανισμοί βελτιστοποίησης πρέπει αναγκαστικά να διαφέρουν ριζικά λόγω της αποκεντρωμένης φύσης της αρχιτεκτονικής. Η χρήση *MIMO* σε συνδυασμό με αυτή την τεχνική φαίνεται πολύ ελκυστική [3].

Πλέον, και καθώς πλησιάζουμε το λεγόμενο “*internet of everything*” τα δίκτυα πολύ συχνά εξελίσσονται από ομογενή (*Homogenous Networks-HomNets*) σε ετερογενή (*Heterogenous Networks-HetNets*), τα οποία έχουν πολλαπλούς τύπους κόμβων πρόσβασης. Η *3GPP*, από την Release 12 και μετά, υποστηρίζει ετερογενή δίκτυα, και συγκεκριμένα επιτρέπει σε τέτοια δίκτυα να έχουν διαφορετικά είδη μικρών *cells* και *macrocells*, και να μοιράζονται φασματικούς πόρους, κατάσταση η οποία επιτρέπει την δραματική βελτίωση της φασματικής απόδοσης και την μείωση των κενών

σημείων κάλυψης. Υπάρχουν τρεις βασικές στρατηγικές διαμοιρασμού των φασματικών πόρων [2], [3]:

- Η *Overlay Spectrum Sharing*: στην οποία χρήστες *microcells* μπορούν να χρησιμοποιήσουν φασματικούς πόρους που δεν χρησιμοποιούνται από τους χρήστες *macrocells*.
- Η *Underlay Spectrum Sharing*: στην οποία χρήστες *microcells* και *macrocells* μπορούν να έχουν πρόσβαση στους ίδιους φασματικούς πόρους ταυτόχρονα. Ωστόσο, κρίνεται απαραίτητο να ελεγχτεί η παρεμβολή από τον πομπό των *microcells* στους δέκτες των *macrocells*, θέτοντας κατάλληλα όρια στην ισχύ για αποφυγή *cross-tier* παρεμβολής.
- Η *Hybrid Spectrum Sharing*: στην οποία χωρίζονται οι φασματικοί πόροι σε κατηγορίες, π.χ. κάποιοι πόροι είναι προσβάσιμοι μόνο από του χρήστες *microcells* και κάποιοι σε χρήστες *macrocell* και *microcell*.

Στην ενότητα αυτή παρουσιάζονται διάφορα σενάρια χρήσης σε *massive MIMO* συστήματα τα οποία είναι ερευνητικού ενδιαφέροντος, για τις δύο βασικές κατηγορίες που αναφέρθηκαν προηγουμένως, δηλ. για τα ομογενή και ετερογενή δίκτυα.

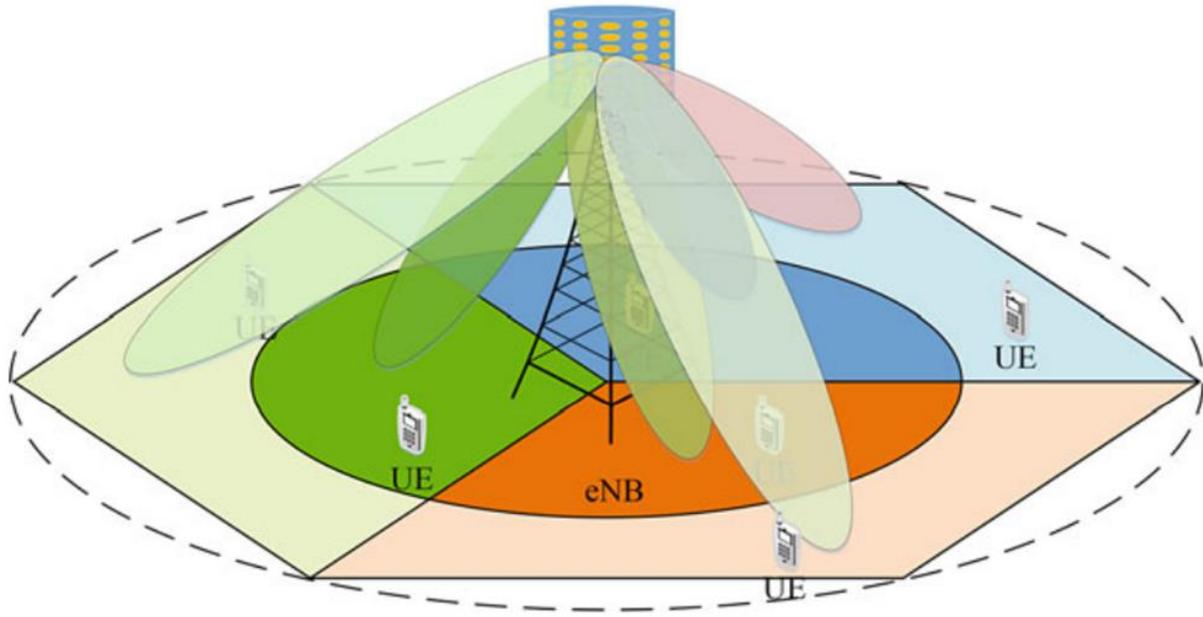
## 2.2. Σενάρια HomNets

### 2.2.1. Η μέθοδος του Multi-layer sectorization

Με την αύξηση των χρηστών και της κίνησης τους σε αστικά περιβάλλοντα, απαιτείται αυξημένη χωρητικότητα του συστήματος για την ικανοποιητική εξυπηρέτηση των χρηστών. Παραδοσιακά, τεχνικές οι οποίες χωρίζουν τομείς χρησιμοποιούνται για να παρέχουν υπηρεσίες σε έναν αυξανόμενο πληθυσμό καθώς αυτές οι τεχνικές αυξάνουν την χωρητικότητα του συστήματος. Σύμφωνα με αυτή την τεχνική τα κόστη εξοπλισμού μειώνονται καθώς μπορεί να χρησιμοποιηθεί ένα

*evolved node b (eNB)* για να εξυπηρετήσει 3 τομείς 120 μοιρών, ή 6 τομείς των 60 μοιρών. Ωστόσο ενώ αυτή η τμηματοποίηση κάνει ποιο αποδοτική χρήση του εύρους ζώνης, αυτή η βελτίωση έχει το μειονέκτημα ότι δημιουργεί αυξημένη παρεμβολή μεταξύ τμημάτων λόγω μη ιδεατής σχετικής θέσης κεραίας με το εκάστοτε τμήμα. Συνεπώς, για να αυξηθεί περαιτέρω η χωρητικότητα του δικτύου, πιο αποδοτικές τεχνικές κρίνονται ως απαραίτητες.

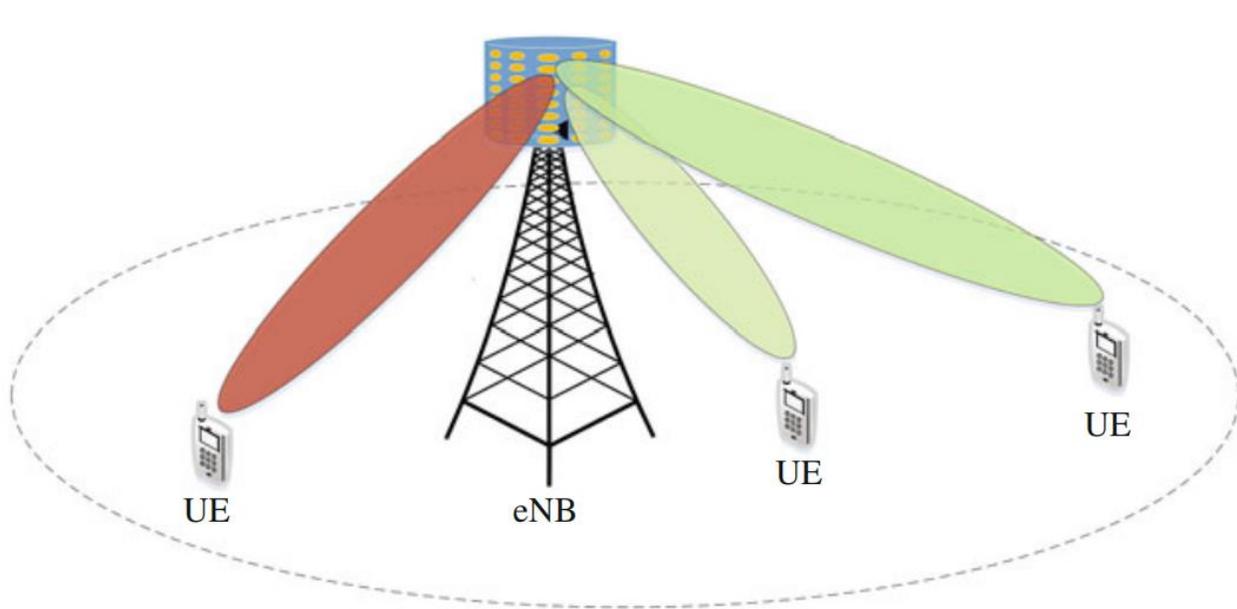
Όπως φαίνεται στην παρακάτω απεικόνιση στην εικόνα 2.1, ακριβής τμηματοποίηση σε *massive MIMO* συστήματα μπορεί να επιτευχθεί με επιλεκτικό οριζόντιο *angular beamforming*, τεχνική που είναι ικανή να μειώσει την παρεμβολή μεταξύ τομέων. Επιπλέον η κάλυψη της κάθε ακτίνας μπορεί να αλλαχθεί επηρεάζοντας την κλίση του τρισδιάστατου *beamforming*. Με αυτό τον τρόπο ένα παραδοσιακά σταθερό τμήμα μπορεί να χωριστεί περαιτέρω σε εσωτερικούς και εξωτερικούς τομείς, ο καθένας από τους οποίους μπορεί να εξυπηρετηθεί από έναν τρισδιάστατο *beamformer* με την ίδια οριζόντια κατεύθυνση αλλά διαφορετική κλίση. Πόροι ίδιας συχνότητας επαναχρησιμοποιούνται από όλους τους τομείς, πράγμα που επιτρέπει, δυνητικά, την σημαντική αύξηση του αριθμού χρηστών που εξυπηρετούνται ή και βελτιώνει την συνολική διακίνηση του δικτύου.



Εικόνα 2.1: Απεικόνιση multi-layer sectorization

### 2.2.2. Η τεχνική Adaptive Beamforming

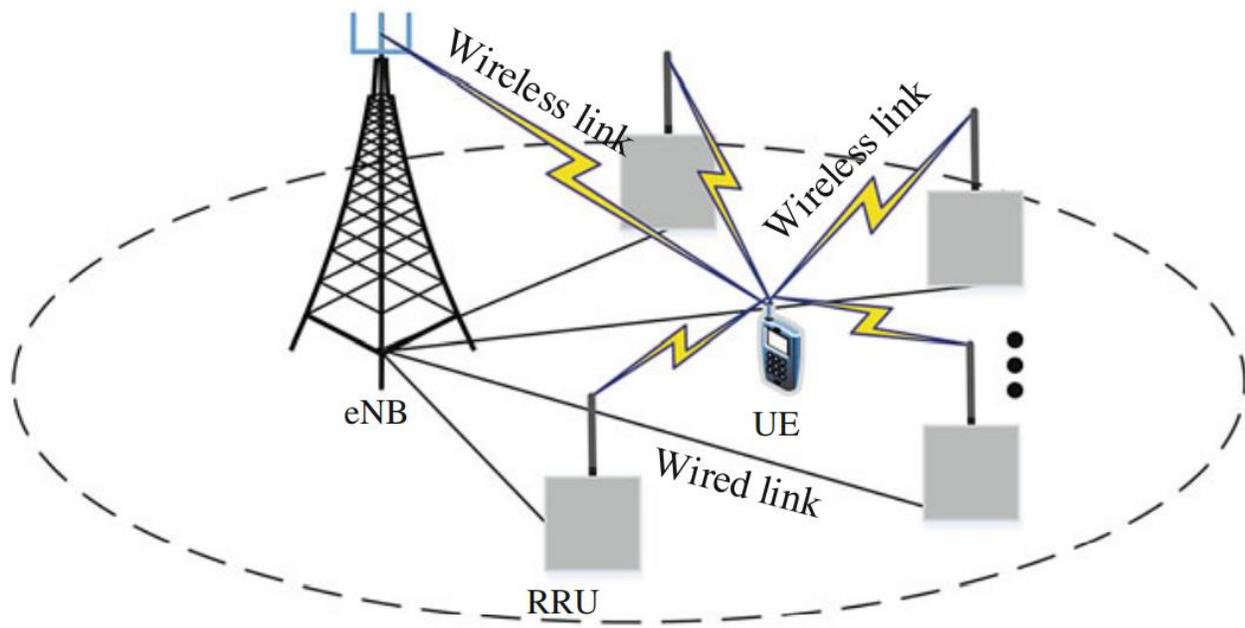
Οι σταθεροί ακτινοβολητές, ονομάζονται σταθεροί γιατί τα βάρη τα οποία πολλαπλασιάζουν τα σήματα σε κάθε στοιχείο της συστοιχίας κεραιών παραμένουν ίδια κατά την διάρκεια λειτουργίας. Αντιθέτως, στους προσαρμοστικούς ακτινοβολητές, τα βάρη συνεχώς ανανεώνονται με βάση τα λαμβανόμενα σήματα, με κατά τέτοιο τρόπο ώστε να ελαχιστοποιηθεί η χωρική παρεμβολή. Αυτή η διαδικασία μπορεί να πραγματοποιηθεί είτε στον τομέα του χρόνου είτε της συχνότητας. Ένας δισδιάστατος προσαρμοστικός ακτινοβολητής έχει λιγότερη ευελιξία από έναν τρισδιάστατο στην αξιοποίηση των φασματικών πόρων στον χωρικό τομέα.



Εικόνα 2.2: Απεικόνιση adaptive beamforming

### 2.2.3. Συνεργασία μεγάλης κλίμακας (Large-Scale Cooperation)

Η βιβλιογραφία σχετικά με το *massive MIMO* αναφέρει διάφορα πλεονεκτήματα για το σενάριο όπου ένας μεγάλος αριθμός κεραιών συσσωρεύονται σ' ένα σημείο. Ωστόσο, τέτοιες αρχιτεκτονικές δημιουργούν κάποιες δυσκολίες, και στην υλική σχεδίαση τους αλλά και στην εγκατάστασή τους. Ως εναλλακτική λύση προσφέρεται το κατακεκολλημένο σύστημα κεραιών (*Distributed Antenna Systems-DASs*) όπου οι κεραιές είναι χωρικά διαχωρισμένες με σκοπό την βελτίωση κάλυψης σε εσωτερικούς χώρους με χρήση συντηρητικού αριθμού κεραιών, όπως στην Εικόνα 2.3.



Εικόνα 2.3: Απεικόνιση large-scale cooperation.

Σχετικές έρευνες [6] έχουν δείξει ότι εκτός από την βελτιωμένη κάλυψη που παρέχει το κατακευμαμένο σύστημα κεραιών, έχει δυνατότητα σημαντικής βελτίωσης της απόδοσης του εύρους ζώνης ακόμα και παρουσία *Inter-Cell Interference (ICI)*. Αυτές οι έρευνες προκάλεσαν ενδιαφέρον για την ανάπτυξη σεναρίων στα οποία το *massive MIMO* σύστημα ευνοείται από την χρήση κατακευμαμένων κεραιών σε αντίθεση με την συγκεντρωτική προσέγγιση.

