

Πολυτεχνική Σχολή

Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Διπλωματική Εργασία

**Εφαρμογή τεχνικών μη εποπτευόμενης**

**μηχανικής μάθησης στην τεχνολογία MIMO για τη βελτίωση της ανάθεσης πόρων σε κινητά δίκτυα επικοινωνιών.**

Πρόδρομος Νικόλαος

Α.Μ. 1072549

Πάτρα, 2024



Πολυτεχνική Σχολή

Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Διπλωματική Εργασία

**Εφαρμογή τεχνικών μη εποπτευόμενης**

**μηχανικής μάθησης στην τεχνολογία MIMO για τη βελτίωση της ανάθεσης πόρων σε κινητά δίκτυα επικοινωνιών.**

Πρόδρομος Νικόλαος

Α.Μ. 1072549

Επιβλέπων

Χρήστος Μπούρας, Καθηγητής

Μέλη Τριμελούς Συμβουλευτικής/Εξεταστικής Επιτροπής

Ιωάννης Γαροφαλάκης, Καθηγητής

Εύη Παπαϊωάννου, Επίκουρη Καθηγήτρια

Πάτρα, 2024

© Copyright συγγραφής Πρόδρομος Νικόλαος, 2024

© Copyright θέματος Χρήστος Μπούρας, Βασίλειος Κόκκινος

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών & Πληροφορικής του Πανεπιστημίου Πατρών δεν υποδηλώνει απαραιτήτως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω θερμά τον υπεύθυνο καθηγητή της διπλωματικής μου εργασίας και Πρύτανη του Πανεπιστημίου Πατρών κύριο Χρήστο Μπούρα για την βοήθεια που μου παρείχε καθ’ όλη τη διάρκεια της εκπόνησης της εργασίας.

Επίσης θα ήθελα να εκφράσω τις ευχαριστίες μου στον Δρ. Βασίλειο Κόκκινο για την καθοδήγηση που μου παρείχε, βοηθώντας με σημαντικά στην ολοκλήρωση αυτής της εργασίας με τις συμβουλές και τις υποδείξεις του.

**Περίληψη**

Η εμφάνιση της τεχνολογίας 5G σηματοδοτεί μια καθοριστική αλλαγή στο τοπίο των τηλεπικοινωνιών, υποσχόμενη γρήγορες ταχύτητες και συνδεσιμότητα. Ωστόσο, η ανάπτυξη και η συντήρηση των δικτύων 5G, ιδιαίτερα σε περιβάλλοντα Πολλαπλών Εισόδων Πολλαπλών Εξόδων (Multiple Input Multiple Output - MIMO), θέτουν σημαντικές προκλήσεις στην κατανομή των πόρων. Αυτή η διπλωματική εργασία προτείνει μια καινοτόμο προσέγγιση για τη βελτιστοποίηση της κατανομής πόρων σε περιβάλλοντα 5G MIMO μέσω αλγορίθμων μη επιβλεπόμενης μηχανικής μάθησης (Machine Learning – ML). Στόχος της έρευνας, είναι η εφαρμογή τέτοιων τεχνικών και πιο συγκεκριμένα συσταδοποίησης (clustering), αλλά και χρήση της μείωσης διαστατικότητας (dimensionality reduction), σε ομάδες χρηστών με διάφορες κατανομές, η σύγκριση των αποτελεσμάτων των τεχνικών αυτών καθώς και η χρήση τους για τη δημιουργία ενός αλγορίθμου αποτελεσματικής κατανομής πόρων. Σκοπός είναι η ευκολότερη κατανόηση και διαχείρισή των δεδομένων των χρηστών και στη συνέχεια η εξασφάλιση βέλτιστης κατανομής των πόρων του δικτύου στους χρήστες μέσω της ανάπτυξης ενός νέου αλγορίθμου κατανομής πόρων.

**Abstract**

The emergence of 5G technology signifies a pivotal change in the telecommunications landscape, promising unprecedented speeds and connectivity. However, the development and maintenance of 5G networks, especially in Multiple Input Multiple Output (MIMO) environments, pose significant challenges in resource allocation. This thesis proposes an innovative approach for optimizing resource allocation in 5G MIMO environments through unsupervised learning algorithms. The research aims to apply these techniques which involve clustering as well as dimensionality reduction, on user groups with various distributions, compare the results of the algorithms and use them to make an effective resource allocation algorithm. The goal is to facilitate the easier understanding and management of each user’s data and subsequently ensure the optimal allocation of network resources to users through developing a new resource allocation algorithm.

**Περιεχόμενα**

1 7

Εισαγωγή 7

1.1 Σημασία του προβλήματος 7

1.2 Στόχοι της εργασίας 9

1.3 Μεθοδολογία προσέγγισης 9

1.4 Συνεισφορά 10

1.5 Διάρθρωση της υπόλοιπης διπλωματικής εργασίας 11

2 12

Δίκτυα κινητής τηλεφωνίας 5ης γενιάς 12

2.1 Εισαγωγή στην 5η γενιά των δικτύων κινητής τηλεφωνίας 12

2.2 Αρχιτεκτονική του δικτύου 5G 13

2.2.1 Τα επίπεδα της αρχιτεκτονικής 5G 14

2.2.2 Φάσμα δικτύων 5G 16

2.3 Νέες τεχνολογίες στο 5G 16

2.3.1 Διαμόρφωση δέσμης 16

2.3.2 Massive MIMO 17

2.3.3 mmWave 17

2.3.4 Διαχωρισμός δικτύου 17

2.3.5 Βελτιωμένη κινητή ευρυζωνική σύνδεση (Enhanced Mobile Broadband - eMBB) 18

2.3.6 Επικοινωνίες υπεραξιόπιστες και με χαμηλή καθυστέρηση (Ultra Reliable Low Latency Communications - URLLC) 18

2.3.7 Μαζικές επικοινωνίες τύπου μηχανής (Massive Machine-Type Communications - mMTC) 19

2.3.8 Συγχώνευση φορέων 19

2.3.9 Επικοινωνία πλήρους διπλής κατεύθυνσης 19

2.3.10 Επικοινωνία συσκευής προς συσκευή (Device-to-Device - D2D) 20

3 21

Η κατανομή πόρων στα δίκτυα 5G 21

3.1 Η ανάγκη για κατανομή πόρων στο 5G MIMO 21

3.2 Η έννοια του RB 21

3.3 Τεχνικές κατανομής πόρων στα δίκτυα 5G 22

3.4 Προκλήσεις στην κατανομή πόρων στα δίκτυα 5G MIMO 23

4 25

ML στα δίκτυα 5G 25

4.1 Η ανάγκη της ML στο 5G MIMO 25

4.2 Τεχνικές κατανομής πόρων με χρήση ML 26

4.3 Η μη επιβλεπόμενη ML στην κατανομή πόρων 27

4.4 Οι προκλήσεις στη χρήση μη επιβλεπόμενης ML στο 5G MIMO 28

5 30

Πειραματική διαδικασία 30

5.1 Εισαγωγή στην πειραματική διαδικασία 30

5.2 Φυσική σημασία των χαρακτηριστικών 31

5.3 Οι υπηρεσίες και τα χαρακτηριστικά τους 32

5.4 Περιγραφή του περιβάλλοντος και η μηχανική των χαρακτηριστικών 33

5.5 Προεπεξεργασία των χαρακτηριστικών 38

5.6 Η σχέση των χαρακτηριστικών 39

5.7 Οι αλγόριθμοι που χρησιμοποιήθηκαν 41

5.8 PCA, scaling και βελτιστοποίηση των αλγορίθμων 46

5.8.1 Αρχική εκτίμηση συστάδων με χρήση δενδρογράμματος και HAC 46

5.8.2 Βελτιστοποίηση του αλγορίθμου K-Means 47

*5.8.2.1* *Άθροισμα τετραγώνων εντός συστάδας (Within Cluster Sum of Squares - WCSS)* 47