Το πλεονέκτημα του κατακευμαμένου *massive MIMO* είναι αναμενόμενο, καθώς τα σήματα, φτάνοντας από κατακευμαμένες κεραιές σε κάθε χρήστη, υπόκεινται ανεξάρτητα τυχαία επίπεδα μεγάλης κλίμακας εξασθένισης, συνεπώς οδηγώντας σε πιθανά κέρδη χωρητικότητας σε σχέση με την αντίστοιχη συγκεντρωτικού τύπου τοποθετημένη κεραία. Δυστυχώς είναι μία δύσκολη πρόκληση να επιτευχθούν αυτά τα κέρδη συντονίζοντας *intra-cell* παρεμβολές, ειδικά σε σεναρία με δεκάδες ή ακόμα και εκατοντάδες Remote Radio Units (RRUs) σε ένα κελί. Αν και η πλήρης συνεργασία αποτελεί μια αποτελεσματική μέθοδο για να εξαλειφθεί το *intra-cell interference*, εντούτοις δεν είναι πρακτικό, λόγω της εξάρτησης του από την πλήρη γνώση κατάστασης πληροφορίας καναλιού (*Channel State*

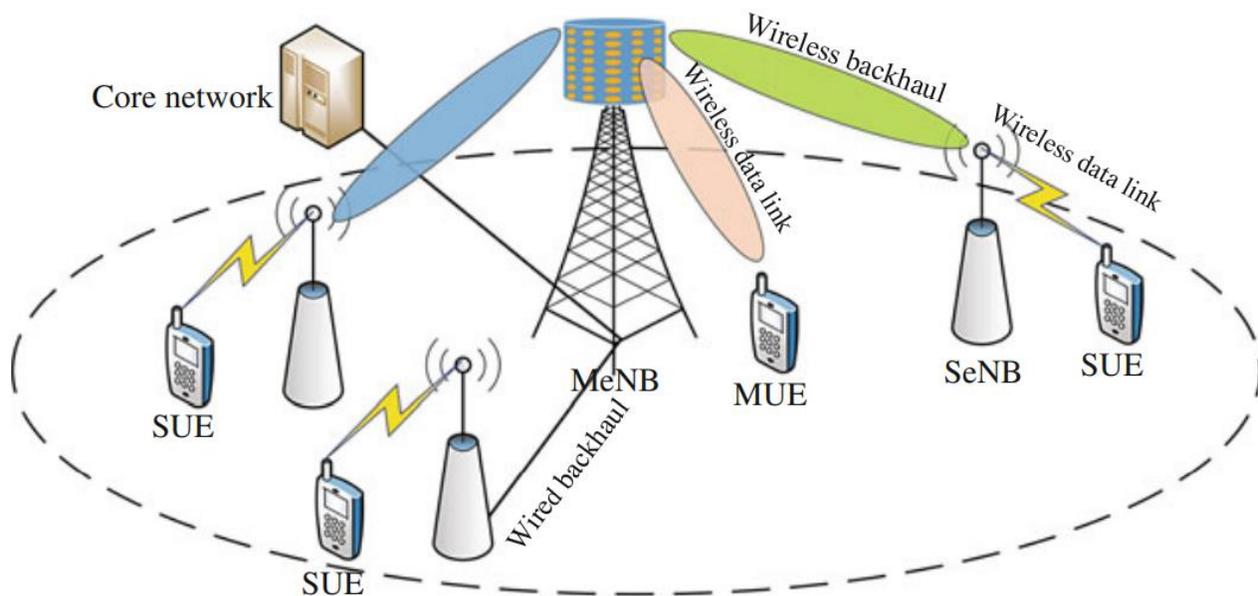
*Information-CSI*). Για να επιτευχθεί μια ισορροπία μεταξύ της επίδοσης και του υπολογιστικού κόστους, αποδοτικοί μεγάλης κλίμακας συνεργατικοί αλγόριθμοι είναι πολύ σημαντικοί. Γενικά, θεωρείται ότι τα κατανεμημένα *massive MIMO* και οι εγκαταστάσεις μικρών cells δεν είναι απαραίτητα ανταγωνιστικές τεχνολογίες, αλλά θα μπορούσαν κάλλιστα να είναι συνεργατικές και να λειτουργούν παράλληλα.

## 2.3. Σενάρια HetNets

### 2.3.1. Ασύρματα backhaul δίκτυα

Ένα ετερογενές δίκτυο με πυκνά μικρά κελιά θεωρείται μία υποσχόμενη αρχιτεκτονική ως προς την αποδοτική χρήση ενέργειας και εύρους ζώνης. Συνήθως αποτελείται από πολλαπλούς τύπους ασύρματης πρόσβασης κόμβους, με τους μικρότερους να είναι συνδεδεμένη με οπισθόζευξη στους μεγαλύτερους μακρο-κόμβους (*Macro-cell evolved Node B-MeNB*). Γενικά προτιμάται ασύρματη σύνδεση έναντι της ενσύρματης λόγω της ευκολίας εγκατάστασης του υλικού. Σε αυτά τα σενάρια ένα *massive MIMO* χρησιμοποιείται τους μακρο-κόμβους οι οποίοι συνήθως έχουν μεγάλο βαθμό ελευθερίας έτσι ώστε να μπορεί να υποστηρίξει πολλαπλές οπισθοζεύξεις στο ετερογενές δίκτυο [7].

Όπως φαίνεται και στην Εικόνα 3.4, το ίδιο φάσμα μπορεί να επαναχρησιμοποιηθεί μεταξύ ασύρματων οπισθοζεύξεων, δηλαδή μπορεί να 'αντιμετωπίσει' τα *SeNBs* ως χρήστες που μέσω των *MeNBs* επικοινωνούν ασύρματα με οπισθοζεύξη. Καθώς η τοποθεσία ενός *eNB* είναι συνήθως στατική, το κανάλι της ασύρματης οπισθόζευξης είναι ημισταθερό στην κλίμακα του χρόνου (*quasi-static time varying*) Συνεπώς το *MeNB* είναι ικανό να εξουδετερώσει τυχών παρεμβολές μεταξύ της ασύρματης οπισθοζεύξης και των *MUEs* μέσω precoding.

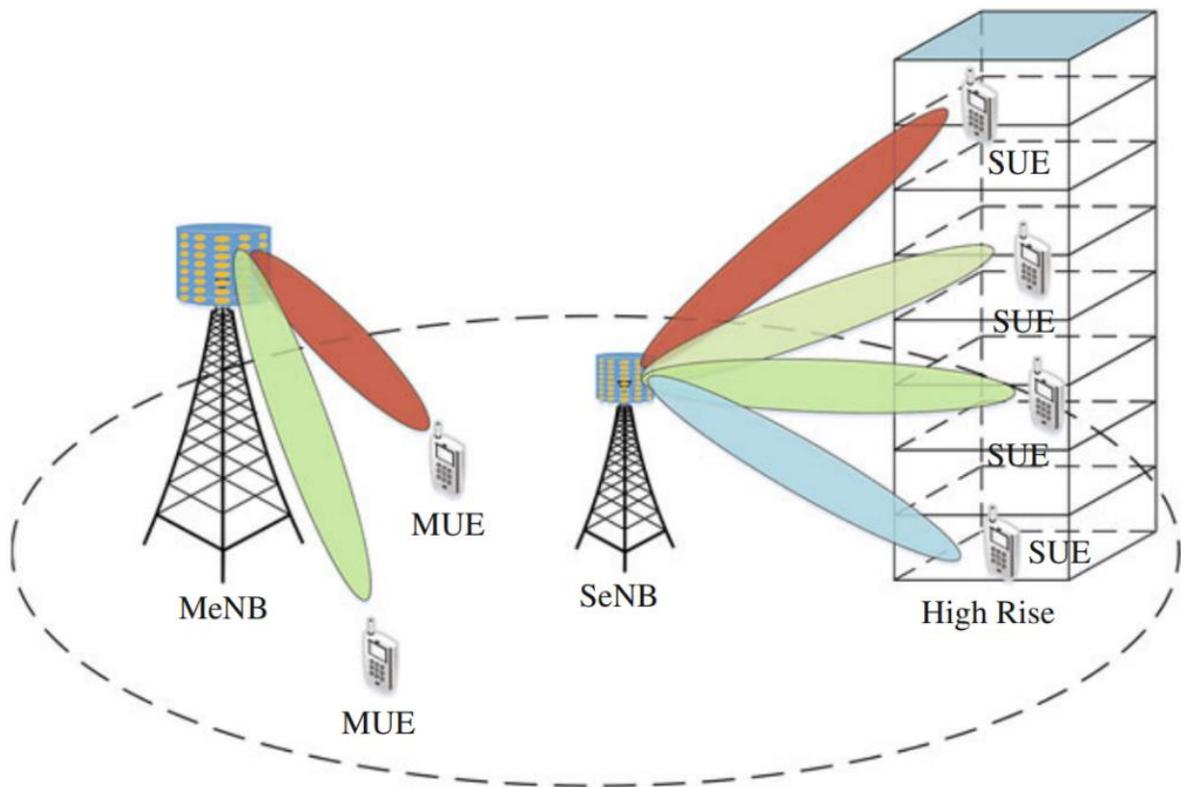


Εικόνα 2.4: Απεικόνιση wireless backhaul

### 2.3.2. Κάλυψη Hotspot

Στατιστικές έρευνες δείχνουν ότι η πλειοψηφία της κινητής ροής δεδομένων προέρχεται από το εσωτερικό κτιρίων, όπως π.χ. σούπερ-μάρκετ, γραφεία, σχολεία, κ.α. Συνεπώς η απαίτηση για υψηλής ποιότητας κάλυψη εσωτερικών χώρων κρίνεται ως ένα τα πολύ σημαντικό ζήτημα για τα ετερογενή δίκτυα. Καθώς αυτή η μεταβαλλόμενη ροή δεδομένων, σε μεγάλο βαθμό, προέρχεται από διαφορετικά ύψη (π.χ. πολυώροφα κτίρια) παραδοσιακά συστήματα κεραιών με στατική κλίση τα οποία είναι κυρίως σχεδιασμένα για κίνηση χρηστών στο επίπεδο του εδάφους δεν κρίνονται κατάλληλα για αυτή την χρήση. Αντιθέτως τα massive συστήματα κεραιών είναι ικανά να προσαρμόσουν και το αζιμούθιο και την κλίση των δεσμών (beams) τους. Αυτό επιτρέπει μετάδοση κατευθείαν στους χρήστες σε διαφορετικούς ορόφους του κτιρίου [8] και έτσι βελτιώνει την μέγιστη δυνατότητα διακίνησης του συστήματος. Ωστόσο, όταν η εσωτερική κάλυψη του κτιρίου πραγματοποιείται από MeNB με massive AA (Antenna Array), το εύρος κλίσεων που μπορεί να επιτύχει παραμένει μικρό σε σχέση με αυτό που μπορεί να επιτύχει ένα SeNB και η γωνιακή ανάλυση

δεν μπορεί να ικανοποιήσει τις ανάγκες των χρηστών όπως δείχνεται σχηματικά και στην Εικόνα 3.5. Είναι γνωστό ότι η μικρή απόσταση μεταξύ χρηστών και *SeNBs* οδηγεί σε μειωμένες απώλειες κατά την διαδρομή (*path losses*). Καταλήγοντας, τα *SeNBs* εξοπλισμένα με *massive AA* θεωρούνται πιο κατάλληλα για εσωτερική κάλυψη κτιρίων, με προφανές μειονέκτημα το κόστος εγκατάστασης.



Εικόνα 2.5: Απεικόνιση hotspot coverage

## **Βιβλιογραφία**

- [1] Kamal, M.A., Raza, H.W., Alam, M.M., Su'ud, M.M. and Sajak, A.B. 2021. 'Resource Allocation Schemes for 5G Network: A Systematic Review'. *Sensors*, 21(19), p.6588. <https://doi.org/10.3390/s21196588>.
- [2] Zhao, L., Zhao, H., Zheng, K. and Xiang, W. 2018. Massive MIMO in 5G networks: selected applications. Switzerland: Springer International Publishing.
- [3] Xu, Y., Gui, G.H., Gacanin and Adachi, F. 2021. 'A Survey on Resource Allocation for 5G Heterogeneous Networks: Current Research, Future Trends, and Challenges'. *IEEE Communications Surveys & Tutorials*, 23(2), p. 668-695, doi: 10.1109/COMST.2021.3059896.
- [4] Chien, W.C., Huang, S.Y., Lai, C.F. and Chao, H.C. 2020. 'Resource management in 5G mobile networks: Survey and challenges'. *Journal of Information Processing Systems*, 16(4), pp.896-914.
- [5] Barri, E., Bouras, C., Kokkinos, V., A. Koukouvela. 2021. 'A Mechanism for Improving the Spectral Efficiency in mu-MIMO for 5G and Beyond Networks'. *MobiWac '21: Proceedings of the 19th ACM International Symposium on Mobility Management and Wireless Access*. p. 11-16. Pages 11–16. doi.org/10.1145/3479241.3486681.
- [6] Zheng, K., Zhao, L., Mei, J., Shao, B., Xiang, W. and Hanzo, L., 2015. Survey of large-scale MIMO systems. *IEEE Communications Surveys & Tutorials*, 17(3), pp.1738-1760.
- [7] Ge, X., Cheng, H., Guizani, M. and Han, T., 2014. 5G wireless backhaul networks: challenges and research advances. *IEEE network*, 28(6), pp.6-11.
- [8] Jiang, D. and Liu, G., 2017. An overview of 5G requirements. *5G Mobile Communications*, pp.3-26.

## Κεφάλαιο 3<sup>ο</sup>: Εισαγωγή στη Μηχανική Μάθηση

Η Μηχανική Μάθηση στηρίζεται στην εφαρμογή ενός συνόλου μεθόδων που βασίζονται σε δεδομένα, και είναι ικανές να μιμούνται, να κατανοούν και να βοηθούν εργασίες επεξεργασίας ανθρώπινων και βιολογικών πληροφοριών. Σε αυτή την επιδίωξη, προκύπτουν πολλά σχετικά ζητήματα, όπως ο τρόπος συμπίεσης δεδομένων, ερμηνείας και επεξεργασίας τους. Συχνά αυτές οι μέθοδοι δεν κατευθύνονται απαραίτητα στη μίμηση άμεσης ανθρώπινη επεξεργασίας αλλά μάλλον στην ενίσχυση τους, όπως π.χ. στην πρόβλεψη του χρηματιστηρίου, στην ταχεία ανάκτηση πληροφοριών κλπ. Σε αυτή τη προσέγγιση, η χρήση θεωρίας πιθανοτήτων είναι απαραίτητη αφού αναπόφευκτα τα περιορισμένα δεδομένα μας και η περιορισμένη κατανόηση του προβλήματος περιέχουν αυξημένη αβεβαιότητα. Με την ευρεία έννοια, η Μηχανική Μάθηση και οι σχετικές πτυχές της στοχεύουν στο να «μάθουν κάτι χρήσιμο» για το περιβάλλον μέσα στο οποίο λειτουργεί ο οργανισμός. Η Μηχανική Μάθηση είναι επίσης στενά συνδεδεμένη με την Τεχνητή Νοημοσύνη, με τη Μηχανική Μάθηση να τοποθετεί περισσότερα έμφαση στη χρήση δεδομένων για την οδήγηση και την προσαρμογή του μοντέλου.

Η Μηχανική Μάθηση αφορά ουσιαστικά την εξαγωγή συμπερασμάτων από μεγάλα σύνολα δεδομένων. Συχνά το κίνητρο είναι να παραχθεί τελικά ένας αλγόριθμος που να μπορεί είτε να μιμηθεί είτε να βελτιώσει την ανθρώπινη/βιολογική απόδοση. Σε γενικές γραμμές, τα δύο κύρια υποπεδία της μηχανικής μάθησης είναι η εποπτευόμενη μάθηση (*supervised learning*) και η μη επιβλεπόμενη μάθηση (*unsupervised learning*). Στην εποπτευόμενη μάθηση ο στόχος είναι η ακριβή πρόβλεψη, ενώ στην μη επιβλεπόμενη μάθηση ο στόχος είναι να περιγράψουν και να συμπυκνωθούν τα δεδομένα. Και στις δύο περιπτώσεις, κάποιος ενδιαφέρεται για μεθόδους που γενικεύουν καλά σε δεδομένα πάνω στα οποία δεν έχουν εκπαιδευτεί προηγουμένως. Με αυτή την έννοια, γίνεται

διάκριση μεταξύ δεδομένων που χρησιμοποιούνται για την εκπαίδευση ενός μοντέλου και δεδομένων που χρησιμοποιούνται για τον έλεγχο της απόδοσης του εκπαιδευμένου μοντέλου.

### 3.1. Γραμμική Παραγωγή

Η γραμμική παλινδρόμηση είναι ένα στατιστικό εργαλείο για τη μοντελοποίηση της σχέσης μεταξύ ορισμένων «επεξηγηματικών» μεταβλητών και ορισμένων πραγματικά αποτιμώμενων αποτελεσμάτων. Θεωρώντας το μαθησιακό πρόβλημα, το σύνολο τομέα (*domain set*)  $X$  είναι ένα υποσύνολο του  $\mathbb{R}^d$ , για κάποιες τιμές του  $d$ , και το σύνολο ετικετών  $Y$  είναι το σύνολο των πραγματικών αριθμών. Το ζητούμενο είναι μια γραμμική συνάρτηση  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  που προσεγγίζει καλύτερα τη σχέση μεταξύ των μεταβλητών μας (για παράδειγμα, πρόβλεψη του βάρους ενός μωρού ως συνάρτηση της ηλικίας του και του βάρους του κατά τη γέννηση). Η υποθετική κατηγορία των προβλέψεων της γραμμικής παλινδρόμησης είναι απλώς το σύνολο των γραμμικών συναρτήσεων:

$$\mathcal{H}_{\text{reg}} = L_d = \{x \mapsto \langle \mathbf{w}, \mathbf{x} \rangle + b : \mathbf{w} \in \mathbb{R}^d, b \in \mathbb{R}\}$$

Στη συνέχεια ορίζεται μια συνάρτηση απώλειας για παλινδρόμηση. Ενώ στην ταξινόμηση ο ορισμός της απώλειας (*loss*) είναι ξεκάθαρος, καθώς το  $(h, (x, y))$  απλώς υποδεικνύει εάν το  $h(x)$  προβλέπει σωστά το  $y$  ή όχι. Εάν όμως το βάρος ενός μωρού είναι 3 κιλά, και οι δυο προβλέψεις 3,00001 κιλών και 4 κιλών είναι «λανθασμένες», αλλά σαφώς προτιμάται το πρώτο από το δεύτερο. Επομένως, πρέπει να οριστεί πόσο θα «τιμωρηθούμε» από την απόκλιση μεταξύ  $h(x)$  και  $y$ . Ένας συνηθισμένος τρόπος είναι η χρήση της συνάρτησης τετραγωνικής απώλειας (*squared loss function*), δηλαδή:  $\ell(h, (x, y)) = (h(x) - y)^2$

Για αυτή τη συνάρτηση απώλειας (*loss function*), η εμπειρική συνάρτηση κινδύνου ονομάζεται Μέσο Τετράγωνο Σφάλμα (*Mean Squared Error*), δηλαδή:

$$L_S(h) = \frac{1}{m} \sum_{i=1}^m (h(\mathbf{x}_i) - y_i)^2$$

Φυσικά, υπάρχει μια ποικιλία άλλων συναρτήσεων απώλειας που μπορούν να χρησιμοποιηθούν, όπως για παράδειγμα, η συνάρτηση απώλειας απόλυτης τιμής:

$$\ell(h, (\mathbf{x}, y)) = |h(\mathbf{x}) - y|$$

Ο κανόνας *ERM* (*Empirical Risk Minimization*) για τη συνάρτηση απώλειας απόλυτης τιμής μπορεί να εφαρμοστεί χρησιμοποιώντας γραμμικό προγραμματισμό.

### 3.2. Μηχανές Διανύσματος Υποστήριξης

Στη μηχανική μάθηση, οι μηχανές διανύσματος υποστήριξης (*Support Vector Machine-SVM*, επίσης δίκτυα διανυσμάτων υποστήριξης) είναι εποπτευόμενα μοντέλα μάθησης με σχετικούς αλγόριθμους μάθησης που επεξεργάζονται δεδομένα για την ανάλυση ταξινόμησης και παλινδρόμησης. Τα *SVM* αναπτύχθηκαν στα AT&T Bell Laboratories από τον V. Vapnik και τους συνεργάτες του [1]. Τα *SVM* είναι μια από τις πιο ισχυρές μεθόδους πρόβλεψης, που βασίζονται σε στατιστικά πλαίσια μάθησης ή τη θεωρία VC (Vapnik–Chervonenkis theory) που προτάθηκαν από τους Vapnik (1982, 1995) και Chervonenkis (1974). Λαμβάνοντας υπόψη ένα σύνολο παραδειγμάτων εκπαίδευσης, το καθένα επισημαίνεται ότι ανήκει σε μία από τις δύο κατηγορίες (*Binary Classifier*), ένας αλγόριθμος εκπαίδευσης *SVM* δημιουργεί ένα μοντέλο που εκχωρεί νέα παραδείγματα στη μία ή στην άλλη κατηγορία, καθιστώντας τον έναν μη πιθανολογικό δυαδικό γραμμικό ταξινομητή. Το *SVM* αντιστοιχίζει παραδείγματα εκπαίδευσης σε σημεία στο χώρο, με τέτοιο τρόπο ώστε να μεγιστοποιήσει την απόσταση μεταξύ των δύο κατηγοριών. Στη συνέχεια, τα νέα παραδείγματα χαρτογραφούνται στον ίδιο χώρο και προβλέπεται ότι ανήκουν σε μια κατηγορία με βάση σε ποια πλευρά του υπερ-επιπέδου (*hyperplane*) τοποθετούνται .

Εκτός από την εκτέλεση γραμμικής ταξινόμησης, τα *SVM* μπορούν να εκτελέσουν αποτελεσματικά μη γραμμική ταξινόμηση, χρησιμοποιώντας το κόλπο του πυρήνα (*kernel trick*) [2] χαρτογραφώντας τις εισόδους τους σε χώρους υψηλότερων διαστάσεων (*high dimensional feature spaces*).

Όταν τα δεδομένα δεν φέρουν ετικέτα, η εποπτευόμενη μάθηση δεν είναι δυνατή και απαιτείται μια μη εποπτευόμενη μαθησιακή προσέγγιση, η οποία προσπαθεί να βρει φυσική ομαδοποίηση των δεδομένων και στη συνέχεια να χαρτογραφήσει/τοποθετήσει (*mapping*) τα νέα δεδομένα σε αυτές τις σχηματισμένες ομάδες. Ο αλγόριθμος ομαδοποίησης διανυσμάτων υποστήριξης [3], που δημιουργήθηκε από τους H. Siegelmann και V. Vapnik, εφαρμόζει τα στατιστικά στοιχεία των διανυσμάτων υποστήριξης, που αναπτύχθηκαν στον αλγόριθμο των *SVM*, για την κατηγοριοποίηση δεδομένων χωρίς ετικέτα [4].

### 3.3. Νευρωνικά Δίκτυα Τροφοδότησης

Ένα τεχνητό νευρωνικό δίκτυο (*Artificial Neural Network-ANN*) είναι ένα σύνολο αλγορίθμων που προσπαθεί να αναγνωρίσει τις υποκείμενες σχέσεις σε ένα σύνολο δεδομένων μέσω μιας διαδικασίας που μιμείται τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου. Με αυτή την έννοια, τα νευρωνικά δίκτυα αναφέρονται σε συστήματα νευρώνων, είτε οργανικής είτε τεχνητής φύσης [3].

Ένα *ANN* βασίζεται σε μια συλλογή συνδεδεμένων μονάδων ή κόμβων που ονομάζονται τεχνητοί νευρώνες, οι οποίοι μοντελοποιούν χαλαρά τους νευρώνες σε έναν βιολογικό εγκέφαλο. Κάθε σύνδεση, όπως οι συνάψεις σε έναν βιολογικό εγκέφαλο, μπορεί να μεταδώσει ένα σήμα σε άλλους νευρώνες. Ένας τεχνητός νευρώνας λαμβάνει ένα σήμα και στη συνέχεια το επεξεργάζεται και μπορεί να σηματοδοτήσει τους νευρώνες που συνδέονται με αυτόν. Το "σήμα" σε μια σύνδεση

είναι ένας πραγματικός αριθμός και η έξοδος κάθε νευρώνα υπολογίζεται από κάποια μη γραμμική συνάρτηση του αθροίσματος των εισόδων του. Οι συνδέσεις ονομάζονται “ακμές”. Οι νευρώνες και τα άκρα έχουν συνήθως ένα βάρος που προσαρμόζεται καθώς προχωρά η μάθηση. Το βάρος αυξάνει ή μειώνει την ισχύ του σήματος σε μια σύνδεση. Οι νευρώνες μπορεί να έχουν ένα κατώφλι, τέτοιο ώστε ένα σήμα να αποστέλλεται μόνο εάν το αθροιστικό σήμα υπερβαίνει αυτό το κατώφλι.

Η ιδέα πίσω από τα νευρωνικά δίκτυα είναι ότι πολλοί νευρώνες μπορούν να ενωθούν με συνδέσμους επικοινωνίας για να πραγματοποιήσουν πολύπλοκους υπολογισμούς. Είναι σύνηθες η δομή ενός νευρωνικού δικτύου να περιγράφεται ως ένα γράφημα του οποίου οι κόμβοι είναι οι νευρώνες και κάθε (κατευθυνόμενη) ακμή στο γράφημα συνδέει την έξοδο κάποιου νευρώνα με την είσοδο ενός άλλου νευρώνα. Θα περιορίσουμε την προσοχή μας στις δομές δικτύων ανατροφοδότησης, στις οποίες το υποκείμενο γράφημα δεν περιέχει κύκλους (άκυκλο).

Ένα νευρωνικό δίκτυο τροφοδοσίας περιγράφεται από ένα κατευθυνόμενο άκυκλο γράφημα,  $G = (V, E)$  και μια συνάρτηση βάρους πάνω στα άκρα,  $w : E \rightarrow \mathbb{R}$ . Οι κόμβοι του γραφήματος αντιστοιχούν σε νευρώνες. Κάθε μεμονωμένος νευρώνας μοντελοποιείται ως μια απλή βαθμωτή συνάρτηση,  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ . Στην συνέχεια θα εστιάσουμε σε τρεις πιθανές συναρτήσεις για το  $\sigma$ : τη συνάρτηση πρόσημου,  $\sigma(a) = \text{πρόσημο}(a)$ , τη συνάρτηση κατωφλίου,  $\sigma(a) = \mathbb{1}_{[a>0]}$ , και η σιγμοειδής συνάρτηση,  $\sigma(a) = 1/(1 + \exp(-a))$ , η οποία είναι μια ομαλή προσέγγιση της συνάρτησης κατωφλίου. Η σιγμοειδής συνάρτηση ονομάζεται και «συνάρτηση ενεργοποίησης» (*activation function*) του νευρώνα. Κάθε άκρο στο γράφημα συνδέει την έξοδο κάποιου νευρώνα με την είσοδο ενός άλλου νευρώνα. Η είσοδος ενός νευρώνα λαμβάνεται ως ένα σταθμισμένο άθροισμα των εξόδων όλων των νευρώνων που συνδέονται με αυτόν, όπου η στάθμιση είναι σύμφωνα με την συνάρτηση  $w$  (*weight function*).

Για την απλοποίηση της περιγραφής του υπολογισμού που εκτελείται από το δίκτυο, υποθέτουμε περαιτέρω ότι το δίκτυο είναι οργανωμένο σε επίπεδα. Δηλαδή, το σύνολο των κόμβων μπορεί να αποσυντεθεί σε μια ένωση (μη κενών) ασύνδετων υποσυνόλων,

$$V = \bigcup_{t=0}^T V_t$$

έτσι ώστε κάθε ακμή στο  $E$  να συνδέει κάποιο κόμβο στο  $V_{t-1}$  με κάποιον κόμβο στο  $V_t$ , για κάποιο  $t \in [T]$ .

Το κάτω επίπεδο,  $V_0$ , ονομάζεται επίπεδο εισόδου (*input layer*). Περιέχει  $n+1$  νευρώνες, όπου  $n$  είναι η διάσταση του χώρου εισόδου. Για κάθε  $i \in [n]$ , η έξοδος του νευρώνα  $i$  στο  $V_0$  είναι απλώς  $x_i$ . Ο τελευταίος νευρώνας στο  $V_0$  είναι ο «σταθερός» νευρώνας, ο οποίος λαμβάνει πάντα την τιμή 1.

Συμβολίζουμε με  $v_{t,i}$  τον  $i$ -οστό νευρώνα του  $t$ -οστού layer και με  $o_{t,i}(x)$  την έξοδο του  $v_{t,i}$  όταν το δίκτυο τροφοδοτείται με το διάνυσμα εισόδου  $x$ . Επομένως, για  $i \in [n]$  έχουμε  $o_{0,i}(x) = x_i$  και για  $i = n+1$  έχουμε  $o_{0,i}(x) = 1$ . Συνεχίζοντας στον υπολογισμό ανά επίπεδο (*layer*), υποθέτουμε ότι έχουμε υπολογίσει τις εξόδους των νευρώνων στο επίπεδο  $t$ . Στη συνέχεια, μπορεί να υπολογιστούν οι εξόδοι των νευρώνων στο επίπεδο  $t+1$  ως εξής. Λαμβάνοντας ως σταθερά τα:  $v_{t+1,j} \in V_{t+1}$ , και αν:  $a_{t+1,j}(x)$  σημαίνει την είσοδο σε  $v_{t+1,j}$  όταν το δίκτυο τροφοδοτείται με το διάνυσμα εισόδου  $x$ , τότε ισχύει:

$$a_{t+1,j}(x) = \sum_{r:(v_{t,r}, v_{t+1,j}) \in E} w((v_{t,r}, v_{t+1,j})) o_{t,r}(x)$$

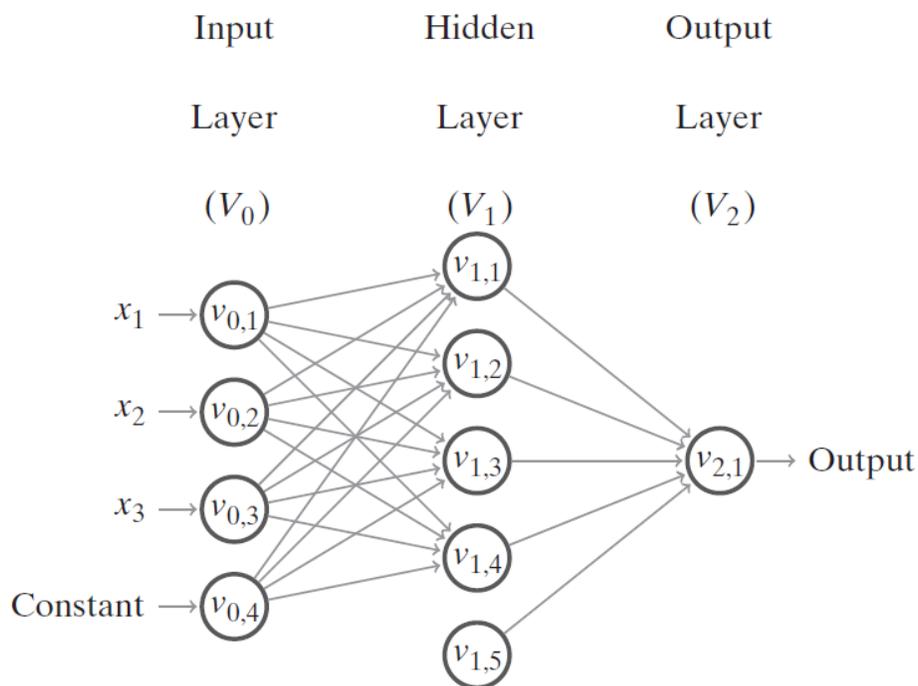
και

$$o_{t+1,j}(x) = \sigma(a_{t+1,j}(x))$$

Δηλαδή, η είσοδος στο  $v_{t+1,j}$  είναι ένα σταθμισμένο άθροισμα των εξόδων των νευρώνων σε  $V_t$  που συνδέονται με  $v_{t+1,j}$ , όπου η στάθμιση είναι σύμφωνα με την συνάρτηση βάρους  $w$  και η έξοδος του  $v_{t+1,j}$  είναι απλώς η εφαρμογή της συνάρτησης ενεργοποίησης  $\sigma$  στην είσοδό του. Τα

επίπεδα  $V_1, \dots, V_{T-1}$  ονομάζονται συχνά κρυφά επίπεδα (*hidden layers*). Το ανώτερο επίπεδο,  $V_T$ , ονομάζεται επίπεδο εξόδου (*output layer*). Σε απλά προβλήματα πρόβλεψης το επίπεδο εξόδου περιέχει έναν μόνο νευρώνα του οποίου η έξοδος είναι η έξοδος του δικτύου βλ. Εικόνα 3.1.