*5.8.2.2* *Elbow point* 48

*5.8.2.3* *Silhouette score* 49

5.8.3 Βελτιστοποίηση του αλγορίθμου GMM 50

*5.8.3.1* *BIC και AIC scores* 51

5.8.4 Βελτιστοποίηση παραμέτρων του DBSCAN 52

5.8.5 Silhouette score για το spectral clustering 53

5.8.6 Δείκτης CH 54

5.8.7 Αποτέλεσμα CH score 55

5.8.8 Τελική επιλογή αλγορίθμου συσταδοποίησης 55

6 57

Αποτελέσματα των πειραμάτων 57

6.1 Τα συμπεράσματα της συσταδοποίησης 57

6.2 Αποτελέσματα και σύγκριση των αλγορίθμων κατανομής πόρων 60

6.3 Συμπεράσματα πειραμάτων 64

6.4 Περιπτώσεις χρήσης των συμπερασμάτων της συσταδοποίησης στο 5G 65

6.5 Μελλοντική εργασία 66

**Λίστα Εικόνων**

**Εικόνα 1:** Τα επίπεδα αρχιτεκτονικής των επιπέδων χρηστών και ελέγχου του δικτύου 5G 15

**Εικόνα 2:** Η δομή ενός πλαισίου πόρων και ενός RB 22

**Εικόνα 3:** Γράφημα της περιοχής του πειράματος, των κεραιών και των χρηστών 34

**Εικόνα 4:** Κατανομή του SNR των χρηστών της κεραίας 1 37

**Εικόνα 5:** Κατανομή του SNR των χρηστών της κεραίας 2 37

**Εικόνα 6:** Γράφημα (scatterplot) των στοιχείων των χρηστών μετά το PCA 2 διαστάσεων 38

**Εικόνα 7:** Γράφημα (heatmap) συσχετίσεων των χαρακτηριστικών των χρηστών 39

**Εικόνα 8:** Δενδρόγραμμα HAC 46

**Εικόνα 9:** WCSS score για αριθμό συστάδων 3 με 8 48

**Εικόνα 10:** Silhouette score K-Means για αριθμό συστάδων 3 με 7 49

**Εικόνα 11:** BIC Και AIC scores του GMM για αριθμό συστάδων 3 με 8 50

**Εικόνα 12:** Silhouette score spectral clustering για αριθμό συστάδων 3 με 8 52

**Εικόνα 13:** Μαθηματικός τύπος για τον δείκτη CH 53

**Εικόνα 14:** Αναλυτικό point-by-point silhouette score για τον αλγόριθμο K-Means 54

**Εικόνα 15:** Γράφημα (boxplot) απαιτήσεων των χρηστών ανάλογα με το cluster τους 56

**Εικόνα 16:** Γράφημα κοντινότερων κεραιών των χρηστών ανάλογα με το cluster τους 57

**Εικόνα 17:** Γράφημα (heatmap) των μέσων τιμών των χαρακτηριστικών των χρηστών ανά cluster 57

**Εικόνα 18:** Γράφημα των τοποθεσιών των χρηστών ανάλογα με το cluster τους 58

**Εικόνα 19:** Ποσοστό RBs που χρησιμοποιήθηκαν από κάθε cluster στον δυναμικό αλγόριθμο 59

**Εικόνα 20:** Κατανομή των απαιτήσεων των χρηστών στις κεραίες στον απλό αλγόριθμο 60

**Εικόνα 21:** Κατανομή των απαιτήσεων των χρηστών στις κεραίες στον δυναμικό αλγόριθμο 60

**Εικόνα 22:** Κατανομή των υπηρεσιών των χρηστών στις κεραίες στον απλό αλγόριθμο 61

**Εικόνα 23:** Κατανομή των υπηρεσιών των χρηστών στις κεραίες στον δυναμικό αλγόριθμο 61

**Εικόνα 24:** Ποσοστό χρηστών που εξυπηρετήθηκαν σε κάθε αλγόριθμο (δυναμικός αλγόριθμος - απλός αλγόριθμος) 62

**Λίστα Πινάκων**

**Πίνακας 1:** Αποτελέσματα WCSS αλγορίθμου K-Means 47

**Πίνακας 2:** Αποτελέσματα silhouette score αλγορίθμου K-Means 49

**Πίνακας 3:** Αποτελέσματα BIC και AIC score αλγορίθμου GMM 50

**Πίνακας 4:** Αποτελέσματα αλγορίθμου DBSCAN για διαφορετικές τιμές του eps 51

**Πίνακας 5:** Αποτελέσματα silhouette score αλγορίθμου spectral clustering 52

**Πίνακας 6:** Αποτελέσματα δείκτη CH για κάθε αλγόριθμο (μετά τη βελτιστοποίηση) 53

**Συντομογραφίες**

|  |  |  |  |
| --- | --- | --- | --- |
| 5G | 5th Generation / Ασύρματα Συστήματα 5ης Γενιάς | | |
| 5GC | | 5G Core Network / Δίκτυο Πυρήνα 5ης Γενιάς | | |
| 5GNR | | 5G New Radio / Νέο Ραδιοδίκτυο 5ης Γενιάς | | |
| 6G | | 6th Generation / Ασύρματα Συστήματα 6ης Γενιάς | | |
| AIC | | Akaike Information Criterion / Κριτήριο Πληροφορίας του Akaike | | |
| AMF | | Access and Mobility Function/Διαχείριση Πρόσβασης και Κινητικότητας | | |
| AR | | Augmented Reality/Επαυξημένη Πραγματικότητα | | |
| B5G | | Beyond 5th Generation/Πέρα από την 5η Γενιά Ασύρματων Δικτύων | | |
| BIC | | Bayesian Information Criterion / Κριτήριο Πληροφορίας του Bayes | | |
| Bps | | Bits per Second | | |
| CH | | Calinsky-Harabasz | | |
| CoMP | | Coordinated MultiPoint / Συντονισμένη Πολυσημειακή Τεχνική | | |
| CRAN | | Cloud Radio Access Network | | |
| CSI | | Channel State Information / Πληροφορίες Κατάστασης Καναλιού | | |
| D2D | | Device-to-Device / Επικοινωνία Συσκευής προς Συσκευή | | |
| DBSCAN | | Density-Based Spatial Clustering of Applications with Noise / Χωρική Ομαδοποίηση Εφαρμογών με Θόρυβο με βάση την Πυκνότητα | | |
| DSA | | Dynamic Spectrum Access / Δυναμική Πρόσβαση Στο Φάσμα | | |
| eCPRI | | Evolved Common Public Radio Interface | | |
| EM | | Expectation Maximization /Αλγόριθμος Προσδοκίας-Μεγιστοποίησης | | |
| eMBB | | Enhanced Mobile Broadband / Βελτιωμένη Κινητή Ευρυζωνική Σύνδεση | | |
| Gbps | | Gigabit per Second | | |
| GMM | | Gaussian Mixture Model / Γκαουσιανό Μοντέλο Μείξης | | |
| HAC | | Hierarchical Agglomerative Clustering / Ιεραρχική Συσσωματωτική Συσταδοποίηση | | |
| HC | | Hierarchical Clustering / Ιεραρχική Συσταδοποίηση | | |
| IoT | | | Internet of Things / Διαδίκτυο των Πραγμάτων | |
| KHz | | Kilohertz | | |
| LTE | | 3GPP Long Term Evolution | | |
| Mbps | | Megabit Per Second | | |
| MEC | | Multi-Access Edge Computing / Παρυφοϋπολογιστική Πολλαπλής Πρόσβασης | | |
| MHz | | Megahertz | | |
| MIMO | Multiple Input Multiple Output / Πολλαπλές Είσοδοι Πολλαπλές Έξοδοι | | |
| ML | | Machine Learning / Μηχανική Μάθηση | | |
| MLP | | Multi-layer Perceptron / Νευρωνικό Δίκτυο Perceptron Πολλών Στρωμάτων | | |
| mMTC | | Massive Machine-Type Communications / Μαζικές Επικοινωνίες Τύπου Μηχανής | | |
| mmWave | | Millimeter wave / Υψηλή Ζώνη | | |
| ms | | Milliseconds | | |
| NFV | | Network Function Virtualization / Εικονικοποίηση Δικτυακής Λειτουργίας | | |
| NG-RAN | | Next Generation Radio Access Network | | |
| PCA | | Principal Component Analysis / Ανάλυση Κύριων Συνιστωσών | | |
| QoS | | Quality of Service / Ποιότητα Σήματος | | |
| RAN | | Radio Access Network | | |
| RB | | Resource Block / Μπλοκ Πόρων | | |
| RL | | Reinforcement Learning / Ενισχυτική Μάθηση | | |
| SBA | | Service Based Architecture / Υπηρεσιοπαγής Αρχιτεκτονική | | |
| SDN | | Software Defined Networking / Δικτύωση Καθορισμένη από Λογισμικό | | |
| SINR | | Signal-to-Interference-plus-Noise Ratio / Λόγος Σήματος Προς Θόρυβο Και Παρεμβολή | | |
| SMF | | Session Management Function / Διαχείριση Συνεδρίας | | |
| SNR | | Signal-to-Noise Ratio / Αναλογία Σήματος Προς Θόρυβο | | |
| t-SNE | | t-Distributed Stochastic Neighbor Embedding | | |
| UE | | User Equipment / Εξοπλισμός Χρήστη | | |
| UPF | | User Plane Function / Λειτουργία Επιπέδου Χρήστη | | |
| URLLC | | Ultra Reliable Low Latency Communications / Επικοινωνίες Υπεραξιόπιστες και Με Χαμηλή Καθυστέρηση | | |
| VoIP | | Voice Over Internet Protocol / Τηλεφωνία μέσω Διαδικτύου | | |
| VR | | Virtual Reality / Εικονική Πραγματικότητα | | |
| WCSS | | Within Cluster Sum of Squares / Άθροισμα Τετραγώνων Εντός Συστάδας | | |

# 

# Εισαγωγή

## Σημασία του προβλήματος

Τα δίκτυα 5ης Γενιάς (5th Generation - 5G) εισάγουν μια επαναστατική προοπτική στις τηλεπικοινωνίες, υποσχόμενα ταχύτητες επικοινωνίας ασύλληπτα υψηλότερες από τις προηγούμενες γενιές, χαμηλότερες καθυστερήσεις και τη δυνατότητα σύνδεσης μεγάλου αριθμού συσκευών. Αυτά τα χαρακτηριστικά ανοίγουν το δρόμο για εφαρμογές που κυμαίνονται από το Διαδίκτυο των Πραγμάτων (Internet of Things - IoT) μέχρι εξελιγμένα συστήματα αυτοματοποιημένης οδήγησης και προηγμένες υπηρεσίες επικοινωνίας [1].

Παρά τις εντυπωσιακές δυνατότητές τους, τα δίκτυα 5G αντιμετωπίζουν σημαντικές προκλήσεις, ειδικά στο πεδίο της κατανομής πόρων. Σε αντίθεση με τους προκατόχους του, το 5G απαιτεί καινοτόμες προσεγγίσεις για να βελτιστοποιήσει την κατανομή πόρων όπως το φάσμα και η ενέργεια. Αυτή η πολυπλοκότητα εντείνεται περαιτέρω από την εμφάνιση των συστημάτων MIMO, τα οποία, ενώ αυξάνουν τη χωρητικότητα και την αποδοτικότητα, απαιτούν περίπλοκη διαχείριση για την πλήρη αξιοποίηση τους.

Στον τομέα της αποδοτικής κατανομής πόρων στα δίκτυα 5G, έχουν γίνει σημαντικές ερευνητικές προσπάθειες που επικεντρώνονται στην ολοκληρωμένη διαχείριση των διαφορετικών απαιτήσεων των υπηρεσιών. Για παράδειγμα, υπάρχουν εργασίες που εξετάζουν την βελτίωση της κατανομής πόρων με την εισαγωγή μεθοδολογιών βασισμένων σε ακέραιο γραμμικό προγραμματισμό, προσαρμοσμένων για την αντιμετώπιση των συγκεκριμένων απαιτήσεων των υπηρεσιών 5G [2].

Παράλληλα, αρκετές εργασίες υπογραμμίζουν τις προκλήσεις και τις ευκαιρίες που παρουσιάζονται με την εισαγωγή των τεχνολογιών 5G και πέρα από το 5G (Beyond 5G - B5G) ενώ εστιάζουν στις στρατηγικές κατανομής πόρων και κοινής χρήσης φάσματος καθώς και σε προηγμένες στρατηγικές για τη δυναμική κατανομή πόρων, δείχνοντας τη δυνατότητα της ML και των "παραδοσιακών" στρατηγικών να διαχειρίζονται την πολυπλοκότητα και τις απαιτήσεις των δικτύων [3],[4],[5],[6],[7],[8].

Οι μέθοδοι ML προσφέρουν τη δυνατότητα δυναμικής και προσαρμοστικής διαχείρισης πόρων, χρησιμοποιώντας τόσο επιβλεπόμενη, μη επιβλεπόμενη όσο και ενισχυτική μάθηση (Reinforcement Learning – RL). Η επιβλεπόμενη μάθηση επιτρέπει την ανάλυση και την πρόβλεψη της συμπεριφοράς των χρηστών και της κίνησης του δικτύου, ενώ η μη επιβλεπόμενη μάθηση όπως οι αλγόριθμοι ομαδοποίησης αποκαλύπτει μοτίβα και δομές από μη επισημασμένα δεδομένα. Ειδικότερα, τεχνικές όπως η βαθιά RL έχουν αποδειχθεί ικανές να βελτιστοποιήσουν την κατανομή ραδιοφωνικών πόρων, επιτυγχάνοντας μια ισορροπία μεταξύ διαφόρων απαιτήσεων ποιότητας υπηρεσίας (Quality of Service - QoS) σε διαφορετικά τμήματα του δικτύου.

Εργασίες που βασίζονται στην ML επιδεικνύουν την εφαρμογή προσαρμοστικών μοντέλων όπως Q-learning, Νευρωνικού Δικτύου Perceptron με Πολλά Στρώματα (Multi-layer Perceptron - MLP) και του αλγορίθμου “Χωρική Ομαδοποίηση Εφαρμογών με Θόρυβο με βάση την Πυκνότητα” (Density-Based Spatial Clustering of Applications with Noise - DBSCAN) για τη βελτιστοποίηση της κατανομής δικτυακών πόρων στα δίκτυα 5G, την αποτελεσματικότερη διαχείριση σφαλμάτων και την ελαχιστοποίηση των παρεμβολών. Ενώ άλλες εργασίες προσπαθούν να επισημάνουν την αυξανόμενη πολυπλοκότητα στα δίκτυα, τον τεράστιο όγκο δεδομένων που αυτά θα έχουν και τον κρίσιμο ρόλο της τεχνητής νοημοσύνης στα δίκτυα 5G MIMO για τη βελτίωση της αποδοτικότητας, της αξιοπιστίας και της απόδοσης των συστημάτων 5G, ενώ χρησιμοποιούνται μοντέλα Hidden Markov και K-Means για τη βελτιστοποίηση της κάλυψης και της χωρητικότητας των καναλιών [9],[10],[11],[12],[13],[14],[15].