Το  $T$  εκφράζει τον αριθμό των επιπέδων στο δίκτυο (εξαιρουμένου του  $V_0$ ), ή το «βάθος» του δικτύου. Το μέγεθος του δικτύου είναι  $|V|$ . Το "πλάτος" του δικτύου είναι  $\max_t |V_t|$ . Μια απεικόνιση ενός πολυεπίπεδου νευρωνικού δικτύου τροφοδοσίας βάθους 2, μεγέθους 10 και πλάτους 5, δίνεται παρακάτω. Ας σημειωθεί ότι υπάρχει ένας νευρώνας στο κρυφό επίπεδο που δεν έχει εισερχόμενα άκρα. Αυτός ο νευρώνας θα εξαγάγει τη σταθερά  $\sigma(0)$  [5].



Εικόνα 3.1: Γραφική αναπαράσταση ενός νευρωνικού δικτύου

## **Βιβλιογραφία**

- [1] Cortes, C. & Vapnik, V., 1995. Support-vector networks. *Machine learning*, 20(3), pp.273–297.
- [2] Smola, A.J. and Schölkopf, B., 1998. *Learning with kernels* (Vol. 4). GMD-Forschungszentrum Informationstechnik.
- [3] Shalev-Shwartz, S. and Ben-David, S., 2014. *Understanding machine learning: From theory to algorithms*. Cambridge University Press.
- [4] Tang, Y., Zhang, Y.Q., Chawla, N.V. and Krasser, S., 2008. SVMs modeling for highly imbalanced classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1), pp.281-288.
- [5] Schmidhuber, J., 2015. Deep learning in neural networks: An overview. *Neural Networks* vol.61, pp.85-117.

## Κεφάλαιο 4<sup>ο</sup>: Περιγραφή Πειραματικής Διαδικασίας

Σε αυτή την ενότητα, περιγράφεται το υιοθετημένο σύστημα *MIMO* και η πλευρά του δικτύου του προβλήματός μας. Χρησιμοποιείται το σύστημα επικοινωνίας που εισάγεται στο *DeepMIMO Dataset*, όπου μια ομάδα χρηστών εξυπηρετείται από πολλές κεραιές ταυτόχρονα. Το μοντέλο του συστήματός θα μπορούσε εύκολα να απεικονιστεί ως ένα κανονικό περιβάλλον *MIMO* με πολλούς σταθμούς βάσης (*BS*). Θεωρείται ότι πολλοί χρήστες λαμβάνουν σήματα από πολλές κεραιές και ότι οι σταθμοί βάσης είναι εξοπλισμένοι με πολλές κεραιές ( $M$ ).

Με αυτό το μοντέλο περιβάλλοντος δημιουργούνται δεδομένα που θα χρησιμεύουν ως είσοδο για τον αλγόριθμο βελτιστοποίησης. Τα δεδομένα είναι με βάση συγκεκριμένες παραμέτρους που ορίζονται από τους δημιουργούς του και το σύνολο δεδομένων είναι ανοιχτού κώδικα και διαθέσιμο για χρήση στον κάθε χρήστη. Περιέχει δεδομένα από τρία διαφορετικά σενάρια, αλλά αυτό που ενδιαφέρει εν προκειμένω είναι αυτό στο οποίο παράγονται σήματα από σταθερά *BS*. Οι παράμετροι που ορίστηκαν από τους αρχικούς δημιουργούς του συνόλου δεδομένων χωρίζονται σε δύο κατηγορίες. Πρώτον, οι παράμετροι που περιέχει το σύστημα ως τυπικές πληροφορίες για τα κανάλια του. Οι παράμετροι για κάθε *BS* και για κάθε κανάλι είναι οι γωνίες αναχώρησης (*AoD*) από τους σταθμούς, οι γωνίες αφίξεως στον χρήστη (*AoA*), η ισχύς που λαμβάνει ο χρήστης, η φάση του μονοπατιού και η καθυστέρηση αυτού. Δεύτερον, οι παράμετροι που έχουν γεωμετρικό νόημα στο περιβάλλον μας, όπως ο αριθμός των *BS*, ο αριθμός των κεραιών  $M$  σε κάθε *BS*, το διάστημα της κεραιάς ( $d$ ), το εύρος ζώνης του συστήματος ( $B$ ), το σύνολο των υποφορέων *OFDM* ( $K$ ) που χρησιμοποιούνται για τους υπολογισμούς στα κανάλια και τον αριθμό των διαδρομών των καναλιών ( $L$ ).

Χρησιμοποιώντας αυτά τα δεδομένα και λαμβάνοντας υπόψη τις παραμέτρους, υπολογίζονται οι πληροφορίες κατάστασης καναλιού (*Channel State Information-CSI*). Κατασκευάζεται πρώτα το  $M \times N$  διάνυσμα καναλιού  $h_k^{b,u}$  για κάθε ζεύγος *BS* και για κάθε χρήστη που ο ερευνητής ενεργοποιεί. Οι υπολογισμοί βασίζονται στον αριθμό των υπο-μεταφορέων (subcarriers), χρησιμοποιώντας όλα τα δεδομένα που αναφέρθηκαν παραπάνω, και τα οποία είναι απαραίτητα για αυτούς τους υπολογισμούς, και ως εκ τούτου  $h_k^{b,u}$  δίνεται:

$$h_k^{b,u} = \sum_{l=1}^L \sqrt{\frac{\rho_l}{K}} e^{j(\theta_l^{b,u} + \frac{2\pi k}{K} \tau_l^{b,u} B)} \alpha(\varphi_{az}^{b,u}, \varphi_{el}^{b,u})$$

Από την επεξεργασία των δεδομένων αυτών, παράγονται τα αποτελέσματα του ληφθέντος σήματος. Με τη γνώση του διανύσματος καναλιού μεταξύ του *BS* και του χρήστη, που είναι η Πληροφορία Κατάστασης Καναλιού για αυτές τις δύο παραμέτρους, το σήμα που λαμβάνεται από το *BS* και πιο συγκεκριμένα στον υποφορέα  $K$  υπολογίζεται, ως ακολούθως:

$$y_k = \sum_{n=1}^N h_{k,n}^T x_{k,n} + v_k$$

όπου  $x_{k,n}$  είναι η ισχύς μετάδοσης (για παράδειγμα ισχύς εκπομπής 30dBm) και όπου προστέθηκε επίσης ο θόρυβος  $v_k$  στον υποφορέα  $K$ . Το σύνολο δεδομένων που θα χρησιμοποιηθούν, ορίζονται από ένα σενάριο ανίχνευσης ακτίνων [1], και περιέχει πληροφορίες σχετικά με τα σήματα, το εύρος ζώνης, τη γωνία άφιξης και αναχώρησης (*AoA*, *AoD*), κλπ. Σε αυτό το σύνολο δεδομένων εφαρμόζεται ένας αλγόριθμος μηχανικής μάθησης (*Machine Learning*), για την εξαγωγή της φασματικής απόδοσης ως έξοδο. Μετά την επεξεργασία, τα δεδομένα αποθηκεύονται και ταξινομούνται υπό την μορφή μήτρας. Το ίδιο το σύνολο δεδομένων περιέχει τις πληροφορίες διεκπεραίωσης του καναλιού για κάθε χρήστη (γραμμές στην μήτρα), σύμφωνα με το βιβλίο κωδικών που δημιουργήθηκε από τα *BS* (σειρές του πίνακα) [3]. Λαμβάνοντας υπόψη ότι κάθε γραμμή του πίνακα είναι ένας χρήστης, και ότι

οι στήλες είναι οι ιδιότητες του χρήστη, οι οποίες είναι προκαθορισμένες από βιβλίο κωδικών, κατασκευάζονται υποθέσεις για να προβλέψει την κατάσταση διεκπεραίωσης κάθε χρήστη και την προτεινόμενη υπηρεσία για τον χρήστη. Συνοπτικά, αρχικά, χρησιμοποιούμε ένα εποπτευόμενο μοντέλο εκμάθησης που συνδέει τον αριθμό των υποφορέων *OFDM* στους *BS* που έχουμε επιλέξει να ενεργοποιήσουμε με το δiάνυσμα διαμόρφωσης δέσμης που δημιουργείται από κάθε έναν από αυτούς. Με αυτόν τον αλγόριθμο σχηματισμού δέσμης, επιτυγχάνεται η συλλογή κάθε πληροφορία που είναι απαραίτητη για το κανάλι, ώστε να συνεχιστεί ο υπολογισμός της εξόδου για τα πειράματά. Αυτή η έξοδος είναι ο ρυθμός του καναλιού για κάθε ζεύγος *BS* και χρηστών, και ονομάζεται επίσης ως ο επιτεύξιμος ρυθμός του καναλιού. Η περιεκτικότητά του μετρείται σε bps/Hz και υπολογίζεται ως εξής: για κάθε *BS*  $n$  ο επιτεύξιμος ρυθμός του καναλιού  $R_n^{(p)}$ , με δiάνυσμα σχηματισμού δέσμης  $f_p$ , δίνεται από:

$$R_n^{(p)} = \frac{1}{|K|} \sum_{k \in |K|} \log_2 \left( 1 + SNR |f_p^T h_k^{n,u}|^2 \right)$$

Στη συνέχεια, χρησιμοποιώντας τον επιτεύξιμο ρυθμό του καναλιού, υπολογίζεται η απόδοση του καναλιού. Αυτή η διαδικασία απομονώνει το *Bandwidth* από το *Dataset* που έχει χρησιμοποιηθεί μέχρι εκείνο το σημείο. Το σύνολο δεδομένων με τον επιτεύξιμο ρυθμό του καναλιού περιέχει πληροφορίες με τη μορφή bps/Hz. Αυτό το περιεχόμενο δείχνει τη φασματική απόδοση του καναλιού. Η φασματική απόδοση συνήθως εκφράζεται ως "*bits per second per hertz*" ή bit/s/Hz. Μπορεί επίσης να οριστεί ως ο ρυθμός δεδομένων του καναλιού διαιρούμενο με το εύρος ζώνης του καναλιού σε Hz:

Φασματική απόδοση = ρυθμός δεδομένων καναλιού (σε bps)/Κανάλι Εύρος ζώνης (σε Hz).

Κατ' αρχάς δημιουργείται το εικονικό περιβάλλον και τα χαρακτηριστικά του συστήματος όπως προτείνονται από το *DeepMIMO dataset*. Καθορίζονται όλες οι απαραίτητες παράμετροι που

θα καθορίσουν το 5G σύστημα *MIMO* που προσομοιώνεται. Αναπαράγονται όλες οι είσοδοι και τα σενάρια στο MATLAB, και γίνεται η εξαγωγή διαφόρων αποτελεσμάτων και συνόλων δεδομένων των καναλιών. Το σύνολο δεδομένων (*dataset*) που θα χρησιμοποιηθεί στο πρώτο επίπεδο επεξεργασίας εξόδου είναι τα δεδομένα κατάστασης καναλιού. Αυτό το σύνολο δεδομένων χρησιμοποιείται από ένα αλγόριθμο *Machine Learning*, ο οποίος εκτελεί πιο λεπτομερείς υπολογισμούς καναλιού όπως η φασματική απόδοση του καναλιού για συγκεκριμένες συνθήκες. Στη συνέχεια, προετοιμάζεται το (*Spectral Efficiency dataset*)/ σύνολο δεδομένων φασματικής απόδοσης για να ληφθεί η απόδοση του καναλιού για το χρήστη για τις ίδιες συγκεκριμένες συνθήκες. Στη συνέχεια εκτελείται το μοντέλο βελτιστοποίησης που μελετά το σύστημά μας και παρέχει τις μεγιστοποιημένες λύσεις που βασίζονται στο σύνολο δεδομένων και τα κριτήρια διεκπεραίωσης που ορίστηκαν

Στο παρελθόν αυτό έχει δοκιμαστεί με κλασσικούς αλγόριθμους, μια επέκταση του *Hungarian algorithm* [2]. Στην παρούσα εργασία ο σκοπός είναι να χρησιμοποιηθούν νευρωνικά δίκτυα όχι με σκοπό βελτίωσης της λύσης, όπως συχνά χρησιμοποιούνται σε προβλήματα στα οποία αποτυγχάνουν οι κλασσικοί αλγόριθμοι, αλλά για μείωση της πολυπλοκότητας. Συγκεκριμένα, ο *Hungarian algorithm* είναι πολυπλοκότητας  $O(n^3)$  και σε πραγματικά massive συστήματα MIMO, κεντρικά ελεγχόμενα και για πολλές χιλιάδες χρήστες, μπορεί να δημιουργήσει πολύ μεγάλες χρονικές καθυστερήσεις. Τα προβλήματα αυτής της προσέγγισης αυξάνονται αν πάρουμε υπόψιν ότι σε αντίθεση με τις προσομοιώσεις που αναγκαστικά χρησιμοποιούνται, οι χρήστες δεν είναι στατικοί στον χώρο. Αυτό πρακτικά σημαίνει ότι ένας τέτοιος αλγόριθμος επίλυσης του γραμμικού προβλήματος ανάθεσης (*linear assignment problem*), όπως αναφέρεται στην βιβλιογραφία, θα πρέπει να εκτελείται κάθε φορά που συμβαίνουν αλλαγές θέσης ή χρηστών, κάτι το οποίο σε μια αστική περιοχή μπορεί να συμβαίνει μόνιμα. Βέβαια ο *Hungarian* εγγυάται βέλτιστη λύση. Για

προφανείς λόγους η χρήση νευρωνικών δικτύων δεν θα παράγει την βέλτιστη λύση, ωστόσο πειραματικά παρατηρήθηκαν σημαντικές βελτιώσεις στον χρόνο εκτέλεσης.

Οπότε η διαδικασία έχει ως εξής:

- Αρχικά παράγεται το *dataset*, για διαφορετικά πλήθη χρηστών, με το *DeepMimoDataset generator* [3], το οποίο επιστρέφει τα στοιχεία όλων των κεραιών των *BS* ως προς όλους του χρήστες.
- Στην συνέχεια, εκτελείτε ο *Hungarian algorithm* και βρίσκεται η βέλτιστη λύση ανάθεσης κεραιών σε χρήστες.
- Ακολούθως, τροφοδοτείται το κάθε *data point*, πλην μερικών, με την ορθή του λύση από τον *Hungarian algorithm* σε ένα *Feedforward Net* και ξεκινά η διαδικασία επιβλεπόμενης μάθησης.
- Τα εναπομείναντα *data points* χρησιμοποιούνται για να αξιολογηθεί η απόδοση του νευρωνικού δικτύου σε παραδείγματα στα οποία δεν έχει εκπαιδευτεί, παράγοντας πολύ σημαντικός για την αποφυγή *overfitting*.
- Στην συνέχεια, εφαρμόζεται το πλέον εκπαιδευμένο *Feedforward Net* και συγκρίνεται η συνολική ροή πληροφοριών που παράγει σε σχέση με την βέλτιστη λύση, καθώς και τον χρόνο εκτέλεσης του σε αντιπαράθεση αυτών του *Hungarian algorithm*.(βλ.Παράρτημα Α, *measurements and plots*)

Σε κάποια παραδείγματα, και για λιγότερους από 1000 χρήστες, παρατηρήθηκε επιτάχυνση της λύσης κατά 10-20 φορές. Είναι αναμενόμενο σε μεγαλύτερο αριθμό χρηστών να παρατηρηθεί ακόμα μεγαλύτερη χρονική διαφορά, καθώς ο *Hungarian algorithm* είναι  $O(n^3)$  πολυπλοκότητας, ενώ τα *Feedforward Neural Nets* τείνουν να επιδεικνύουν  $O(n)$  πολυπλοκότητες, χωρίς αυτό να είναι θεωρητικά αποδείξιμο. Συνεπώς, και εφόσον οι απώλειες σε συνολικό throughput είναι

περιορισμένες, είναι μια ενδιαφέρουσα προσέγγιση. Όντως, όπως παρουσιάζεται και στην εικόνα 5.1, υπολογίστηκαν απώλειες συνολικής ροής δεδομένων της τάξεως του 12%-30% με περιθώριο βελτίωσης. Τα *datasets* που δημιουργούνται είναι πολύ ογκώδη και χρονοβόρα να δημιουργηθούν σε μεγάλη κλίμακα από έναν προσωπικό υπολογιστή. Ως εκ τούτου το μοντέλο έχει εκπαιδευτεί με λιγότερα από 1000 παραδείγματα, το καθένα με διαφορετικό αριθμό χρηστών. Αν δεν υπήρχαν οι περιορισμοί χώρου και υπολογιστικής ισχύος, το επιθυμητό θα ήταν να δημιουργηθούν μερικές χιλιάδες παραδείγματα, γεγονός το οποίο θα βελτίωνε σημαντικά την ακρίβεια.

Μερικά διαδικαστικά ζητήματα όσον αφορούν την υλοποίηση του κώδικα. Ο προσομοιωτής εξάγει πίνακες με αριθμό γραμμών ίσο με το πλήθος των χρηστών και αριθμό στηλών όσο και οι επιμέρους κυψέλες των κεραιών, στην περίπτωση μας 1000. Αυτοί οι πίνακες είναι γεμάτοι με το σήμα του κάθε χρήστη προς κάθε κεραία. Δηλαδή η κάθε γραμμή  $N$  περιέχει το σήμα του  $N$ -οστού χρήστη ως προς την κεραία με αριθμο ίσο με αυτόν της στήλης, π.χ. το σήμα του  $N$ -οστού χρήστη ως προς την κεραία 2 βρίσκεται στην θέση  $(N,2)$ , προς την κεραία 3 στην θέση  $(N,3)$  κ.ο.κ. Σε αυτή την μορφή είναι έτοιμος να επεξεργαστεί από τον *Hungarian Algorithm* (βλ.Παράρτημα A *matchpairs*) όχι όμως από τα δίκτυα μας. Όντως η συνάρτηση *matchpairs* που εκτελεί τον *Hungarian Algorithm* επεξεργάζεται τον πίνακα και μας επιστρέφει την λίστα ζευγαρίων κεραιών με χρηστών που εγγυάται την μέγιστη συνολική ροή δεδομένων. Με γνώμονα αυτή την λίστα χαρτογραφούμε σε έναν νέο πίνακα άσσους για κάθε γραμμή-χρήστη στην κατάλληλη κεραία. Αυτό είναι ένα σημαντικό βήμα γιατί σε πρακτικούς όρους απλοποιεί την πληροφορία που πρέπει να μάθει το νευρωνικό δίκτυο (βλ.Παράρτημα A *preprocess*).

Χρησιμοποιήθηκε η “*scaled conjugate descent*” *backpropagation* ως “*training function*” του δικτύου που είναι κατάλληλη, δηλαδή έντονα παραλληλοποιήσιμη, για εκπαίδευση σε κάρτες

γραφικών. Επίσης έχειδειχτεί ότι είναι σημαντικά πιο γρήγορη από απλο *backpropagation* και άλλες “*conjugate gradient methods*” [4], [5].

Σε γενικές γραμμές οι κάρτες γραφικών αφιερώνουν σημαντικά περισσότερα transistors σε αριθμητικές λογικές μονάδες και σημαντικά λιγότερες σε κρυφές μνήμες και ελέγχους ροής σε σχέση με τους κεντρικούς επεξεργαστές. Επιπροσθέτως ένα μοντέρνος μη-επαγγελματικής χρήσεως επεξεργαστής μπορεί να έχει 4, η 16, ακόμα και 32 πυρήνες, ο καθένας εκ των οποίων εκτελεί σειριακά εντολές, στην περίπτωση των νευρωνικών δικτύων αυτές είναι απλές αριθμητικές πράξεις κινητής υποδιαστολής. Σε αντίθεση μια μοντέρνα κάρτα γραφικών έχει χιλιάδες τέτοιους πυρήνες οι οποίοι κατά κανόνα θα έχουν πιο μικρή συχνότητα ρολογιού (αριθμός κύκλων ανά δευτερόλεπτο, μετριέται σε Hz). Το υπολογιστικό σύστημα που χρησιμοποιήθηκε για τα πειράματα αυτά είχε επεξεργαστή 8 πυρήνων με συχνότητα ρολογιού 3.2 GHz και κάρτα γραφικών με 2432 πυρήνες με συχνότητα ρολογιού 1.607 GHz. Η εκπαίδευση των δικτύων είναι εγγενώς έντονα παραλληλοποιήσιμη διαδικασία και ως εκ τούτου βλέπει σπουδαία χρονική βελτίωση από την χρήση κάρτα γραφικών. Στο πρώτο δίκτυο που δοκιμάστηκε, *Net 1* το οποίο παρουσιάζεται παρακάτω, η διαφορά χρόνο εκπαίδευσης με τον επεξεργαστή ενάντια της κάρτας γραφικών ήταν περίπου μία τάξης μεγέθους καλύτερη για την κάρτα γραφικών. Με αυτές τις πληροφορίες κρίνεται ασύμφορη η χρήση επεξεργαστών και γίνεται χρήση κάρτας γραφικών.

## **Βιβλιογραφία**

- [1] Alkhateeb, A. ‘DeepMIMO: A Generic Deep Learning Dataset for Millimeter Wave and Massive MIMO Applications’, Proc. of Information Theory and Applications Workshop (ITA), Feb., 2019.
- [2] Tzanakos, D., Allayiotis, F., Kokkinos, V. and Bouras, C., A mechanism for 5G MIMO performance optimization and evaluation.
- [3] <http://www.DeepMIMO.net>
- [4] Moller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(4), p. 525-533. doi.org/10.1016/S0893-6080(05)80056-5.
- [5] Braun, H., Riedmiller, M., Rprop: A fast and robust backpropagation learning strategy. *Proc. of the ACNN*, 1993.

## Κεφάλαιο 5<sup>ο</sup>: Βελτιστοποίηση Ανάθεσης Πόρων με Τεχνικές Μηχανικής Μάθησης - Πειραματικά Αποτελέσματα

Διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων δοκιμάστηκαν, με διαφορετικό αριθμό νευρώνων όσο και διαφορετικά πλήθη κρυφών επίπεδο (*hidden layers*). Το υπολογιστικό κόστος εκπαίδευσης αυτών των δικτύων διαφέρει και σε γενικές γραμμές αυξάνεται γραμμικά με τον αριθμό νευρώνων και κρυφών επιπέδων. Δεν κρίνεται απαραίτητη η σύγκριση τους ως προς αυτόν τον παράγοντα καθώς είναι ένα υπολογιστικό κόστος που προκύπτει μόνο μία φορά κατά την δημιουργία του δικτύου το οποίο αντικαθιστά τους παραδοσιακούς αλγόριθμους οι οποίοι χρησιμοποιούνται για την αποτελεσματική ανάθεση χρηστών με κυψέλες κεραιών. Οι δύο μετρικές που εξετάζονται είναι:

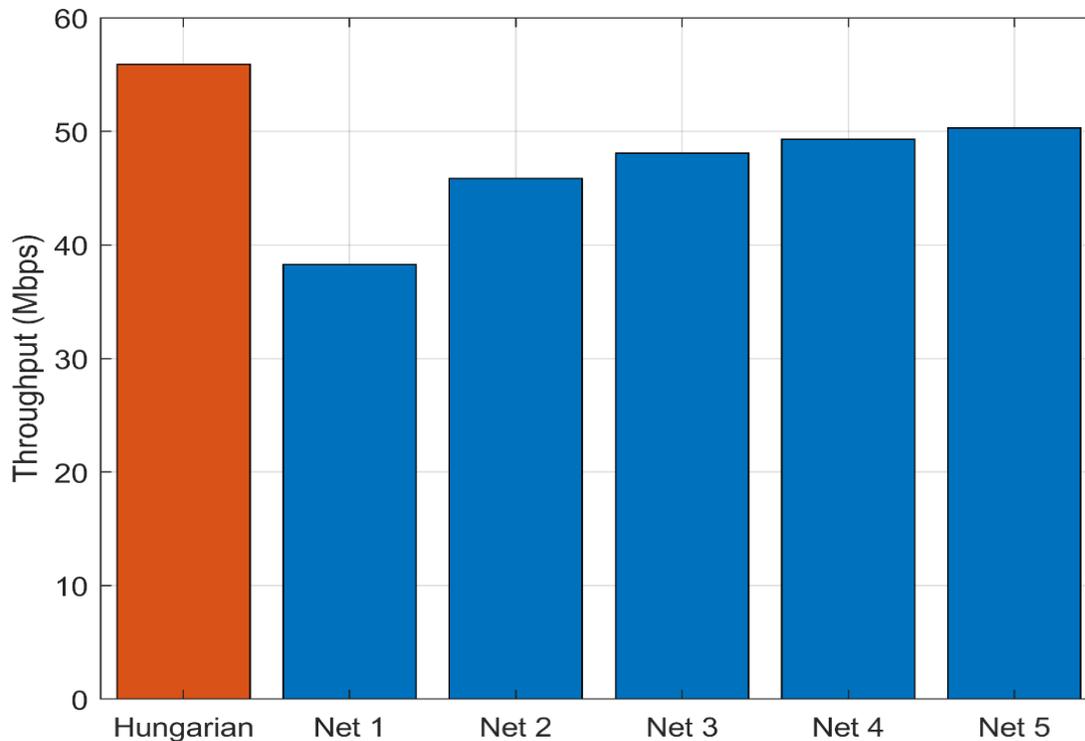
- Η αποτελεσματικότητα, η οποία ορίζεται ως ο συνολικός όγκος ροής δεδομένων (*throughput*) που επιτυγχάνεται από του χρήστες με την εκάστοτε ανάθεση χρηστών με κυψέλες κεραιών.
- Η ταχύτητα εκτέλεσης του νευρωνικού δικτύου η οποία ουσιαστικά εκφράζει το μέτρο της υπολογιστικής πολυπλοκότητας του. Σημαντικός παράγοντας είναι ότι ενώ τα δίκτυα εκπαιδεύονται με την κάρτα γραφικών η ίδια η εκτέλεση τους γίνεται μέσω του κεντρικού επεξεργαστή, όπως και του *Hungarian*, γεγονός το οποίο επιτρέπει την άμεση σύγκριση τους.

Πρώτα εξετάζονται οι σχετικές επιδόσεις των δικτύων σε αποτελεσματικότητα και ταχύτητα εκτέλεσης μεταξύ τους, αλλά και σε σχέση με τον *Hungarian*, και ακολουθεί μια σύντομη εξέταση των επιμέρους δικτύων και των αρχιτεκτονικών τους.

Στην Εικόνα 5.1 παρατηρείται ότι οι πιο απλές αρχιτεκτονικές ενός κρυφού επιπέδου, δηλαδή των δικτύων 1 και 2 έχουνε δυσκολία στο 'να μάθουν' τα χαρακτηριστικά της τοπολογίας κάτι που αποτυπώνεται στην αδυναμία τους σε σχέση με πιο περίπλοκες αρχιτεκτονικές να επιτύχουν σωστή

ανάθεση χρηστών σε κεραίες. Παρατηρείται 31% απώλεια όγκου ροής δεδομένων σε σχέση με την βέλτιστη λύση την οποία επιτυγχάνει ο *Hungarian* για το πρώτο δίκτυο και 18% για το δεύτερο δίκτυο. Η διαφορά μεταξύ του 1<sup>ου</sup> και 2<sup>ου</sup> δικτύου εξηγείται από την δραματική αύξηση νευρώνων στο 2<sup>ο</sup>, το οποίο έχει 1000 νευρώνες έναντι των 100 του πρώτου, που του επιτρέπει μια πολύ πιο ακριβή γενίκευση. Στην συνέχεια παρατηρούνται συντηρητικές βελτιώσεις με την αύξηση επιπέδων αλλά και νευρώνων. Περαιτέρω αυξήσεις επιπέδων και νευρώνων φάνηκαν να αυξάνουν την αποτελεσματικότητα σε πολύ μικρό βαθμό με το δίκτυο 5 να χρησιμοποιεί μια αρχιτεκτονική τριών κρυφών επιπέδων με 100, 150 και 100 νευρώνες να επιτυγχάνει τα βέλτιστα αποτελέσματα, στα οποία έχουμε πτώση της αποτελεσματικότητας σε σχέση με τον *Hungarian* επιπέδου 10%. Οι μικρές διαφορές σε αποτελεσματικότητα παρά την ριζική διαφοροποίηση των αρχιτεκτονικών οδηγούν στο συμπέρασμα ότι μεγαλύτερη ποσότητα δεδομένων εκπαίδευσης απαιτείται για περαιτέρω αύξηση της αποτελεσματικότητας. Για τεχνικούς λόγους ήταν ανέφικτο να πραγματοποιηθεί αυτό, καθώς τα παραδείγματα είναι του μεγέθους της τάξης του 1GB έκαστο.

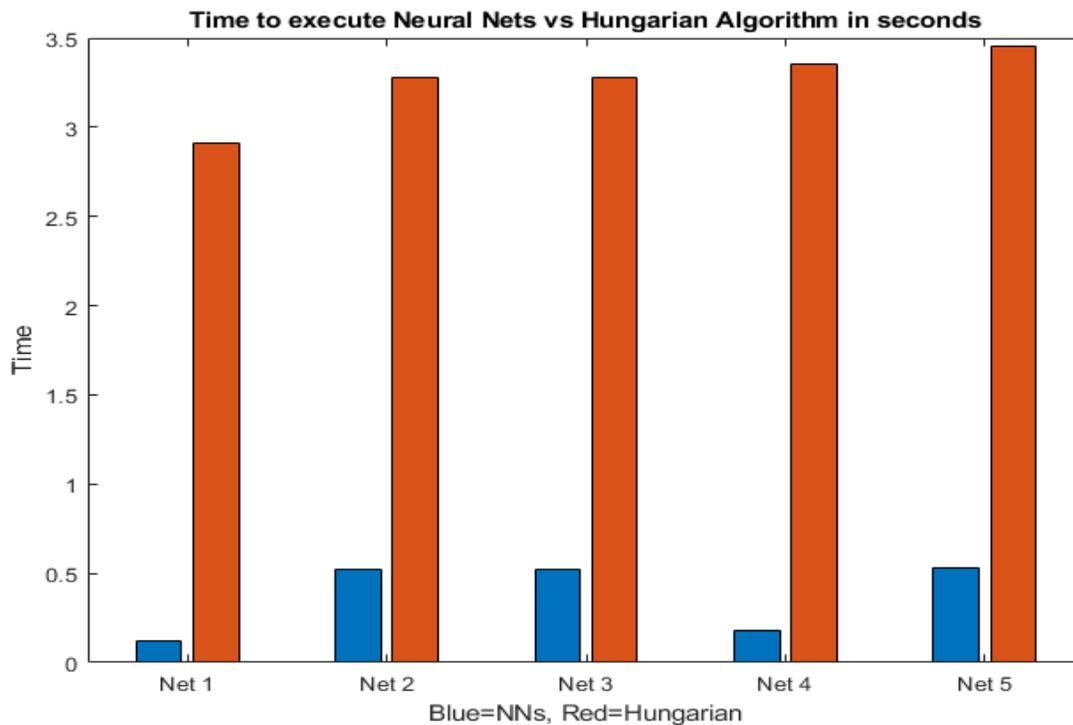
Παρατηρώντας την σχετικά μικρή αύξηση του συνολικού όγκου ροής δεδομένων καθώς αυξάνει η πολυπλοκότητα του δικτύου μπορούμε να συμπεράνουμε ότι ένα δίκτυο πολύ παρόμοιο με το 5<sup>ο</sup> έχει την απαραίτητη πολυπλοκότητα για να μάθει την τοπολογία. Βέβαια αυτό το συμπέρασμα θα ήταν πιο ασφαλές αν ήταν εφικτή η εκπαίδευση των δικτύων με σημαντικά περισσότερα παραδείγματα.



Εικόνα 5.1: Μέσος όγκος ροής δεδομένων διαθέσιμος στους χρήστες

Στην Εικόνα 5.2 γίνεται σύγκριση της ταχύτητας εκτέλεσης του *Hungarian* με την εκάστοτε αρχιτεκτονική. Οι συγκρίσεις αυτές γίνονται μεταξύ ίδιων παραδειγμάτων, ωστόσο είναι στην πραγματικότητα περιορισμένης αποδεικτικής αξίας για τις χρονικές επιδόσεις μεταξύ επιμέρους δικτύων. Αυτό φαίνεται από την διαφορά χρόνου εκτέλεσης του *Hungarian* ο οποίος έχει μεγάλη απόκλιση σε χρόνο εκτέλεσης ενώ εκτελείται σε πανομοιότυπα παραδείγματα και, προφανώς, έχει πάντα την ακριβώς ίδια υπολογιστική πολυπλοκότητα και συγκεκριμένα εκτελεί ακριβώς τις ίδιες πράξεις. Αυτό συμβαίνει διότι στα μοντέρνα λειτουργικά συστήματα και επεξεργαστές ανά πάσα χρονική στιγμή εκτελούνται διεργασίες κρυφές από τον χρήστη οι οποίες κάνουν χρήση πόρων όπως RAM, επεξεργαστικής ισχύς αλλά ακόμα και προβλέψεις αναγκών πόρων σε επίπεδο κρυφής μνήμης οι οποίες επηρεάζουν την πραγματική υπολογιστική ισχύ η οποία είναι διαθέσιμη στα προγράμματα.