Στο πεδίο της κατανομής πόρων σε δίκτυα 5G MIMO, οι υφιστάμενες εργασίες έχουν εστιάσει κυρίως στη βελτίωση της αποδοτικότητας και της ποιότητας των υπηρεσιών μέσω διάφορων τεχνικών, όπως ο διαμοιρασμός ο διαμοιρασμός φάσματος (spectrum sharing) και η δυναμική διαχείριση των πόρων. Για παράδειγμα, υπάρχουν εργασίες που εισάγουν μεθόδους βασισμένες σε ακέραιο γραμμικό προγραμματισμό στοχεύοντας στην καλύτερη εξυπηρέτηση ετερογενών απαιτήσεων υπηρεσιών, και άλλες που αναδεικνύουν τις προκλήσεις και τις ευκαιρίες της κοινής χρήσης φάσματος [16],[17].

Παρ' όλα αυτά, παρατηρείται ένα κενό στην έρευνα σχετικά με την ενσωμάτωση και την πλήρη αξιοποίηση της ML για τη διαχείριση των πόρων σε περιβάλλοντα 5G MIMO. Αν και υπάρχουν εργασίες όπως που εφαρμόζουν τεχνικές ML για προβλέψεις κίνησης και διαχείριση φάσματος, υπάρχει ανάγκη για βαθύτερη εξερεύνηση της ικανότητας της ML να προσαρμόζει δυναμικά τις αποφάσεις κατανομής πόρων σε πραγματικό χρόνο, ειδικά σε πολύπλοκες και μεταβαλλόμενες συνθήκες δικτυακού φορτίου [18]. Επίσης, η περαιτέρω έρευνα για την αξιοποίηση μη επιβλεπόμενων και ενισχυτικών μοντέλων μάθησης που μπορούν να αναγνωρίζουν και να ανταποκρίνονται σε πολυδιάστατα προβλήματα κατανομής πόρων είναι κρίσιμη, εφόσον οι τεχνολογίες αυτές προσφέρουν δυνατότητες για πιο βιώσιμες και αποδοτικές λύσεις.

Το κύριο κενό λοιπόν, στην υφιστάμενη βιβλιογραφία εστιάζεται στην ανάγκη για ανάπτυξη και αξιολόγηση εξελιγμένων στρατηγικών ML και δυναμικής ανακατανομής των πόρων που θα μπορούσαν να ανταποκρίνονται σε διαφορετικές ανάγκες υπηρεσιών και ποιότητας, καθώς και να διασφαλίζουν την αποδοτική και δίκαιη κατανομή πόρων σε ένα πολύπλοκο δίκτυο όπως το δίκτυο 5G.

Έτσι, κεντρικό σημείο της διπλωματικής εργασίας αποτελεί το πώς διάφοροι αλγόριθμοι συσταδοποίησης μπορούν να εφαρμοστούν αποτελεσματικά στο πλαίσιο του 5G για να επιτευχθεί βέλτιστη κατανομή πόρων. Αυτή η εργασία επιδιώκει να αποδείξει τη χρησιμότητα των μεθόδων μη επιβλεπόμενης ML στα περιβάλλοντα 5G.

## Στόχοι της εργασίας

Οι κύριοι στόχοι αυτής της διπλωματικής εργασίας είναι η παρουσίαση τεχνικών μη επιβλεπόμενης ML για ομαδοποίηση των χρηστών, η ανάλυση και η ανάπτυξη ενός αποδοτικού συστήματος για τη δυναμική κατανομή πόρων στα δίκτυα 5G MIMO χρησιμοποιώντας τις τεχνικές αυτές και τέλος η σύγκριση της αποτελεσματικότητας του συστήματος αυτού με τους «παραδοσιακούς» αλγορίθμους κατανομής πόρων. Οι ειδικότεροι στόχοι της εργασίας περιλαμβάνουν:

* **Κατανόηση των προκλήσεων και των απαιτήσεων της διαχείρισης πόρων στα δίκτυα 5G**: Εξετάζοντας τις τεχνολογικές προκλήσεις, η εργασία θα παρουσιάσει την ανάγκη για πιο ευέλικτες και δυναμικές μεθόδους κατανομής πόρων για την επίτευξη σύνδεσης από όλους τους χρήστες ακόμα και με χαμηλότερη ποιότητα καθώς και την ανάγκη για ομαδοποίηση των χρηστών με σκοπό την ευκολότερη διαχείρισή τους.
* **Εξερεύνηση της εφαρμογής των αλγόριθμων ομαδοποίησης στη διαχείριση πόρων και σύγκριση των αλγορίθμων**: Η εργασία θα αναλύσει διάφορους αλγόριθμους ομαδοποίησης όπως K-Means, DBSCAN, και άλλους για την αποδοτική ομαδοποίηση των χρηστών με βάση τις απαιτήσεις τους και την ποιότητα του δικτυακού τους σήματος καθώς και άλλα χαρακτηριστικά τους.
* **Ανάπτυξη αλγορίθμων για τη δυναμική κατανομή πόρων και σύγκρισή τους με απλές μεθόδους**: Η εργασία, θα χρησιμοποιήσει την εφαρμογή νέων μεθόδων δυναμικής κατανομής πόρων με τη χρήση των αποτελεσμάτων της ομαδοποίησης, στοχεύοντας στην βελτιστοποίηση της κατανομής των μπλοκ πόρων (resource blocks – RBs) σε αυτούς τους ομαδοποιημένους χρήστες και συνεπώς στην βελτιστοποίηση της απόδοσης και της ποιότητας του δικτύου. Η απόδοση αυτή ύστερα θα συγκριθεί με εκείνη ενός παραδοσιακού αλγορίθμου κατανομής που βασίζεται μόνο στην ποιότητα του σήματος των χρηστών στις 2 κεραίες. Στόχος του αλγορίθμου που θα προταθεί, θα είναι η ικανοποίηση των αναγκών όλων των χρηστών ακόμα και σε χαμηλότερη ποιότητα. Αναμένεται ότι ο νέος αλγόριθμος θα αποδειχθεί ανώτερος, προσφέροντας αποδοτικότερη κατανομή πόρων.

## Μεθοδολογία προσέγγισης

Η μεθοδολογία της προσέγγισης που παρουσιάζεται σε αυτή την διπλωματική εργασία έχει τα εξής στάδια :

* **Προσομοίωση δικτύου 5G MIMO:** Ανάπτυξη μίας προσομοίωσης που μιμείται το περιβάλλον ενός δικτύου 5G MIMO με δύο σταθμούς τύπου macrocell καθώς και χρήστες με διαφορετικές απαιτήσεις και διαφορετικές συντεταγμένες σε διάφορες κατανομές, με σκοπό την ανάλυση ενός συστήματος 5G MIMO και του σήματος των χρηστών.
* **Διαχείριση δεδομένων και ομαδοποίηση χρηστών:** Χρήση τεχνικών μη εποπτευόμενης ML για μείωση των διαστάσεων όπως η Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis - PCA) και για συσταδοποίηση όπως K-Means clustering, DBSCAN, Hierarchical clustering (HC), spectral clustering, Γκαουσιανό Μοντέλο Μείξης (Gaussian Mixture Model - GMM) για την ομαδοποίηση των χρηστών βάσει διαφόρων κριτηρίων όπως η ποιότητα του σήματος, η απαίτηση τους σε Megabit per second (Mbps), η μέγιστη αποδεκτή καθυστέρηση της υπηρεσίας τους, η κοντινότερη κεραία σε αυτούς και ο λόγος σήματος προς θόρυβο (Signal-to-Noise Ratio - SNR) και ύστερα την ανάλυση και επεξεργασία των συστάδων που δημιουργούνται.
* **Βελτιστοποίηση και σύγκριση των αλγορίθμων ομαδοποίησης:** Βελτιστοποίηση των αλγορίθμων ομαδοποίησης με την εύρεση των βέλτιστων παραμέτρων με χρήση συναρτήσεων και μεθόδων όπως elbow point για την εκτίμηση του βέλτιστου αριθμού συστάδων ή silhouette και calinsky-harabasz (CH) scores για την εκτίμηση της πυκνότητας των συστάδων που δημιουργούνται από τους αλγορίθμους και της απόστασης των συστάδων. Σκοπός αυτών, η εύρεση αριθμού συστάδων που επιτυγχάνει τη μέγιστη δυνατή πυκνότητα εντός συστάδας και τη μέγιστη δυνατή απόσταση μεταξύ διαφορετικών συστάδων, κάτι που υποδηλώνει καλύτερη ομαδοποίηση των χρηστών.
* **Ανάπτυξη αλγορίθμων για την κατανομή πόρων:** Εστιάζοντας σε "hotspots" χρηστών στο χώρο, κάθε ένα με τις δικές του απαιτήσεις και χαρακτηριστικά, γίνεται σχεδιασμός και εφαρμογή ενός αλγορίθμου που είναι βασισμένος στα αποτελέσματα των αλγορίθμων συσταδοποίησης των χρηστών και αποσκοπεί στη δυναμική ρύθμιση της κατανομής των πόρων (σε μορφή RB) στα δίκτυα 5G. Ύστερα, πραγματοποιείται η σύγκρισή του με τον «παραδοσιακό» τρόπο κατανομής πόρων ο οποίος βασίζεται μόνο στην απόσταση και στο καλύτερο σήμα.

## Συνεισφορά

Η εργασία αυτή στον τομέα των δικτύων 5G MIMO συμβάλλει σημαντικά στην κατανόηση των αλγορίθμων συσταδοποίησης και στη βελτίωση της διαχείρισης και κατανομής πόρων με τη χρήση αλγορίθμων δυναμικής κατανομής. Εφόσον η μεθοδολογία που παρουσιάζεται επικεντρώνεται στην προσομοίωση ενός περιβάλλοντος 5G MIMO και εξυπηρετεί τις ανάγκες των χρηστών με τη χρήση μη επιβλεπόμενης ML, αξιοποιώντας τα δεδομένα των χρηστών και των κεραιών ώστε να παραχθούν τα απαιτούμενα συμπεράσματα, καταφέρνει να κλείσει ένα κρίσιμο κενό στην υφιστάμενη βιβλιογραφία. Αυτό το καταφέρνει με τη χρήση συσταδοποίησης, διαστατικής μείωσης και ενός βελτιωμένου αλγορίθμου κατανομής με σκοπό τη βελτιστοποίηση της διαχείρισης των δικτυακών πόρων σε περιβάλλοντα υψηλής πολυπλοκότητας, αλλά ταυτόχρονα τη δικαιοσύνη και την αποδοτικότητα στη διανομή των πόρων, εξασφαλίζοντας τη βέλτιστη εμπειρία για όλους τους χρήστες.

## Διάρθρωση της υπόλοιπης διπλωματικής εργασίας

Η Διπλωματική Εργασία δομείται ως εξής :

Το δεύτερο κεφάλαιο, εστιάζει στην τεχνολογία 5G και στις λειτουργίες που αυτό προσφέρει. Δίνει επίσης μια πλήρη επισκόπηση του πώς η τεχνολογία 5G φέρνει την επανάσταση στις τηλεπικοινωνίες, επισημαίνοντας τα χαρακτηριστικά της και τα φάσματα στα οποία λειτουργεί, την αρχιτεκτονική της αλλά και προβλήματα που αυτή η τεχνολογία αντιμετωπίζει. Επίσης, εμβαθύνει στην εξερεύνηση των νέων λειτουργιών που επιτρέπουν την ταυτόχρονη σύνδεση εκατοντάδων συσκευών στο δίκτυο, χωρίς να αυξάνονται οι καθυστερήσεις υποβιβάζοντας την ταχύτητα μεταφοράς δεδομένων.

Στο τρίτο κεφάλαιο, συζητείται η κατανομή πόρων στα δίκτυα 5G, δίνεται η έννοια του RB και παρουσιάζονται οι τεχνικές κατανομής διαφορετικών πόρων και οι προκλήσεις που πρέπει να αντιμετωπιστούν.

Στο τέταρτο κεφάλαιο, γίνεται εξερεύνηση στην ML στα δίκτυα 5G και στον συνδυασμό της με την κατανομή πόρων, τονίζοντας την αναγκαιότητα για ML στα δίκτυα αυτά και στον τρόπο με τον οποίο αυτή βοηθά την κατανομή πόρων. Συζητούνται διάφορες μέθοδοι και τεχνολογίες ML (επιβλεπόμενης, βαθιάς και ενισχυτικής μάθησης) που έχουν χρησιμοποιηθεί σε άλλες έρευνες για την κατανομή πόρων σε δίκτυα πέμπτης γενιάς και τέλος δίνεται έμφαση στην μη επιβλεπόμενη ML και στα προβλήματα που υπάρχουν στη χρήση της στο 5G.

Στο πέμπτο κεφάλαιο, αναλύεται η πειραματική διαδικασία, το περιβάλλον του πειράματος, η «μηχανική» των χαρακτηριστικών, ο χειρισμός και η προεπεξεργασία τους, οι αλγόριθμοι που χρησιμοποιήθηκαν, η βελτιστοποίησή τους και η σύγκριση των αποτελεσμάτων τους.

Στο έκτο κεφάλαιο, αναλύονται τα αποτελέσματα με τη μορφή γραφημάτων, αποδεικνύεται η αποτελεσματικότητα της προσέγγισης σε σχέση με την παραδοσιακή προσέγγιση και αναφέρονται πραγματικές περιπτώσεις όπου η προσέγγισή που προτείνεται μπορεί να βοηθήσει στην αποτελεσματικότερη κατανομή πόρων και διαχείριση χρηστών σε περιβάλλοντα 5G. Επίσης, συζητούνται τρόποι με τους οποίους αυτή η έρευνα μπορεί να επεκταθεί στο μέλλον.

Στο τέλος, βρίσκεται το κεφάλαιο της Βιβλιογραφίας που χρησιμοποιήθηκε για την συγγραφή της διπλωματικής και το Παράρτημα με τον κώδικα που χρησιμοποιήθηκε στην εργασία.

# 

# Δίκτυα κινητής τηλεφωνίας 5ης γενιάς

## Εισαγωγή στην 5η γενιά των δικτύων κινητής τηλεφωνίας

Η επικείμενη εποχή των δικτύων 5G αποτελεί μια επαναστατική φάση για τις τηλεπικοινωνίες, φέρνοντας μαζί της μεγάλες ταχύτητες και αξιόπιστες επικοινωνίες. Τα 5G δίκτυα, με την ικανότητά τους να υποστηρίζουν ταχύτητες Gigabit per second (Gbps), πολύ μικρή καθυστέρηση, και τεράστια δικτυακή χωρητικότητα, ανταποκρίνονται στην αυξανόμενη ζήτηση για δεδομένα και συνδεσιμότητα. Με ταχύτητες που ξεπερνούν τα 20 Gbps, το 5G υπόσχεται ασύγκριτες εμπειρίες στην ψηφιακή ψυχαγωγία, τις online υπηρεσίες, και την επικοινωνία σε πραγματικό χρόνο, μειώνοντας την καθυστέρηση σε ελάχιστα χιλιοστά του δευτερολέπτου.

Η καινοτομία του 5G δεν περιορίζεται στην ταχύτητα και την καθυστέρηση. Η χωρητικότητά του που δίνει τη δυνατότητα σύνδεσης αμέτρητων συσκευών προσφέρει το έδαφος για την ανάπτυξη ευφυών πόλεων και βιομηχανιών, ενώ η χρήση διαφόρων ζωνών συχνοτήτων διασφαλίζει την εξυπηρέτηση διάφορων αναγκών, από την κάλυψη εκτεταμένων αγροτικών εκτάσεων μέχρι την υποστήριξη υψηλής χωρητικότητας σε αστικά κέντρα. Η αρχιτεκτονική του, με νέες τεχνολογίες, βελτιστοποιεί την απόδοση και προσφέρει εξατομικευμένες εμπειρίες.

Κάθε ζώνη του 5G προσφέρει μοναδικά οφέλη, με τη χαμηλή ζώνη να παρέχει εκτεταμένη κάλυψη και βαθιά διείσδυση εσωτερικών χώρων. H μεσαία ζώνη προσφέρει μια ισορροπία κάλυψης και χωρητικότητας, και η υψηλή ζώνη (millimeter wave - mmWave)[[1]](#footnote-1) παρέχει υπερυψηλές ταχύτητες. Η ποικιλία στις ζώνες συχνοτήτων που μας προσφέρουν τα δίκτυα 5G, διασφαλίζει την εξυπηρέτηση μίας ευρείας γκάμας περιπτώσεων χρήσης και περιβαλλόντων, από πυκνοκατοικημένες αστικές περιοχές που απαιτούν υψηλή χωρητικότητα έως εκτεταμένες αγροτικές περιοχές που χρειάζονται κάλυψη [19]

Στον πυρήνα της αρχιτεκτονικής του 5G βρίσκεται ο καινοτόμος σχεδιασμός του δικτύου, το οποίο περιλαμβάνει προηγμένες τεχνολογίες όπως διαμόρφωση δέσμης (beamforming), Massive MIMO, network slicing, και Συγχώνευση Φορέων (carrier aggregation). Αυτές οι τεχνολογίες συνεργάζονται για να βελτιστοποιήσουν την απόδοση των δικτύων 5G, επιτρέποντάς τους να διανέμουν δυναμικά πόρους, να ενισχύουν την ποιότητα του σήματος, να αυξάνουν την ροή δεδομένων, και να παρέχουν εξατομικευμένες υπηρεσίες που ανταποκρίνονται στις διάφορες ανάγκες των χρηστών και των εφαρμογών.

Συνοπτικά, τα δίκτυα 5G σηματοδοτούν ένα σημαντικό ορόσημο στην εξέλιξη των τηλεπικοινωνιών, χαρακτηριζόμενα από την υψηλή τους ταχύτητα, τη χαμηλή τους καθυστέρηση, και τις εκτεταμένες δυνατότητες συνδεσιμότητας. Ως ο κορμός της μελλοντικής ψηφιακής υποδομής, το 5G ανοίγει νέους ορίζοντες για τεχνολογική καινοτομία και έχει τη δυνατότητα να μεταμορφώσει διάφορους τομείς, συμπεριλαμβανομένων της υγείας, των μεταφορών και της ψυχαγωγίας, εισάγοντας μια νέα εποχή στα ασύρματα δίκτυα.

## Αρχιτεκτονική του δικτύου 5G

Η αρχιτεκτονική του 5G, ένα παράδειγμα προσαρμοστικότητας και κλιμακωσιμότητας, υποστηρίζει από ταχύτατη κινητή ευρυζωνικότητα έως ένα πλήθος δικτύων IoT. Στο κέντρο αυτής της τεχνολογικής ανάπτυξης είναι οι Εξοπλισμοί Χρήστη (User Equipment - UE), που εκτείνονται πέραν των παραδοσιακών smartphones και περιλαμβάνουν διάφορες συσκευές IoT και χάρη σε τεχνολογίες, όπως το Massive MIMO και το beamforming, σε αυτά τα UE βελτιστοποιείται η εμπειρία σύνδεσης, δίνοντας την ευκαιρία για μια αναβαθμισμένη ψηφιακή ζωή.

Η αρχιτεκτονική 5G αρχίζει με το Next-Generation Radio Access Network (NG-RAN), το οποίο ενσωματώνει τα gNodeBs και ng-eNBs, υποστηρίζοντας την ασύρματη επικοινωνία με τις συσκευές χρηστών. Η ικανότητά του να λειτουργεί με ομαλότητα σε ένα ευρύ φάσμα συχνοτήτων, από τις mmWave μέχρι και κάτω των 6 GHz, το καθιστά εξαιρετικά ευέλικτο, εξασφαλίζοντας τόσο την υψηλή χωρητικότητα όσο και την πλήρη κάλυψη ώστε να μπορεί να ανταπεξέλθει στις απαιτητικές ανάγκες δεδομένων των υπηρεσιών 5G, προσφέροντας αξιόπιστη συνδεσιμότητα για ένα ευρύ φάσμα εφαρμογών.

Το Δίκτυο Πυρήνα 5ης  Γενιάς (5th Generation Core Network - 5GC) φέρνει μια ουσιαστική αλλαγή. Με κεντρικό στοιχείο την Υπηρεσιοπαγή Αρχιτεκτονική (Service-Based Architecture SBA), το 5GC αναπτύσσει μια δομή όπου οι δικτυακές λειτουργίες συνεργάζονται σε ένα περιβάλλον, φέρνοντας την επανάσταση στην ευελιξία και στην αποκρισιμότητα του δικτύου.

Λειτουργίες κλειδιά όπως η Διαχείριση Πρόσβασης και Κινητικότητας (Access and Mobility Function - AMF), η Διαχείριση Συνεδρίας (Session Management Function - SMF), και η Λειτουργία Επιπέδου Χρήστη (User Plain Function - UPF) αποτελούν τη βάση για την ομαλή λειτουργία του 5G. Η AMF διαχειρίζεται τη συνδεσιμότητα και την κινητικότητα, η SMF τη δημιουργία και διατήρηση των συνεδριών, ενώ η UPF την προώθηση και δρομολόγηση δεδομένων εξασφαλίζοντας την ανταπόκριση του δικτύου στις δυναμικά μεταβαλλόμενες απαιτήσεις, και υποστηρίζοντας ένα φάσμα υπηρεσιών νέας γενιάς με μεγαλύτερη αποδοτικότητα[[2]](#footnote-2).

Η Παρυφοϋπολογιστική Πολλαπλής Πρόσβασης (Multi-Access Edge Computing - MEC) αποτελεί μια μετασχηματιστική προσέγγιση στην αρχιτεκτονική 5G που τοποθετεί «στρατηγικά» τους υπολογιστικούς πόρους στην περιφέρεια του δικτύου, σημαντικά πιο κοντά στους τελικούς χρήστες. Αυτή η αρχιτεκτονική καινοτομία μειώνει δραστικά την καθυστέρηση, ενισχύει τις ταχύτητες επεξεργασίας δεδομένων. Στον χώρο του IoT, η MEC επιτρέπει την ταχεία επεξεργασία και ανάλυση τεράστιων ροών δεδομένων, ενδυναμώνοντας τις έξυπνες υποδομές και τα συνδεδεμένα οικοσυστήματα.

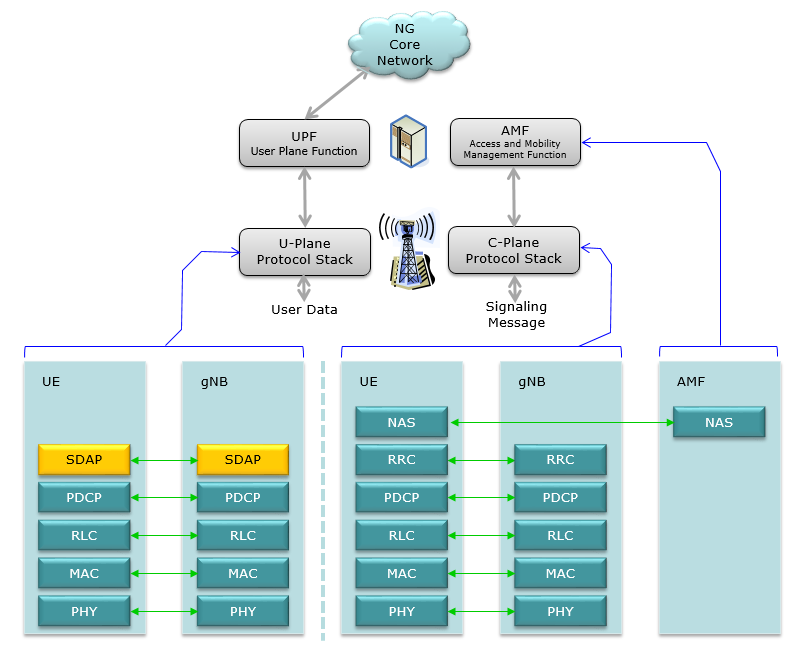
Η αρχιτεκτονική Cloud-Radio Access Network (CRAN) ανοίγει ένα νέο κεφάλαιο στον σχεδιασμό δικτύων, παρέχοντας μια σημαντική εξέλιξη από τα παραδοσιακά μοντέλα. Μέσω της εικονικοποίησης (virtualization) στο cloud, απελευθερώνει ένα πεδίο νέων δυνατοτήτων σε κλιμακωσιμότητα, ευελιξία και αποτελεσματικότητα.