Ουσιαστικά, αυτό μπορεί να αντιμετωπιστεί σαν σφάλμα μέτρησης. Η διαφορά μεταξύ *Hungarian* και των δικτύων, είναι τέτοια ώστε να μπορεί να βγει το συμπέρασμα ότι τα δίκτυα παρουσιάζουν σαφώς μικρότερη πολυπλοκότητα καθώς οι διαφορές είναι πολλές τάξεις μεγέθους μεγαλύτερες από τα σφάλματα μετρήσεων. Αυτό όμως δεν ισχύει στον ίδιο βαθμό στις επιμέρους συγκρίσεις των αρχιτεκτονικών. Τα νευρωνικά δίκτυα είναι έντονα παραλληλοποιήσιμα και στην εκτέλεση τους και όχι μόνο στην εκπαίδευσή τους. Ως εκ τούτου, υπάρχει η δυνατότητα να εκτελεστούν με κάρτες γραφικών, ωστόσο για να υπάρχει καλή σύγκριση μεταξύ *Hungarian* και δικτύων θεωρήθηκε απαραίτητο να εκτελεστούν παραδοσιακά, δηλαδή με επεξεργαστή. Αυτό γίνεται όχι μόνο για καλύτερη σύγκριση, αλλά και καθώς τα συστήματα που χρησιμοποιούνται σε συστοιχίες κεραιών κάνουν χρήση χαμηλού κόστους μικροεπεξεργαστών οι οποίοι είναι απίθανο να έχουν κάρτες γραφικών, καθώς αυτές έχουν πολύ μεγάλο κόστος και απαιτήσεις ενέργειας.



Εικόνα 5.2: Ταχύτητα εκτέλεσης

**Χαρακτηριστικά Δικτύων**

Δίκτυα	Αριθμός Επιπέδων	Αριθμός Νευρώνων	Αριθμός Βαρών
Δίκτυο 1	1	100	205924
Δίκτυο 2	1	1024	2099200
Δίκτυο 3	2	200	216024
Δίκτυο 4	3	250	169874
Δίκτυο 5	3	350	236174

**Στοιχεία Εκπαίδευσης Δικτύων**

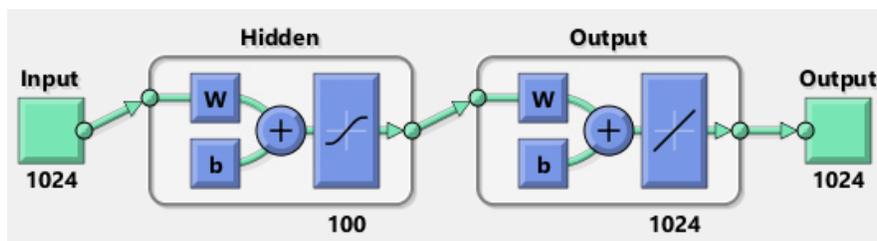
Error Histograms: είναι τα ιστογράμματα μεταξύ επιθυμητών τιμών και αυτών που προέβλεψε το νευρωνικό δίκτυο, ο άξονας X αντιπροσωπεύει το σφάλμα (γίνεται να είναι και αρνητικό καθώς είναι απλά η διαφορά μεταξύ επιθυμητών και προβλεπόμενων). Ο άξονας των Y αντιπροσωπεύει το πλήθος των παραδειγμάτων απο το *dataset*.

Validation performance graphs: εξετάζουν το μέσο τετραγωνικό σφάλμα μεταξύ των παρατηρήσεων/παραδειγμάτων και της προβλεπόμενης τιμής καθώς και την απόδοση αυτού στο validation δείγμα, γενικά το σφάλμα μειώνεται μετα απο περισσότερες εποχές εκπαίδευσης. Όταν αρχίσει να αυξάνεται αυτο συνήθως ενδεικνύει πρόβλημα overfitting, κάτι το οποίο συμβαίνει όταν τα παραδείγματα δεν είναι αρκετά διαφορετικά και οδηγούν σε αδυναμία γενίκευσης του μοντέλου. Υπάρχει σχετική προστασία απέναντι σε αυτό το πρόβλημα καθώς μετά απο 6 διαδοχικές αυξήσεις σφάλματος τερματίζει την εκπαίδευση. Εφόσον το σφάλμα συνεχίζει να μειώνεται, η εκπαίδευση

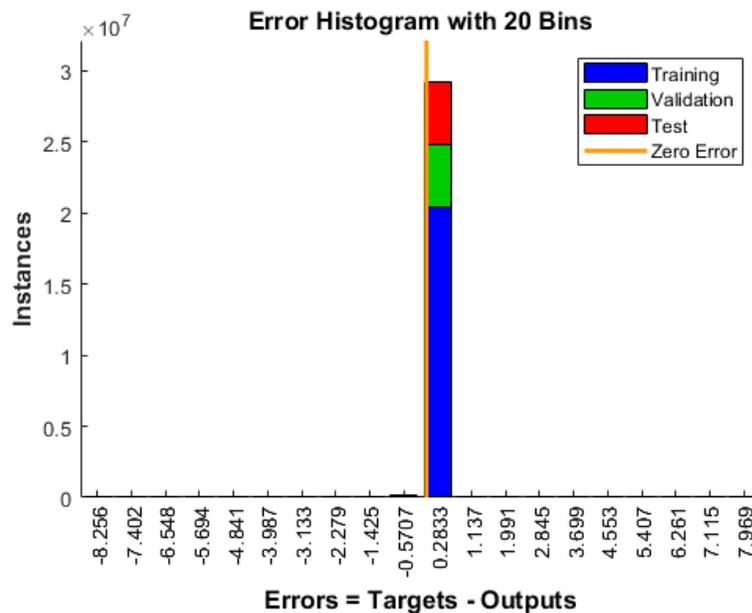
μπορεί να συνεχίσει, σχεδόν σε όλα τα παραδείγματα που ακολουθούν μπορούσε να επιτευχθεί ελαφρώ μεγαλύτερη ακρίβεια αλλά για λόγους υλοποίησης δεν συνεχίστηκε μετά τις 1000 εποχές.

Training state plots: Δίνουν την τιμή της gradient του backpropagation σε κάθε εποχή σε λογαριθμική κλίμακα, πρακτικά δείχνει με τι ταχύτητα συγκλίνει το μοντέλο προς τα τοπικά ελάχιστα της συνάρτησης ενεργοποίησης. *Validation fails* θα είχαμε μόνο σε περίπτωση *overfitting* ή *overtraining*.