Τα δίκτυα backhaul και fronthaul αποτελούν τους πυλώνες της αρχιτεκτονικής 5G, εξασφαλίζοντας αποδοτικότητα και υψηλή απόδοση, υποστηρίζοντας την παροχή υπηρεσιών από υψηλής ποιότητας βίντεο μέχρι βιομηχανικές εφαρμογές IoT και έξυπνες πόλεις. Το backhaul συνδέει το NG-RAN με το κεντρικό δίκτυο του 5G, διευκολύνοντας την ροή δεδομένων μέσα στο δίκτυο. Το fronthaul, από την άλλη, δημιουργεί μια σύνδεση μεταξύ των ραδιοφωνικών μονάδων και των κεντρικών μονάδων, κρίσιμη για τη λειτουργία των Cloud-RAN ή διανεμημένων Radio Access Network (RAN) και υιοθετεί προηγμένες τεχνολογίες όπως το CRAN ή το evolved Common Public Radio Interface (eCPRI), στοχεύοντας στη μείωση της καθυστέρησης και στη μεγιστοποίηση της ακρίβειας του συγχρονισμού. Η ομαλή λειτουργία τους λοιπόν είναι κρίσιμη για την εκπλήρωση της υπόσχεσης του 5G για έναν κόσμο πιο συνδεδεμένο και τεχνολογικά προηγμένο.

### Τα επίπεδα της αρχιτεκτονικής 5G

Η τεχνολογία 5G λοιπόν, διαθέτει μια σύνθετη αρχιτεκτονική που περιλαμβάνει πολλαπλά επίπεδα τόσο στη δομή του δικτύου όσο και στη στοίβα πρωτοκόλλων. Κάθε επίπεδο εκτελεί διακριτές λειτουργίες και χρησιμοποιεί συγκεκριμένα πρωτόκολλα για την εξασφάλιση της αποδοτικότητας και της ασφάλειας στην επικοινωνία. Στην Εικόνα 1, παρουσιάζονται τα παρακάτω επίπεδα της αρχιτεκτονικής του 5G :

* **Το φυσικό επίπεδο (PHY)** αποτελεί τη βάση της στοίβας πρωτοκόλλων 5G και είναι υπεύθυνο για τη φυσική μετάδοση και λήψη των σημάτων μέσω του αέρα. Εδώ χρησιμοποιούνται τεχνικές όπως η μοντελοποίηση, η κωδικοποίηση και η αποκωδικοποίηση, καθώς και η εκτίμηση του καναλιού για τη βελτιστοποίηση της απόδοσης του σήματος.
* **Το επίπεδο MAC** διαχειρίζεται την πρόσβαση στους ραδιοφωνικούς πόρους και τον χρονοπρογραμματισμό της μετάδοσης δεδομένων μεταξύ πολλαπλών χρηστών. Αυτό το επίπεδο είναι κρίσιμο για την αποδοτική κατανομή των πόρων και τη διασφάλιση της δίκαιης πρόσβασης στο δίκτυο.
* **Στο επίπεδο RLC**, τα δεδομένα που μεταφέρονται θα υποστούν διαδικασίες ελέγχου σφαλμάτων, διαχωρισμού και επανασυναρμολόγησης. Το RLC είναι απαραίτητο για την εξασφάλιση της αξιοπιστίας και της συνέπειας των δεδομένων που μεταδίδονται πάνω στη ραδιοφωνική διασύνδεση.
* **Το PDCP** χειρίζεται τη συμπίεση και αποσυμπίεση των IP πακέτων, καθώς και την κρυπτογράφηση και την αποκρυπτογράφηση των δεδομένων του χρήστη. Αυτό το επίπεδο είναι ζωτικής σημασίας για τη διασφάλιση της ασφάλειας και του απορρήτου των μεταδιδόμενων δεδομένων.
* **Το RRC** επιτηρεί και διαχειρίζεται την κατάσταση των ραδιοπόρων μεταξύ του UE και του δικτύου. Αυτό περιλαμβάνει τη διαμόρφωση, τη διαχείριση και την απελευθέρωση των ραδιοπόρων, καθώς και τη διαχείριση της κινητικότητας των χρηστών.
* **Το SDAP** ασχολείται με τη διαχείριση της ποιότητας της υπηρεσίας. Αυτό το επίπεδο εξασφαλίζει ότι διάφοροι τύποι κυκλοφορίας λαμβάνουν τις κατάλληλες QoS παραμέτρους.
* **Το NAS** είναι υπεύθυνο για τη διαχείριση σημαντικών λειτουργιών όπως η εγγραφή στο δίκτυο, η ταυτοποίηση, η ασφάλεια, και η διαχείριση κινητικότητας. Αυτό το επίπεδο εξασφαλίζει την ομαλή λειτουργία του δικτύου και τη συνέχιση της υπηρεσίας ανεξαρτήτως της φυσικής τοποθεσίας του χρήστη.



**Εικόνα 1:** Τα επίπεδα αρχιτεκτονικής των επιπέδων χρηστών και ελέγχου του δικτύου 5G

### Φάσμα δικτύων 5G

Η τεχνολογία 5G χρησιμοποιεί μια ευρεία γκάμα συχνοτήτων, οι οποίες κατηγοριοποιούνται ευρέως σε τρεις κύριες ζώνες φάσματος: τη χαμηλή ζώνη, τη μεσαία ζώνη, και το mmWave. Κάθε ζώνη έχει συγκεκριμένους ρόλους και χαρακτηριστικά, προσφέροντας διαφορετικές ισορροπίες μεταξύ κάλυψης, χωρητικότητας και ταχύτητας μετάδοσης δεδομένων.

* Η χαμηλή ζώνη φάσματος, συχνά αναφερόμενη ως "στρώμα κάλυψης", παρέχει εξαιρετική διείσδυση σε κτίρια και εμπόδια, καθιστώντας την ιδανική για εκτεταμένη κάλυψη σε αγροτικές και δυσπρόσιτες περιοχές. Η συγκεκριμένη ζώνη υποστηρίζει βασικές υπηρεσίες φωνής και δεδομένων με αξιόπιστη συνδεσιμότητα, αλλά χαρακτηρίζεται από σχετικά χαμηλότερες ταχύτητες μετάδοσης και χωρητικότητα σε σύγκριση με τις μεγαλύτερες ζώνες.
* Η μεσαία ζώνη προσφέρει μια ισορροπία μεταξύ κάλυψης και χωρητικότητας. Είναι πιο κατάλληλη για αστικές περιοχές όπου η πυκνότητα των χρηστών και ο όγκος δεδομένων είναι υψηλότερος. Υποστηρίζει υψηλότερες ταχύτητες και χαμηλότερες καθυστερήσεις από τη χαμηλή ζώνη, καθιστώντας την ιδανική για την παροχή πιο απαιτητικών υπηρεσιών όπως η αναπαραγωγή βίντεο υψηλής ευκρίνειας, αλλά και για υπηρεσίες IoT που απαιτούν σταθερή συνδεσιμότητα και γρήγορη ανταπόκριση.
* Το mmWave, προσφέρει τις μέγιστες ταχύτητες και χωρητικότητα, τοποθετώντας την στην κορυφή του φάσματος για εφαρμογές όπως υπερ-υψηλής ευκρίνειας βίντεο, πραγματικού χρόνου παιχνίδια, και εξελιγμένες επιχειρηματικές εφαρμογές. Ωστόσο, τα μειονεκτήματα της υψηλής ζώνης περιλαμβάνουν περιορισμένη κάλυψη και αδύναμη διείσδυση μέσα από στερεά εμπόδια. Για να αντιμετωπιστούν αυτά τα προβλήματα, απαιτείται ένα πυκνό δίκτυο από κεραίες τύπου small cell, δηλαδή microcell, picocell και femtocell καθώς και αναμεταδότες σήματος για να διασφαλίζεται η συνεχής και αξιόπιστη συνδεσιμότητα.

## Νέες τεχνολογίες στο 5G

Η τεχνολογία 5G αναδεικνύεται ως ένας παράγοντας πρωτοποριακών αλλαγών με την εισαγωγή καινοτομιών που επιτρέπουν υψηλότερες ταχύτητες και χαμηλές καθυστερήσεις. Από το Beamforming και το Massive MIMO, μέχρι την εξερεύνηση του φάσματος mmWave και τις καινοτομίες όπως το network slicing και η επικοινωνία πλήρους διπλής κατεύθυνσης (Full Duplex), οι τεχνολογίες αυτές διευκολύνουν τη δημιουργία ενός αξιόπιστου, αποδοτικού και εξαιρετικά γρήγορου δικτύου.

### Διαμόρφωση δέσμης

Η διαμόρφωση δέσμης ή αλλιώς beamforming αποτελεί μια βασική τεχνολογία των δικτύων 5G, βελτιώνοντας με προσοχή την ακρίβεια και την αποδοτικότητα της ασύρματης επικοινωνίας μέσω της κατευθυνόμενης εκπομπής συγκεντρωμένων δεσμών ραδιοκυμάτων απευθείας προς συγκεκριμένους χρήστες. Αυτή η πολύπλοκη μέθοδος χρησιμοποιεί πλέγματα κεραιών για την εκπομπή συγχρονισμένων σημάτων, τα οποία, μέσω σύνθετων αλγορίθμων επεξεργασίας, προσαρμόζουν τη φάση και το πλάτος τους για να εστιάσουν ενέργεια στις επιθυμητές κατευθύνσεις, μειώνοντας ταυτόχρονα την ενέργεια προς τις κατευθύνσεις που δεν χρειάζεται να πάει το σήμα. Η τέτοιου είδους δυναμική διαμόρφωση σήματος είναι κρίσιμη για τη διατήρηση συνδέσμων επικοινωνίας, εξασφαλίζοντας βελτιωμένη ποιότητα σήματος και αξιόπιστη συνδεσιμότητα. Επιπλέον, με την ελαχιστοποίηση της διάχυσης σήματος και της παρεμβολής, η διαμόρφωση δέσμης ενισχύει σημαντικά την χωρητικότητα και την απόδοση του δικτύου, επιτρέποντας τις διάφορες και απαιτητικές εφαρμογές που προβλέπονται για το 5G, από τις εμπειρίες επαυξημένης πραγματικότητας έως τα αυτόνομα οχήματα, λειτουργώντας έτσι ως ένας καίριος παράγοντας του 5G.