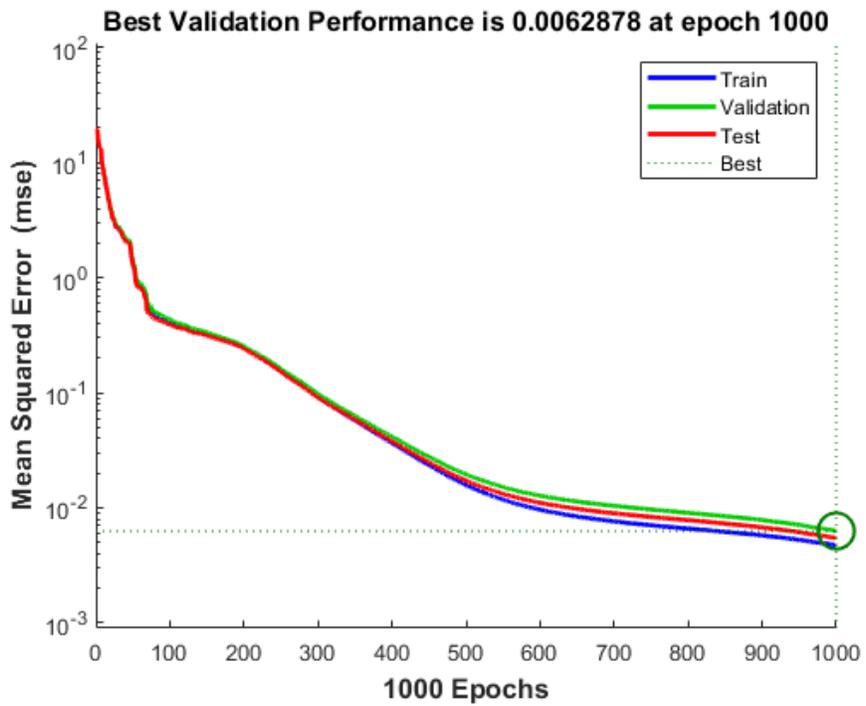
## Net 1



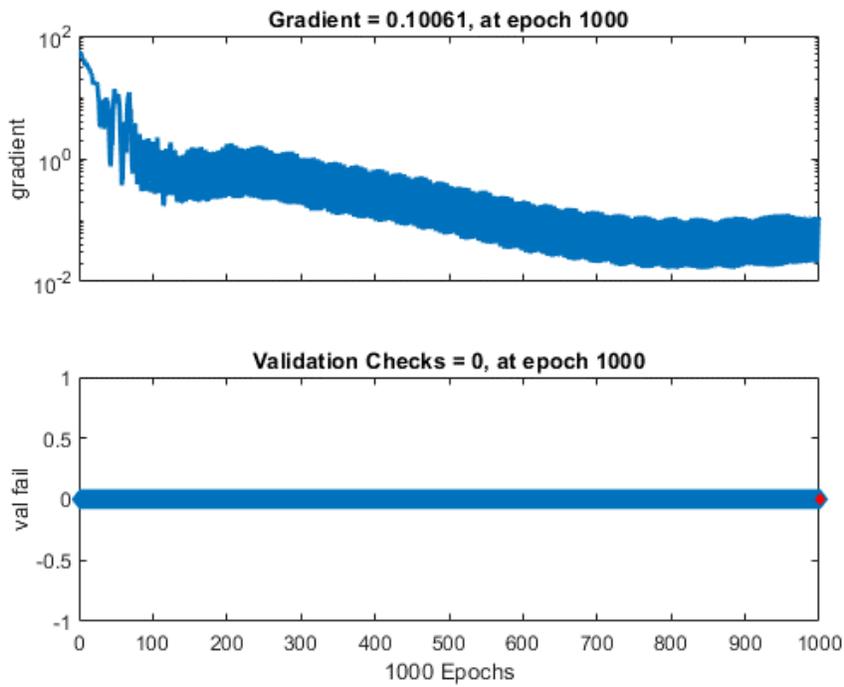
Εικόνα 5.3: Architecture of Net 1



Εικόνα 5.4: Error Histogram Net 1

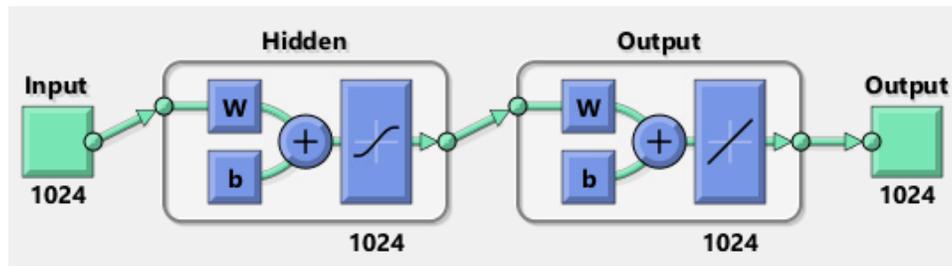


Εικόνα 5.5: Best Validation Performance Net 1



Εικόνα 5.6: Training State Plot Net 1

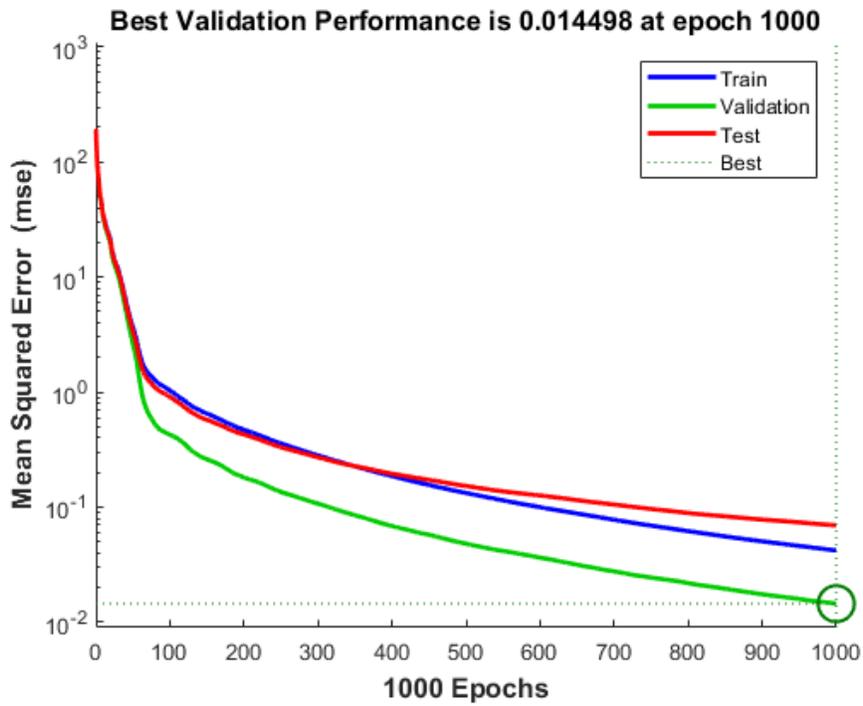
## Net 2



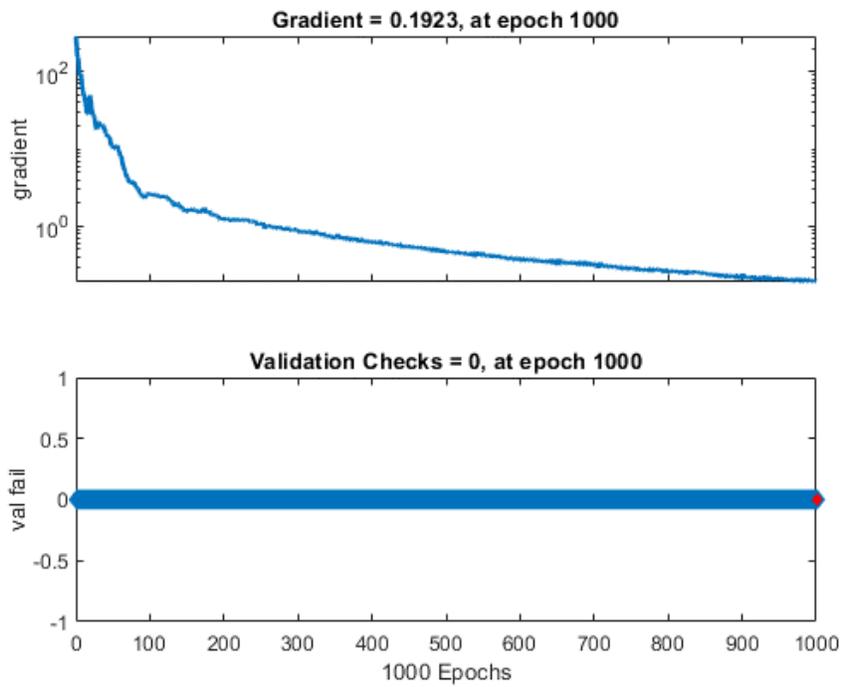
Εικόνα : 5.7: Architecture of Net 2



Εικόνα 5.8: Error Histogram Net 2

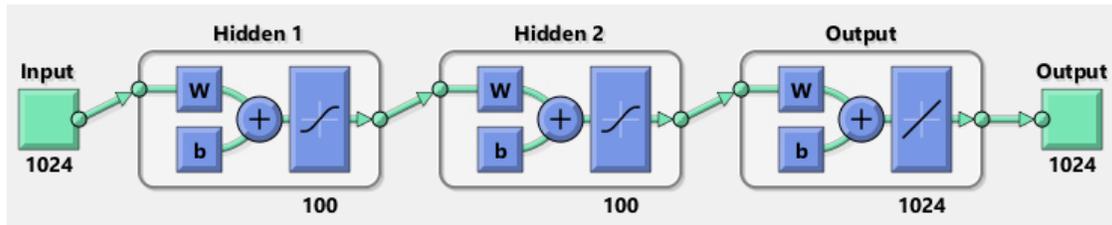


Εικόνα 5.9: Best Validation Performance Net 2

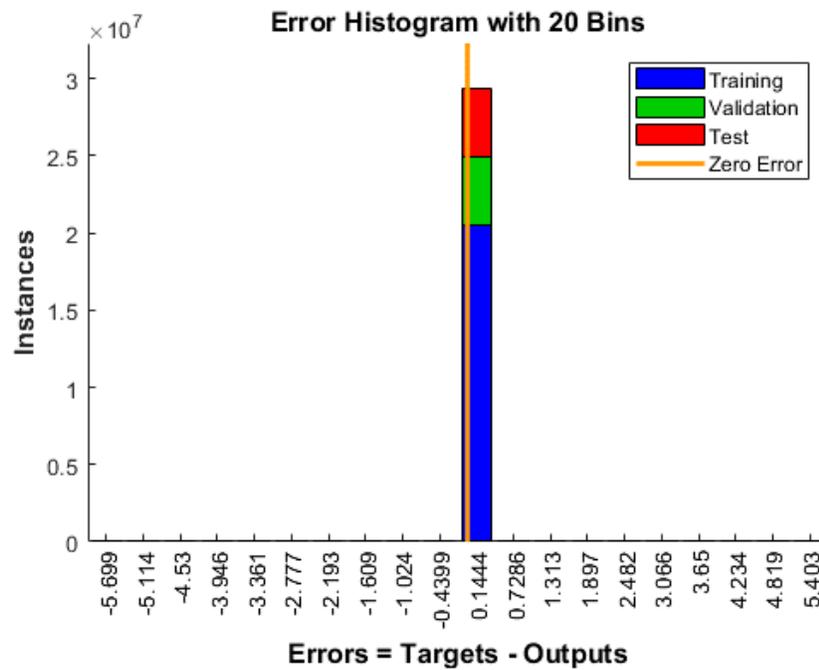


Εικόνα 5.10: Training State Plot Net 2

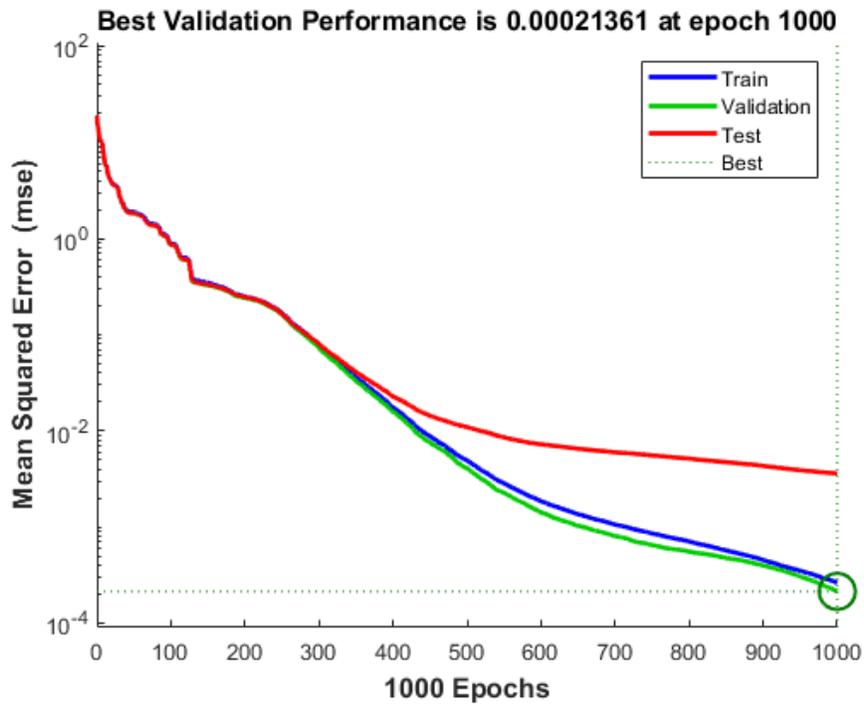
### Net 3



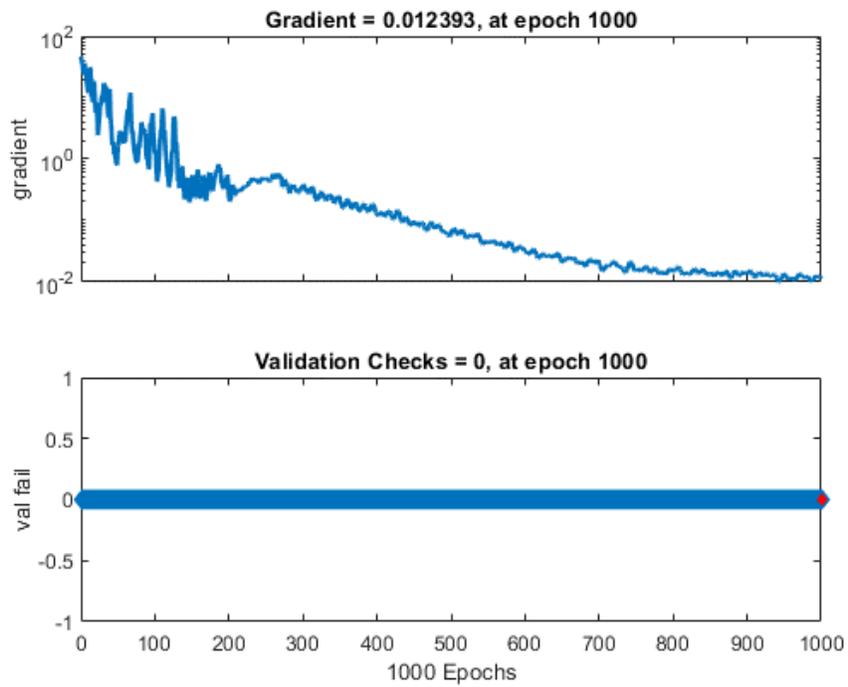
Εικόνα 5.11: Architecture of Net 3



Εικόνα 5.12: Error Histogram Net 3

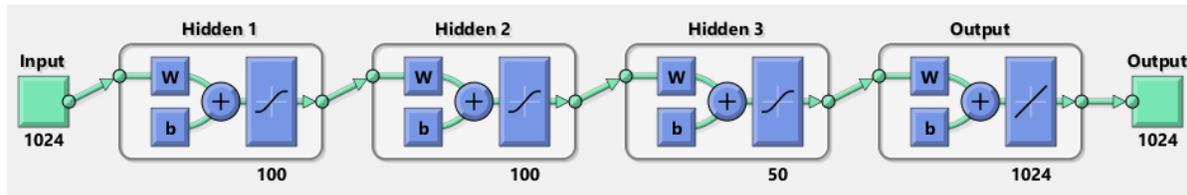


Εικόνα 5.13: Best Validation Performance Net 3

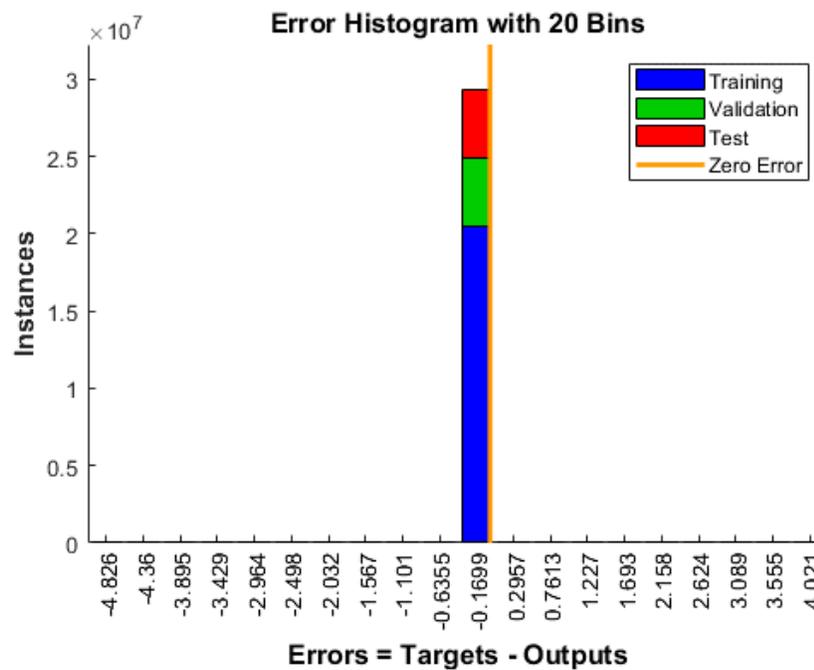


Εικόνα 5.14: Training State Plot Net 3

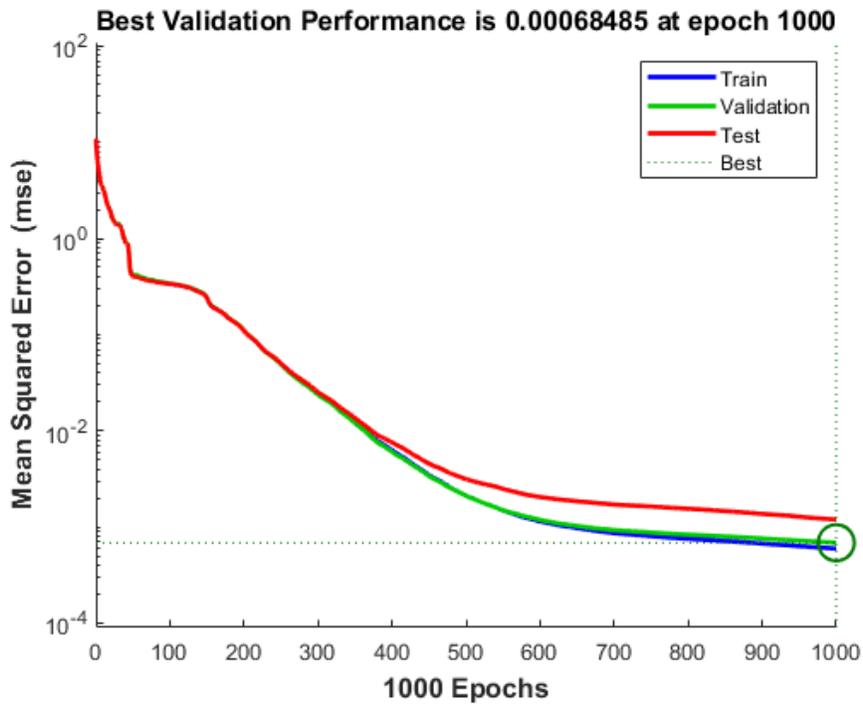
### Net 4



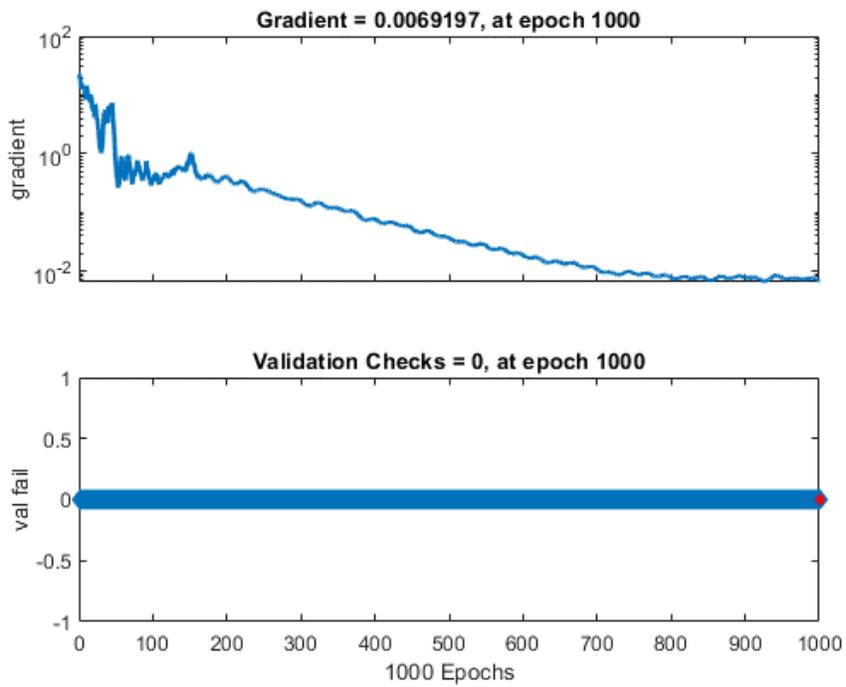
Εικόνα 5.15: Architecture of Net 4



Εικόνα 5.16: Error Histogram Net 4

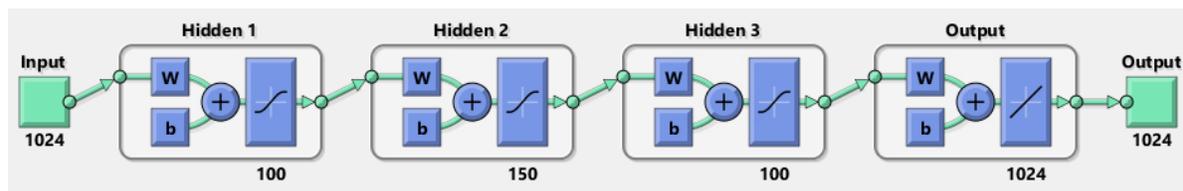


Εικόνα 5.17: Best Validation Performance Net 4

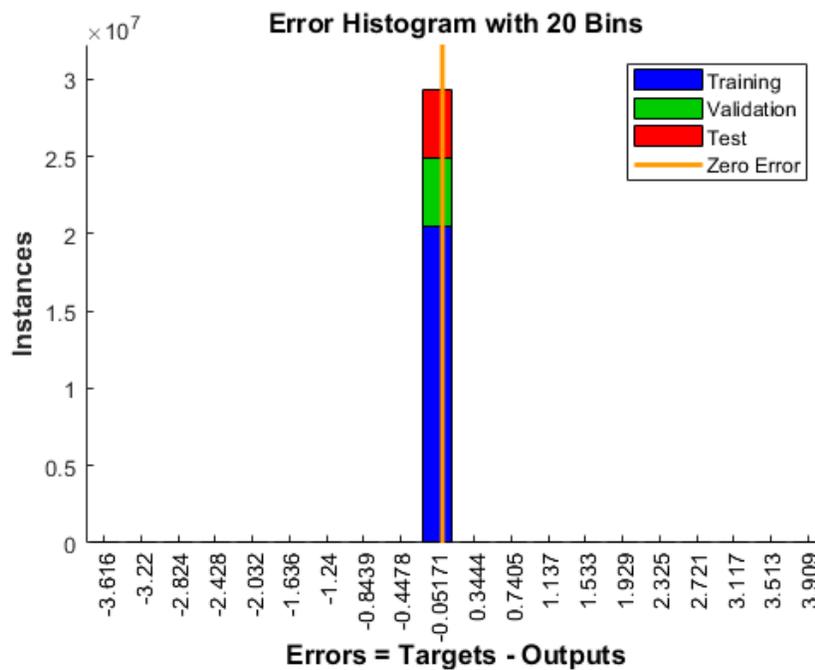


Εικόνα 5.18: Training State Plot Net 4

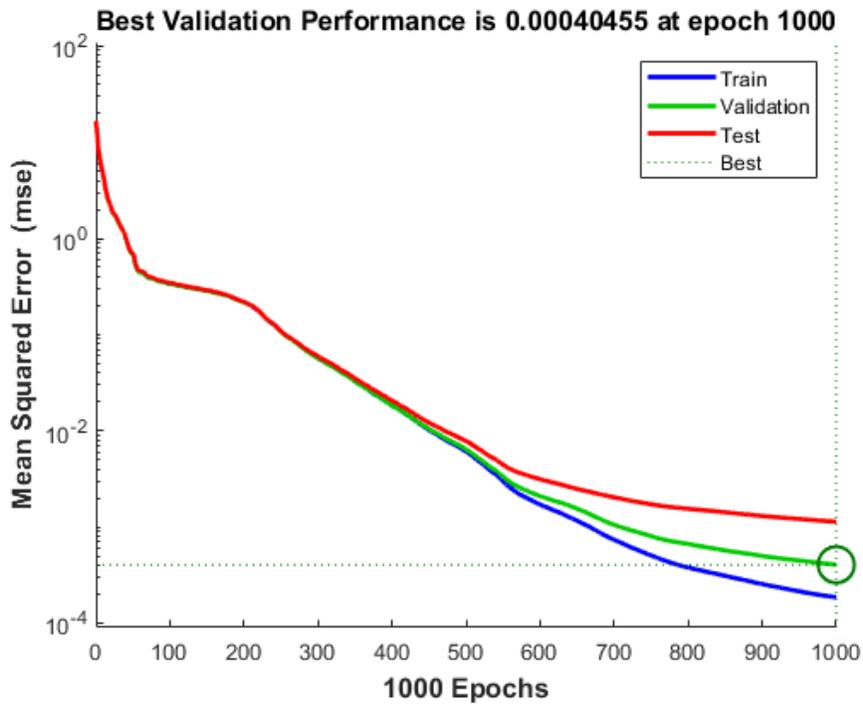
## Net 5



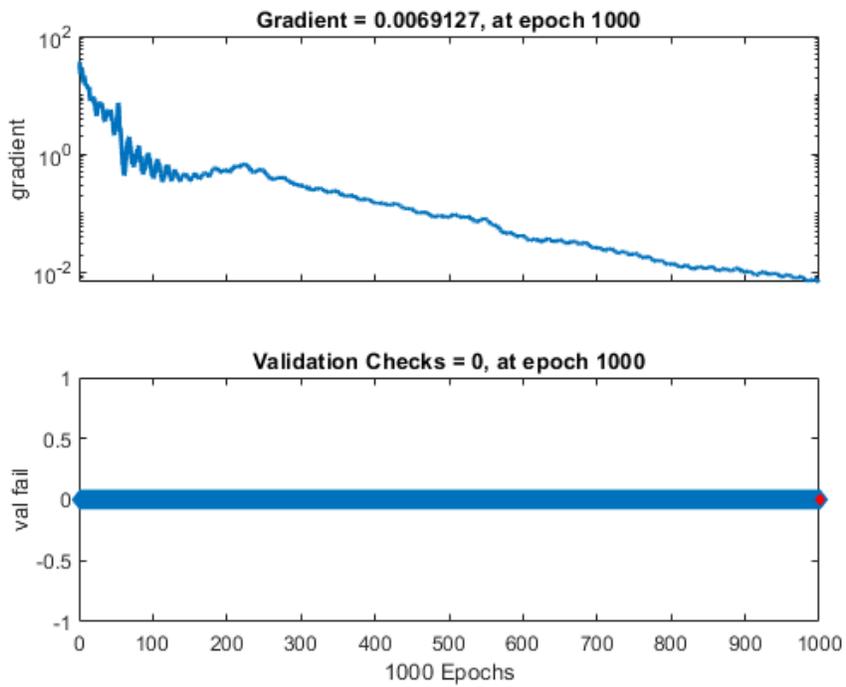
Εικόνα 5.19: Architecture of Net 5



Εικόνα 5.20: Error Histogram Net 5



Εικόνα 5.21: Best Validation Performance Net 5



Εικόνα 5.22: Training State Plot Net 5

## Κεφάλαιο 6<sup>ο</sup>: Συμπεράσματα και Μελλοντική Εργασία

Σε αυτή την ερευνητική εργασία ερευνήθηκε η λύση του προβλήματος ανάθεσης με τεχνικές μηχανικής εκμάθησης με σκοπό της μείωση της πολυπλοκότητας.

Τα αποτελέσματα που παρατηρήθηκαν είναι ενθαρρυντικά. Όπως ήταν αναμενόμενο δεν παρατηρήθηκε ακρίβεια ανάθεσης ίση με τον *Hungarian* που βρίσκει βέλτιστη λύση, ωστόσο η χρονική βελτίωση είναι αρκετά σημαντική ώστε να ενθαρρύνει την χρήση νευρωνικών δικτύων σε προβλήματα στα οποία μια ικανοποιητική αλλά γρήγορη λύση υπερτερεί μία βέλτιστης αλλά αργής λύσης.

Μελλοντική έρευνα απλά με πιο ισχυρά μηχανήματα και μεγαλύτερες δυνατότητες αποθήκευσης δεδομένων για *dataset* με πολλές χιλιάδες παραδείγματα θα ήτανε μια πολύ σημαντική βελτίωση, πρακτικά στην παρούσα εργασία το πρόβλημα που μελετήθηκε περιέχει και τον παράγοντα προσπάθειας εξαγωγής ικανοποιητικού αποτελέσματος με περιορισμένα δεδομένα. Ωστόσο λόγω της φύσεως της συλλογής δεδομένων του συγκεκριμένου προβλήματος θα ήταν μικρής δυσκολίας σε επίπεδο πειραματικό, απλή απαίτηση αποθηκευτικού χώρου και υπολογιστικής ισχύς για να παραχθούν τα 'σωστά' παραδείγματα μέσω των κλασικών αλγόριθμων. Παρομοίως και στην πραγματική περίπτωση όπου τέτοια δεδομένα θα συλλέγονταν απο συστοιχίες κεραιών σε αστικές περιοχές.

Προτείνεται έρευνα σε πιο περίπλοκες τοπολογίες για να αξιολογηθεί η ικανότητα των δικτύων να τις 'μάθει' καθώς στην συγκεκριμένη εργασία έχει χρησιμοποιηθεί μια απλή τοπολογία με 2 κεραιές, με 500 κυψέλες έκαστη και σε αμιγώς αστική περιοχή. Η υπόθεση είναι ότι μεγαλύτερος αριθμός βαρών, και κατα συνέπεια νευρώνων, θα χρειαστεί.

Εξίσου σημαντικό κρίνεται να γίνει περαιτέρω σύγκριση παραδοσιακών αλγόριθμων με νευρωνικά δίκτυα, καθώς σε μεγαλύτερους αριθμούς χρηστών αναμένεται ακόμα χειρότερη απόδοση τους, όντας πολυπλοκότητας  $O(n^3)$ , γεγονός το οποίο ευνοεί την σχετικά μικρή πολυπλοκότητα των δικτύων.

## Παράρτημα Α: Κώδικας

**Matchpairs**, η συνάρτηση που χρησιμοποιείται για την ανάθεση με αλγοριθμικό τρόπο, πρακτικά είναι μια βελτιωμένη έκδοση του *Hungarian*.

```
function [matchings, unassignedRows, unassignedCols] = ...
    matchpairs(Cost, costUnmatched, goal)

if ~isfloat(Cost) || ~isreal(Cost) || ~ismatrix(Cost)
    error(message('MATLAB:matchpairs:InvalidCost'));
elseif any(isnan(Cost), 'all')
    error(message('MATLAB:matchpairs:NonFiniteCost'));
end

if ~isfloat(costUnmatched) || ~isreal(costUnmatched) ||
~isscalar(costUnmatched)
    error(message('MATLAB:matchpairs:InvalidCostUnmatched'));
elseif ~isfinite(costUnmatched)
    error(message('MATLAB:matchpairs:NonFiniteCostUnmatched'));
end

if nargin > 2
    if ~( (ischar(goal) && isrow(goal)) || (isstring(goal) &&
isscalar(goal)) )
        error(message('MATLAB:matchpairs:InvalidOption'));
    end
    ind = strncmpi(goal, {'min', 'max'}, max(strlength(goal), 1));
    if nnz(ind) ~= 1
        error(message('MATLAB:matchpairs:InvalidOption'));
    end
    minimize = ind(1);
end
```

```

else
    minimize = true;
end

if ~minimize
    Cost = -Cost;
    costUnmatched = -costUnmatched;
end

cls = 'double';
if isa(Cost, 'single') && isa(costUnmatched, 'single')
    cls = 'single';
end

% Create a larger cost matrix, with dummy rows and columns
% to account for the possibility of not matching.
[m, n] = size(Cost);

% If costUnmatched is less than or equal to zero, the zero elements
% in C can safely be treated as if they were Inf, because they will
% never be used in a matching. This allows us to construct a sparse
% padded matrix.
useSparse = issparse(Cost) & costUnmatched <= 0;

if ~useSparse
    paddedCost = Inf(m+n, m+n, cls);

    paddedCost(1:m, 1:n) = Cost;
    paddedCost(m+1:end, n+1:end) = Cost.';
    for ii=1:m
        paddedCost(ii, n+ii) = 2*costUnmatched;
    end
    for jj=1:n

```

```

        paddedCost(m+jj, jj) = 2*costUnmatched;
    end
    [colToRow, rowToCol] =
matlab.internal.graph.perfectMatching(paddedCost);
else
    % Note: This treats structural zero as equivalent to Inf, which
is okay if
    % costUnmatched is less than or equal to zero.
    [i, j, v] = find(Cost);