### Massive MIMO

Το Massive MIMO, ένα απαραίτητο στοιχείο του τεχνολογικού συνόλου του 5G, εκμεταλλεύεται τη δύναμη των εκτεταμένων πλεγμάτων κεραιών στους βασικούς σταθμούς, αξιοποιώντας έναν μεγάλο αριθμό των στοιχείων της κεραίας για να διευκολύνει την ταυτόχρονη μετάδοση πολλαπλών ροών δεδομένων. Η εισαγωγή του Massive MIMO αποτελεί ένα ορόσημο στην τεχνολογική εξέλιξη, ενθαρρύνοντας τεράστια βήματα προόδου στη διαπερατότητα και φασματική αποδοτικότητα των δικτύων. Με την αξιοποίηση περισσότερων χωρικών καναλιών, το Massive MIMO, υπερβαίνει τα κλασσικά όρια και υποστηρίζει αποτελεσματικά τις ολοένα και αυξανόμενες ανάγκες για δεδομένα. Η ανάπτυξή του είναι κρίσιμη, καθώς παίζει έναν ουσιαστικό ρόλο στην καλύτερη αξιοποίηση του διαθέσιμου φάσματος, διασφαλίζοντας αξιόπιστη συνδεσιμότητα και σταθερή ροή υπηρεσιών σε μια πληθώρα εφαρμογών 5G, όπως την ενσωμάτωση του IoT.

### mmWave

Η τεχνολογία mmWave αποτελεί ένα μετασχηματιστικό στοιχείο του 5G, εξερευνώντας το υψηλότερο φάσμα συχνοτήτων μεταξύ 30 GHz και 300 GHz για να ξεκλειδώσει εκτεταμένα εύρη ζώνης και να διευκολύνει τη μετάδοση δεδομένων με πρωτοφανείς ταχύτητες. Αν και το άλμα αυτό στις συχνότητες mmWave εισάγει μια νέα εποχή υπερ-γρήγορης συνδεσιμότητας, συνοδεύεται από τις δικές της λεπτομέρειες. Τα σήματα mmWave διαθέτουν φυσικά μικρότερα εύρη διάδοσης και εμφανίζουν αυξημένη ευαισθησία σε φυσικά εμπόδια και ατμοσφαιρικές συνθήκες. Παρόλα αυτά, το mmWave υποστηρίζει την πρόσβαση σε υπερ-υψηλές ταχύτητες διαδικτύου σε πυκνοκατοικημένες περιοχές και σε ειδικές εφαρμογές όπου η ταχύτατη ανταλλαγή δεδομένων είναι κορυφαία. Η ενσωμάτωσή του στις υποδομές 5G είναι καθοριστική για την υποστήριξη της εκθετικής αύξησης στην κυκλοφορία δεδομένων.

### Διαχωρισμός δικτύου

Ο διαχωρισμός δικτύου ή αλλιώς network slicing αναδεικνύεται ως μια κύρια καινοτομία στην αρχιτεκτονική του 5G, προσφέροντας τη δυνατότητα να διαχωρίζεται μια ενιαία φυσική δικτυακή υποδομή σε πολυάριθμα εικονικά τμήματα ή slices, το καθένα εξοπλισμένο έτσι ώστε να εξυπηρετεί διακριτές υπηρεσίες ή ομάδες πελατών. Επιτρέπει λοιπόν στους διαχειριστές δικτύων να προσφέρουν εξατομικευμένα δικτυακά περιβάλλοντα που ικανοποιούν διαφορετικές απαιτήσεις, από την παροχή υψηλού εύρους ζώνης και υπερ-χαμηλής καθυστέρησης μέχρι την υποστήριξη εκτεταμένων οικοσυστημάτων IoT. Άρα με το network slicing, τα δίκτυα 5G επιτυγχάνουν εξαιρετική ευελιξία, βελτιώνοντας σημαντικά την αξιοποίηση πόρων και επιτρέποντας μια εξατομικευμένη εμπειρία υπηρεσίας.

### Βελτιωμένη κινητή ευρυζωνική σύνδεση (Enhanced Mobile Broadband - eMBB)

Η eMBB αποτελεί την κύρια τεχνολογία πίσω από τις εφαρμογές του 5G και ο στόχος του είναι να απογειώσει τις ταχύτητες δεδομένων και την χωρητικότητα δικτύου, ανταποκρινόμενο στην εντεινόμενη ανάγκη για εφαρμογές υψηλού εύρους ζώνης. Σχεδιασμένο για να φέρει στο προσκήνιο νέα επίπεδα ψηφιακών εμπειριών - από την καταγραφή βίντεο υπερυψηλής ευκρίνειας (4Κ και 8Κ) έως τα περιβάλλοντα Εικονικής Πραγματικότητας (Virtual Reality – VR) και Επαυξημένης Πραγματικότητας (Augmented Reality – AR), και τις αξιόπιστες υπηρεσίες cloud - το eMBB κατατάσσεται ως ένα σημαντικό στοιχείο του 5G. Αυτή η εξέλιξη που προσφέρει το eMBB στη συνδεσιμότητα υπόσχεται έναν ανεπτυγμένο κόσμο ψυχαγωγίας, επικοινωνίας και επαγγελματικών εμπειριών[[3]](#footnote-3).

### Επικοινωνίες υπεραξιόπιστες και με χαμηλή καθυστέρηση (Ultra Reliable Low Latency Communications - URLLC)

Η τεχνολογία URLLC έχει πρωταγωνιστικό ρόλο εντός του περιβάλλοντος του 5G, δημιουργημένη με μεγάλη προσοχή για να υποστηρίζει εφαρμογές όπου η άμεση ανταπόκριση και η άριστη αξιοπιστία δεν είναι απλώς επιθυμίες αλλά απόλυτες ανάγκες. Με στόχο να συμπιέσει την καθυστέρηση σε επίπεδα που πλησιάζουν το 1 χιλιοστό του δευτερολέπτου, εξασφαλίζοντας ταυτόχρονα την ακλόνητη αξιοπιστία, το URLLC αποκαλύπτεται ως κεντρικός παράγοντας για πληθώρα ζωτικών υπηρεσιών. Σε σενάρια όπου η παραμικρή καθυστέρηση ή απόκλιση έχει δραματικές επιπτώσεις, το URLLC είναι αυτό που εγγυάται την ικανότητα των δικτύων 5G να πληρούν τις αυστηρές προδιαγραφές αυτών των κρίσιμων εφαρμογών, προάγοντας την καινοτομία και την αξιοπιστία[[4]](#footnote-4).

### Μαζικές επικοινωνίες τύπου μηχανής (Massive Machine-Type Communications - mMTC)

Η τεχνολογία mMTC σηματοδοτεί ένα σπουδαίο στάδιο στην ανάπτυξη του 5G, αποκαλύπτοντας μια εποχή η οποία θα μπορέσει να υποστηρίξει την ενσωμάτωση του IoT. Φέρνοντας κάτω από την ίδια «στέγη» έναν πολύ μεγάλο αριθμό συνδεδεμένων συσκευών, καθεμία με τις δικές της ειδικές ανάγκες σε ροή δεδομένων, καθυστέρηση και οικονομία ενέργειας, το mMTC συμβάλλει στη διαμόρφωση του μέλλοντος του IoT και των δικτύων 5G. Αυτή η καινοτομία παρέχει την υποδομή για την σύνδεση αναρίθμητων συσκευών σε διάφορα πεδία, ενισχύοντας την έξυπνη διαχείριση της αστικής υποδομής, την εξελιγμένη γεωργία, τη βελτίωση της φροντίδας υγείας μέσω τεχνολογικών προόδων, και πολλά άλλα. Μέσα από τον ρόλο του στο 5G, το mMTC μεταμορφώνει την έννοια της απλής συνδεσιμότητας σε μια πραγματικά ενωμένη και ευφυή δικτυακή εμπειρία, προωθώντας την αποδοτικότητα, την αυτοματοποίηση και την ενημερωμένη λήψη αποφάσεων σε κάθε βήμα, θέτοντας τα θεμέλια για ένα μελλοντικό δίκτυο που θα είναι όχι μόνο συνδεδεμένο αλλά και έξυπνο[[5]](#footnote-5).

### Συγχώνευση φορέων

Η Συγχώνευση Φορέων ή carrier aggregation ανοίγει νέους δρόμους στην εποχή του 5G, φέρνοντας κοντά διαφορετικές ζώνες συχνοτήτων για να δημιουργήσει ένα ισχυρό ενιαίο κανάλι επικοινωνίας, μία καινοτομία στη διαχείριση του φάσματος, που κάνει τις ταχύτητες μετάδοσης δεδομένων και το εύρος ζώνης του δικτύου πολύ μεγαλύτερα, προσφέροντας μια ασυναγώνιστη εμπειρία στον χρήστη. Με τη συντονισμένη χρήση πολλαπλών ζωνών, τα δίκτυα 5G αποκτούν την ευελιξία να εκμεταλλεύονται αποδοτικά το διαθέσιμο φάσμα, ανταποκρινόμενα στις αυξημένες απαιτήσεις των διαφόρων εφαρμογών. Αυτή η προσέγγιση είναι καθοριστική για να διασφαλιστεί η άρτια και ταχύτατη συνδεσιμότητα που απαιτούν οι σύγχρονοι χρήστες σε έναν όλο και πιο συνδεδεμένο κόσμο.

### Επικοινωνία πλήρους διπλής κατεύθυνσης

Η τεχνολογία πλήρους διπλής κατεύθυνσης ή full duplex σηματοδοτεί ένα επαναστατικό βήμα στον τομέα των ασύρματων επικοινωνιών, ειδικά εντός του πλαισίου του 5G, επιτρέποντας την ταυτόχρονη αποστολή και λήψη ροών δεδομένων πάνω από μία μόνο ζώνη συχνοτήτων. Αυτή η καινοτομία έχει το δυναμικό να μετασχηματίσει ριζικά τη φασματική αποδοτικότητα, διπλασιάζοντας ουσιαστικά τη χωρητικότητα του καναλιού και ενισχύοντας ένα νέο επίπεδο απόδοσης του δικτύου και παρόλο που βρίσκεται ακόμη στα στάδια της έρευνας και ανάπτυξης, η τεχνολογία πλήρους διπλής κατεύθυνσης είναι προετοιμασμένη να αλλάξει τα δεδομένα, αυξάνοντας την αξιοπιστία. Ξεπερνώντας τα συνηθισμένα όρια της χρόνο-διαίρεσης και της συχνοτικής διαίρεσης, η τεχνολογία πλήρους διπλής κατεύθυνσης φαίνεται να ανοίγει νέους ορίζοντες στην αποδοτικότητα της μετάδοσης δεδομένων. Αυτή η καινοτομία έχει τη δυνατότητα να θέσει νέα πρότυπα στις δυνατότητες των ασύρματων δικτύων, προμηνύοντας ένα μέλλον με ακόμη πιο συνδεδεμένες και άμεσες ψηφιακές αλληλεπιδράσεις [20].

### Επικοινωνία συσκευής προς συσκευή (Device-to-Device - D2D)

Η D2D αποτελεί ακόμα μια κρίσιμη καινοτομία στα δίκτυα 5G, διευκολύνοντας έναν άμεσο "διάλογο" μεταξύ κοντινών κινητών συσκευών, παρακάμπτοντας τη συμβατική διαδρομή μέσω βάσεων σταθμών ή διαύλων του κεντρικού δικτύου. Αυτό το μοντέλο άμεσης αλληλεπίδρασης είναι καθοριστικό στην "ανακούφιση" της συμφόρησης του δικτύου, στη μείωση της καθυστέρησης και στην ενίσχυση της φασματικής αποδοτικότητας ανοίγοντας δρόμους για μια νέα κατηγορία εφαρμογών και υπηρεσιών peer-to-peer, προσανατολισμένων στην εγγύτητα. Επιτρέποντας στις συσκευές να επικοινωνούν ανεξάρτητα από την υποδομή του δικτύου, η επικοινωνία D2D όχι μόνο βελτιστοποιεί την αξιοποίηση των πόρων του δικτύου και βελτιώνοντας την εμπειρία των χρηστών ,αλλά επίσης προετοιμάζει το έδαφος για πιο ανθεκτικά και προσαρμόσιμα πλαίσια ασύρματων επικοινωνιών, ιδίως σε σενάρια όπου η συνδεσιμότητα του δικτύου είναι διαταραγμένη ή υπάρχει υπερβολική συμφόρηση [21].

# 

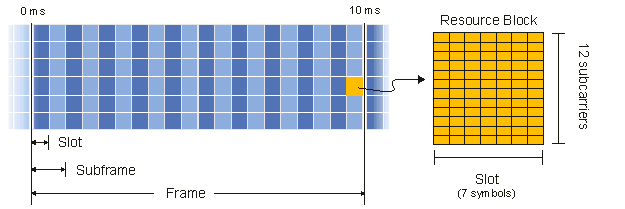
# Η κατανομή πόρων στα δίκτυα 5G

## Η ανάγκη για κατανομή πόρων στο 5G MIMO

Η κατανομή πόρων στα δίκτυα 5G αποτελεί ένα θεμελιώδες στοιχείο για την επιτυχία και τη βιωσιμότητα της πέμπτης γενιάς τεχνολογιών ασύρματων επικοινωνιών. Οι δυνατότητες του 5G να υποστηρίζει υψηλές ταχύτητες δεδομένων, ελάχιστη καθυστέρηση και μεγάλη αξιοπιστία χρειάζονται την αποτελεσματική κατανομή των διαθέσιμων πόρων, όπως το φάσμα ραδιοσυχνοτήτων, η ενέργεια και οι υπολογιστικές δυνατότητες. Χωρίς μια έξυπνη και δυναμική κατανομή πόρων, τα δίκτυα 5G δεν θα μπορούσαν να επιτύχουν τα υψηλά επίπεδα απόδοσης που απαιτούνται για κρίσιμες εφαρμογές. Για παράδειγμα, η τεχνολογία URLLC που απαιτείται για τον αυτόματο έλεγχο μηχανημάτων σε βιομηχανικά περιβάλλοντα ή για χειρουργικές επεμβάσεις με τηλεπικοινωνία, χρειάζεται εξαιρετικά χαμηλή καθυστέρηση και υψηλή αξιοπιστία, που μπορεί να επιτευχθεί μόνο μέσω της προσεκτικής κατανομής πόρων. Επιπλέον, οι πόροι στο 5G πρέπει να είναι ικανοί να προσαρμόζονται σε μεταβαλλόμενες συνθήκες και ανάγκες σε πραγματικό χρόνο, να ανταποκρίνονται σε διακυμάνσεις της κίνησης δεδομένων και στην ποικιλία των συνδεδεμένων συσκευών και υπηρεσιών. Αυτή η ευελιξία είναι απαραίτητη για τη διασφάλιση ότι όλες οι εφαρμογές λειτουργούν με βέλτιστη αποδοτικότητα και απόδοση.

## Η έννοια του RB

Ένα RB, βρίσκεται μέσα σε κάθε πλαίσιο πόρων (frame) και είναι η βασική μονάδα που παρέχεται σε έναν χρήστη στα 3GPP Long Term Evolution (LTE) δίκτυα και στα Νέα Ραδιοδίκτυα 5ης Γενιάς (5G New Radio - 5GNR). Στην Εικόνα 2 παρουσιάζεται η δομή αυτού του block η οποία περιλαμβάνει πολλαπλούς υπό-φορείς με σειριακή διάταξη, καθένας από τους οποίους περιέχει πληροφορίες πόρων. Ένα παράδειγμα είναι ότι στο LTE, ένα RB αποτελείται από 12 υπό-φορείς με απόσταση 15 Kilohertz (KHz) μεταξύ τους, καλύπτοντας συνολικά 180 KHz και έχει διάρκεια 0.5 milliseconds (ms). Στο 5G NR, η δομή των RBs παραμένει παρόμοια, αλλά υπάρχει μεγαλύτερη ευελιξία στην επιλογή των αποστάσεων μεταξύ των υπό-φορέων, κάτι που επιτρέπει τις συχνότητες 15 KHz, 30 KHz, 60 KHz, 120 KHz και 240 KHz.



**Εικόνα 2:** Η δομή ενός πλαισίου πόρων και ενός RB

Τα RBs έχουν ζωτική σημασία για την αποτελεσματικότητα ενός δικτύου, καθώς η δυναμική κατανομή τους στους χρήστες βελτιώνει τη φασματική απόδοση, την εμπειρία του χρήστη και τη δίκαιη κατανομή πόρων και οι διαχειριστές του δικτύου μπορούν μέσω προηγμένων αλγορίθμων που λαμβάνουν υπόψη παράγοντες όπως η ποιότητα του σήματος, η κινητικότητα του χρήστη, και η προτεραιότητα των υπηρεσιών, να μεγιστοποιήσουν τη φασματική απόδοση, να βελτιώσουν την εμπειρία του χρήστη, και να επιτύχουν δίκαιη κατανομή πόρων.

Η κατανομή των RBs επηρεάζει άμεσα την ποιότητα της υπηρεσίας που βιώνουν οι χρήστες που ζητούν υπηρεσίες υψηλής προτεραιότητας, όπως η Τηλεφωνία μέσω Διαδικτύου (Voice over Internet Protocol – VoIP) και Video Calls, που απαιτούν έγκαιρη και αξιόπιστη κατανομή πόρων για να ικανοποιήσουν τις αυστηρές απαιτήσεις τους σε εύρος ζώνης και ελάχιστη καθυστέρηση. Η διαχείριση των RBs παρουσιάζει αρκετές προκλήσεις, ιδιαίτερα σε πυκνά δίκτυα με υψηλές απαιτήσεις από τους χρήστες. Επιπλέον, η διαχείριση της παρεμβολής γίνεται κρίσιμη, καθώς τα γειτονικά RBs μπορεί να προκαλέσουν παρεμβολή μεταξύ τους, υποβαθμίζοντας την ποιότητα του σήματος. Προηγμένες τεχνολογίες όπως το MIMO και το beamforming στο 5G NR βοηθούν στην αντιμετώπιση αυτών των προκλήσεων, βελτιώνοντας την ποιότητα του σήματος και επιτρέποντας μια πιο αποδοτική χρήση των blocks.

## Τεχνικές κατανομής πόρων στα δίκτυα 5G

Η εφαρμογή τεχνικών κατανομής πόρων εξασφαλίζει την βέλτιστη χρήση του φάσματος, τη μείωση των καθυστερήσεων και τη βελτίωση της εμπειρίας του τελικού χρήστη, κρίσιμα στοιχεία για την επίτευξη των υψηλών προσδοκιών από τα δίκτυα πέμπτης γενιάς. Μερικές από τις τεχνικές κατανομής πόρων στα δίκτυα 5G, μπορούν να χωριστούν ανάλογα με τους πόρους τους οποίους διαχειρίζονται, όπως παρακάτω:

* **Δυναμική Πρόσβαση στο Φάσμα (Dynamic Spectrum Access - DSA) που** αναδεικνύεται ως καινοτόμος τεχνολογία στο πλαίσιο των δικτύων 5G, παρέχοντας έναν ευέλικτο μηχανισμό για την ανταπόκριση στις αυξημένες απαιτήσεις χρήσης του φάσματος. Χρησιμοποιώντας προηγμένους αλγορίθμους, το DSA διερευνά δυναμικά το φασματικό περιβάλλον και τη ζήτηση των χρηστών, προσαρμόζοντας την κατανομή των φασματικών πόρων με σκοπό τη μείωση των ζητημάτων που προκύπτουν από τη στατική κατανομή του φάσματος. Αυτή η προσέγγιση βοηθά στη μείωση των παρεμβολών και στην ενίσχυση της συνολικής απόδοσης του δικτύου.
* **Beamforming**, το οποίο στα συστήματα 5G MIMO, είναι μία διαδικασία που κατέχει κρίσιμο ρόλο. Χρησιμοποιώντας την Πληροφορία Κατάστασης Καναλιού (Channel; State Information - CSI), το σύστημα προσαρμόζει δυναμικά το σήμα, βελτιστοποιώντας την εστίαση του προς τους χρήστες και μειώνοντας ταυτόχρονα τις παρεμβολές. Η διαδικασία αυτή ενσωματώνει τον υπολογισμό βέλτιστων διανυσμάτων βάρους, ενισχύοντας τον λόγο σήματος προς θόρυβο και παρεμβολή (Signal-to-Interference-plus-Noise Ratio - SINR) και εξασφαλίζει ισορροπία μεταξύ απόδοσης και πολυπλοκότητας με τεχνικές όπως το υβριδικό beamforming.
* **Network Slicing**, το οποίο επιτρέπει τη δημιουργία προσαρμοσμένων τμημάτων δικτύου που ανταποκρίνονται σε συγκεκριμένες απαιτήσεις υπηρεσιών. Κάθε τμήμα διαχειρίζεται ξεχωριστά τους δικτυακούς πόρους και τις πολιτικές απομόνωσης, προσφέροντας βελτιστοποιημένες λύσεις για εφαρμογές eMBB, URLLC και mMTC. Οι μηχανισμοί slicing λειτουργούν με την υποστήριξη της Δικτύωσης Καθορισμένης από Λογισμικό (Software Defined Networking - SDN) και της Εικονικοποίησης Δικτυακής Λειτουργίας (Network Function Virtualization - NFV), ενισχύοντας την ευελιξία και την αποδοτικότητα του δικτύου.
* **Στρατηγική Ισορροπία Φορτίου** μεταξύ των κεραιών καθίσταται απαραίτητη για την ομοιόμορφη κατανομή της κίνησης και των πόρων στα δίκτυα 5G. Μέσω της εφαρμογής σύνθετων αλγορίθμων και τεχνικών, το δίκτυο διασφαλίζει την ισορροπημένη κατανομή των φορτίων, αποτρέποντας την υπερφόρτωση των κόμβων και εξασφαλίζοντας σταθερή QoS ακόμη και σε περιόδους αιχμής.
* **Η Διαχείριση και ο Συντονισμός των Παρεμβολών,** που είναιαπαραίτητα σε περιβάλλοντα υψηλής πυκνότητας και χρησιμοποιούνται με σκοπό την εξασφάλιση της αποδοτικής λειτουργίας του δικτύου. Χρησιμοποιώντας τεχνικές όπως η Συντονισμένη Πολυσημειακή τεχνική (Coordinated Multi-Point - CoMP) και τα χωρικά φίλτρα (spatial filtering), το δίκτυο βελτιώνει τη δυνατότητα ανίχνευσης των επιθυμητών σημάτων, μειώνοντας τις παρεμβολές και βελτιώνοντας την ποιότητα του σήματος.

## Προκλήσεις στην κατανομή πόρων στα δίκτυα 5G MIMO

Ένα από τα βασικά ζητήματα που αναδύονται αφορά τη διαχείριση της ενισχυμένης χωρικής κάλυψης που προσφέρεται από την τεχνολογία MIMO και ειδικότερα την Massive MIMO. Αυτή η τεχνολογία δημιουργεί μια σημαντική ευκαιρία για βελτίωση της φασματικής αποδοτικότητας και της δικτυακής χωρητικότητας, αλλά ταυτόχρονα ενσωματώνει ένα επίπεδο πολυπλοκότητας στην εκτίμηση του καναλιού, στην εκτίμηση της χρήσης του φάσματος από τους χρήστες και στην επεξεργασία των σημάτων.

Επιπροσθέτως, η ραγδαία μεταβαλλόμενη φύση των δικτύων 5G, με την κινητικότητα των χρηστών και τα μεταβαλλόμενα μοτίβα κίνησης, αντιπροσωπεύει μια ακόμη σημαντική πρόκληση. Οι μηχανισμοί κατανομής πόρων πρέπει να διαθέτουν την ευελιξία και την ικανότητα να προσαρμόζονται γρήγορα στις μεταβαλλόμενες συνθήκες του δικτυακού περιβάλλοντος.