    nz = nnz(Cost);
    indC = sparse(i, j, 1:nnz(Cost), m, n);

    structPaddedCost = [indC, (nz+1)*speye(m, m); (nz+1)*speye(n, n),
indC'];
    vPadded = [v; 2*costUnmatched];
    nzPaddedCost = vPadded(nonzeros(structPaddedCost));

    [colToRow, rowToCol] =
matlab.internal.graph.perfectMatching(structPaddedCost, nzPaddedCost);
end

if isempty(colToRow) && m+n > 0
    % No perfect matching found, must be because of an overflow
inside
    % perfectMatching. Rescale the matrix and try again.
    if ~useSparse
        maxVal = max(abs(paddedCost(isfinite(paddedCost))), [],
'all');
    else
        maxVal = max(abs(nzPaddedCost(isfinite(nzPaddedCost))));
    end
end

```

```

scale = 2^-nextpow2(maxVal);

if ~useSparse
    [colToRow, rowToCol] =
matlab.internal.graph.perfectMatching(paddedCost * scale);
else
    [colToRow, rowToCol] =
matlab.internal.graph.perfectMatching(structPaddedCost,
nzPaddedCost*scale);
end
end
% r = colToRow(c) gives row r matched to column c.
% c = rowToCol(r) gives column c matched to row r.

% Check for ambiguous cases, where a pair can be matched or unmatched
% with no effect on the goal function. Set these cases to be
unmatched:
for c=1:n
    r = colToRow(c);
    if r <= m && Cost(r, c) == 2*costUnmatched
        colToRow(c) = m+n+1;
        rowToCol(r) = m+n+1;
    end
end

% Real row matched to real column: a matched pair in matrix C.
matchedPairs = colToRow(1:n) <= m;
matchings = [colToRow(matchedPairs), find(matchedPairs)];
matchings = reshape(matchings, [], 2); % Correct size in empty case.

if nargout >= 2
    % Real row matched to dummy column: an unassigned row.
    unassignedRows = find(rowToCol(1:m) > n);

```

```

        unassignedRows = unassignedRows(:);
end

if narginout >= 3
    % Real column matched to dummy row: an unassigned column.
    unassignedCols = find(colToRow(1:n) > m);
    unassignedCols = unassignedCols(:);
end

```

## Preprocess δεδομένων

```


```

%preprocess
clear variables

examples=50;
% read files
matfiles = dir('*.mat') ;
N = length(matfiles) ;
iwant = cell(N,1) ; % to save output
for i = 1:N

    iwant{i} = zeros(1024,1024);
    temp=load(matfiles(i).name, 'DL_output');
    iwant{i}(1:size(temp.DL_output, 1), :) = temp.DL_output; % do
what you want, let out put be out
end

```


```

```

desiredM=cell(size(iwant));

for i=1:size(iwant)
    rng(i);
    %size(input{i})

    input{i}=iwant{i};

    costUnmatched = 0; %always 0

    [M,uR,uC] = matchpairs(input{i}, costUnmatched, 'max');
    expected{i} = M;
    %aristero column xrhstes, dexi keraies
    expectedsorted{i}=sortrows(expected{i});

    desiredM{i,1}=zeros(1024,1024);

    tic

    for sot=1:size(input{i},1)

        if sot<=size(expectedsorted{i})
            desiredM{i}(sot,expectedsorted{i}(sot,2))=1;
        end

        %desiredM{1}(1,1024)=0;
        %desiredMtest(sot,sortedMtest(sot,2))=1; alternate way
    end

```

```

end
toc

%sum=input{i}(M(t,1),M(1,t); FIND

save('dato.txt', 'DL_output_band');
temp = matfile('dato.txt');
filename =
'C:\Users\mkour\Desktop\diplomatikh\DeepMIMO_Dataset_Generation_v1.1\
data1.mat';
save(filename, 'DL_output_band');
data = readmatrix('dato.txt');
temp=load('data1.mat');
data1=temp.DL_output_band;
DL_output_band= temp.DL_output_band;
costUnmatched = 0;

```

## Net Setup

```

mu=0.001;
mu_dec=0.2;
neuron_number=[100 150 100];
neurons=num2str(neuron_number);
layer_sizes = [neuron_number];
net = feedforwardnet(layer_sizes);
net = configure(net,input,desiredM'); % desiredM, target
net.layers{1}.transferFcn = 'tansig'
net.layers{end}.size = 1024;

```

```

net.layers{end}.dimensions = 1024;
net.layers{end}.transferFcn = 'purelin'
net.trainParam.mu=mu;
net.trainParam.mu_dec=mu_dec;
net.output.processFcns = {'mapminmax'};
net.input.processFcns = {'mapminmax'};
[net, tr] = train(net,input,desiredM', 'useGPU', 'yes');

```

## Measurements

```

inputtest(1:numusers,:) = DL_output;
tic
y = net(inputtest);
nettime = toc
y = y(1:numusers,:);
isequal(test, Mtest);
%simtest = 100 * numel(find(test == Mtest)) / numel(ytest)
[Q, k] = max(y, [], 2, 'linear');
inputmid = DL_output(k);
tic
[Minput, uR, uC] = matchpairs(DL_output, costUnmatched, 'max');
Minput = sortrows(Minput);
hungtime = toc
Minputlin = sub2ind(size(DL_output), Minput(:,1), Minput(:,2));
inputmidhung = DL_output(Minputlin);
meanband = mean(inputmid);
meanbandhung = mean(inputmidhung);

```

## Throughput Plots

```

close all
y = [meanbandexhibit meanbandnet1 meanbandnet2 meanbandnet3
meanbandnet4 meanbandnet5]
X = categorical({'Hungarian','Net 1','Net 2','Net 3', 'Net 4', 'Net
5'});
X = reordercats(X,{'Hungarian','Net 1','Net 2','Net 3', 'Net 4', 'Net
5'});
b=bar(X,y, 'FaceColor', 'flat' , 'edgecolor', 'black' );
hold on
%b.FaceColor = 'flat';
b.CData(1,:) = [0.8500, 0.3250, 0.0980];
b.CData(2,:) = [0, 0.4470, 0.7410];
b.CData(3,:) = [0, 0.4470, 0.7410];
b.CData(4,:) = [0, 0.4470, 0.7410];
b.CData(5,:) = [0, 0.4470, 0.7410];
b.CData(6,:) = [0, 0.4470, 0.7410];
%title('Total throughput afforded to users')
%xlabel('Blue=NNs, Red=Hungarian')
ylabel('Throughput (Mbps)')
grid on
%l=cell(1,6);
%l{1}='Hungarian'; l{2}='Neural Net 1'; l{2}='Neural Net
2';l{3}='Neural Net 4';l{4}='Neural Net 5'; l{6}='Neural Net 6';
%legend(b,l);

print('bars.png', '-dpng', '-r400');

```

## Time plots

```
y = [net2time net2hungtime; net3time net3hungtime; net4time
net4hungtime; net5time net5hungtime; net6time net6hungtime]
X = categorical({'Net 1','Net 2','Net 3','Net 4', 'Net 5'});
X = reordercats(X,{'Net 1','Net 2','Net 3','Net 4', 'Net 5'});
bar(X,y)
title('Time to execute Neural Nets vs Hungarian Algorithm in seconds')
xlabel('Blue=NNs, Red=Hungarian')
ylabel('Time')
```