Η ένταξη διάφορων δικτυακών τεχνολογιών και αρχιτεκτονικών εντός του 5G εισάγει περαιτέρω πολυπλοκότητα στην αποδοτική κατανομή πόρων, ενώ το network slicing αυξάνει την ανάγκη για στοχευμένες στρατηγικές που μπορούν να ικανοποιήσουν τις εξειδικευμένες απαιτήσεις κάθε κομματιού του δικτύου.

Επιπλέον, η ανάγκη για ενεργειακή αποδοτικότητα και η διαχείριση παρεμβολών σε πυκνά εγκατεστημένα δίκτυα απαιτούν την ανάπτυξη αλγορίθμων και στρατηγικών που θα μπορούν να διαχειριστούν αποτελεσματικά την πολυπλοκότητα και τη δυναμικότητα των περιβαλλόντων 5G. Οι στρατηγικές αυτές κατανομής πόρων θα πρέπει να οργανώνουν με σύνεση τις συχνότητες και τα χρονικά διαστήματα για να περιορίσουν τις παρεμβολές και να ενισχύσουν την ακεραιότητα και την ενεργειακή αποδοτικότητα του σήματος.

Συνεπώς, η αντιμετώπιση των πολυσύνθετων προκλήσεων της κατανομής πόρων στα περιβάλλοντα 5G MIMO είναι ουσιώδης για την απελευθέρωση των δυνατοτήτων αυτών των δικτύων. Οι μελλοντικές προσπάθειες έρευνας και ανάπτυξης θα πρέπει να επικεντρωθούν στη δημιουργία καινοτόμων, προσαρμοστικών και ενεργειακά αποδοτικών αλγορίθμων και στρατηγικών κατανομής πόρων που θα μπορούν να διαχειρίζονται αποτελεσματικά την πολυπλοκότητα και τη δυναμικότητα των περιβαλλόντων 5G, εξασφαλίζοντας τη βέλτιστη απόδοση του δικτύου και την ικανοποίηση των χρηστών.

# 

# ML στα δίκτυα 5G

## Η ανάγκη της ML στο 5G MIMO

Στο πλαίσιο των δικτύων 5G, η ενσωμάτωση αλγορίθμων ML μπορεί να βοηθήσει στην προσπάθεια για μία πρόοδο προς την επίτευξη ανεπτυγμένων επιπέδων δικτυακής νοημοσύνης και προσαρμοστικότητας στα περιβάλλοντα 5G MIMO. Η εφαρμογή της ML επιτρέπει μία «αλλαγή» στην προσέγγιση κατανομής πόρων, χρησιμοποιώντας προβλεπτικές αναλύσεις. Οι εν λόγω αλγόριθμοι έχουν τη δυνατότητα να προβλέψουν τις ανάγκες του δικτύου και των χρηστών αλλά και τις απαιτήσεις των υπηρεσιών, βελτιώνοντας την προληπτική διαμόρφωση και κατανομή των πόρων.

Μέσω της αξιοποίησης τόσο ιστορικών όσο και πραγματικού χρόνου δεδομένων, συμπεριλαμβανομένων της κινητικότητας των χρηστών, των αναγκών για υπηρεσίες και την ποιότητα του σήματος των χρηστών, οι αλγόριθμοι ML είναι ικανοί να αναγνωρίσουν μοτίβα και εξαρτήσεις που δεν είναι εύκολο να εντοπιστούν μέσω των παραδοσιακών ευρετικών μοντέλων. Αυτό επιτρέπει την πρόβλεψη των όγκων κίνησης, της κατανομής των χρηστών και των σημείων συμφόρησης του δικτύου με μεγάλη ακρίβεια.

Περαιτέρω, η ML ενισχύει τις προσαρμοστικές δυνατότητες του δικτύου, επιτρέποντας την επαναληπτική μάθηση και την εξειδίκευση των μοντέλων με σκοπό την ανταπόκρισή τους στα εξελισσόμενα «σενάρια» και στις διαφορετικές απαιτήσεις. Αυτή η συνεχής προσαρμογή εξασφαλίζει την διαρκή βελτιστοποίηση των στρατηγικών κατανομής πόρων, ευθυγραμμίζοντάς τες με τις δυναμικές ανάγκες των υπηρεσιών 5G.

Η ML όχι μόνο αυξάνει την αποδοτικότητα αλλά και προσδίδει ευελιξία στο δίκτυο, επιτρέποντας τον πραγματικού χρόνου επαναπροσδιορισμό των πόρων σύμφωνα με τις μεταβαλλόμενες απαιτήσεις και συνθήκες. Η ευελιξία αυτή είναι κρίσιμη για τη διατήρηση της βέλτιστης απόδοσης του δικτύου, μειώνοντας το λειτουργικό φόρτο και προσφέροντας συνεπή, υψηλής ποιότητας εμπειρία σε όλους τους χρήστες.

Συνολικά, η συνεργασία της ML με τις στρατηγικές κατανομής πόρων στα περιβάλλοντα 5G αποκαλύπτει μια νέα εποχή δικτυακής νοημοσύνης, υποσχόμενη βελτιωμένη αποδοτικότητα, προσαρμοστικότητα και μακροχρόνια ευελιξία του δικτύου. Η ενσωμάτωση της ML στη διαχείριση πόρων προσφέρει λοιπόν τη δυνατότητα στα δίκτυα 5G να ανταποκρίνονται με πιο δυναμικό και προσαρμοσμένο τρόπο στις ανάγκες και τις εξελίξεις, και συμβάλλοντας στην εκπλήρωση της "επανάστασης" του 5G.

## Τεχνικές κατανομής πόρων με χρήση ML

Η συνεισφορά των τεχνικών ML στη βελτίωση της κατανομής πόρων στα δίκτυα 5G έχει αποδειχθεί καταλυτική, αξιοποιώντας πολύπλοκες μεθόδους όπως η επιβλεπόμενη, η ημι-επιβλεπόμενη, η μη επιβλεπόμενη, και η RL. Αυτές οι προσεγγίσεις αντιμετωπίζουν πολυεπίπεδες προκλήσεις, ενισχύοντας την αποδοτικότητα και την ευελιξία του δικτύου.

* Χρησιμοποιώντας επισημασμένα δεδομένα, οι αλγόριθμοι επιβλεπόμενης μάθησης αποκτούν τη δυνατότητα να προβλέπουν την κίνηση του δικτύου και τη συμπεριφορά των χρηστών [22]. Πιο συγκεκριμένα :
  + Προβλεπτική Μοντελοποίηση: Αλγόριθμοι όπως η παλινδρόμηση, τα δέντρα αποφάσεων ή τα νευρωνικά δίκτυα μπορούν να προβλέψουν την κίνηση του δικτύου ή τη συμπεριφορά των χρηστών βάσει ιστορικών δεδομένων. Προβλέποντας με ακρίβεια τη ζήτηση, αυτά τα μοντέλα διασφαλίζουν ότι η χωρητικότητα του δικτύου συμφωνεί με τις ανάγκες των χρηστών.
  + Βελτιστοποίηση της QoS: Η επιβλεπόμενη ML μπορεί να βοηθήσει στην ταξινόμηση της κίνησης του δικτύου και στην πρόβλεψη των απαιτήσεων QoS διαφορετικών υπηρεσιών. Κατανοώντας αυτές τις απαιτήσεις, το δίκτυο μπορεί να κατανείμει πόρους με τέτοιον τρόπο ώστε να δίνεται προτεραιότητα σε κρίσιμες υπηρεσίες, μπορεί να βελτιώνει την εμπειρία του χρήστη και να μεγιστοποιεί τη συνολική απόδοση του δικτύου.
* Η RL είναι ιδιαίτερα κατάλληλη για δυναμικά και πολύπλοκα περιβάλλοντα όπως τα δίκτυα 5G, όπου το μοντέλο σταδιακά μαθαίνει τις καλύτερες ενέργειες που πρέπει να πραγματοποιηθούν σε διαφορετικές καταστάσεις μέσω της αλληλεπίδρασης με το περιβάλλον [23].
  + Δυναμική Κατανομή Πόρων: Η RL μπορεί να χρησιμοποιηθεί για να αναπτύξει πολιτικές που προσαρμόζουν δυναμικά την κατανομή πόρων ανταποκρινόμενες στις αλλαγές των συνθηκών του δικτύου και τις απαιτήσεις των χρηστών. Μαθαίνοντας συνεχώς από τις αλληλεπιδράσεις με το δίκτυο, οι παράγοντες RL μπορούν να εντοπίζουν βέλτιστες στρατηγικές κατανομής που μεγιστοποιούν την αποδοτικότητα του δικτύου και την ικανοποίηση των χρηστών.
  + Οργάνωση «Κομματιών» (slices) του Δικτύου: Στο πλαίσιο της οργάνωσης των network slices, όπου δημιουργούνται διαφορετικά εικονικά δίκτυα πάνω από μια κοινόχρηστη φυσική υποδομή, η RL μπορεί να βελτιστοποιήσει την κατανομή πόρων μεταξύ των κομματιών για να ικανοποιήσει διάφορες και μεταβαλλόμενες απαιτήσεις υπηρεσιών εξασφαλίζοντας ταυτόχρονα την αποδοτική χρήση των πόρων.
* Τεχνικές όπως η ημι-επιβλεπόμενη μάθηση και η μεταφορική μάθηση ή transfer learning μπορούν να βοηθήσουν επίσης. Αξιοποιώντας τόσο επισημασμένα όσο και μη επισημασμένα δεδομένα, η ημι-επιβλεπόμενη μάθηση αποτελεί έναν ιδανικό τρόπο για τη βελτίωση της ακρίβειας της μάθησης. Η μεταφορική μάθηση, από την άλλη πλευρά, ενισχύει την ταχύτητα προσαρμογής σε νέα σενάρια, μειώνοντας την ανάγκη για νέα επισημασμένα δεδομένα.
* Η βαθιά μάθηση ή deep learning, ιδιαίτερα στο πεδίο της αναγνώρισης περίπλοκων προτύπων και της πρόβλεψης κίνησης, αποδεικνύεται επίσης απαραίτητη εφόσον η τεχνική αυτή μπορεί να χρησιμοποιηθεί και για την εύρεση των βέλτιστων παραμέτρων, προσφέροντας αυξημένη ακρίβεια και βελτιστοποίηση στην κατανομή πόρων [24],[25].

Η ενσωμάτωση αυτών των προηγμένων τεχνικών ML στον τομέα της κατανομής πόρων του 5G καθιστά τα δίκτυα πιο δυναμικά, αποδοτικά και ικανά να ανταποκρίνονται στις εξελισσόμενες απαιτήσεις, προωθώντας έναν κόσμο όπου η προσαρμοστικότητα και η αξιοπιστία συναντούν την απαιτούμενη QoS. Οι προαναφερθείσες τεχνικές ML λοιπόν, μπορούν να χρησιμοποιηθούν για βελτιώσεις στα δίκτυα 5G, όπως :

* Πρόβλεψη Κίνησης και Ανίχνευση Ανωμαλιών: Βαθιά νευρωνικά δίκτυα μπορούν να αποκαλύψουν περίπλοκα μοτίβα στα δεδομένα κίνησης, βελτιώνοντας την ακρίβεια της πρόβλεψης ζήτησης και της ανίχνευσης ανωμαλιών, πράγμα που με τη σειρά του υποστηρίζει μια πιο «ανθεκτική» κατανομή πόρων [26].
* Βελτιστοποίηση Συστημάτων MIMO: Σε περιβάλλοντα MIMO, η βαθιά μάθηση μπορεί να βελτιστοποιήσει το beamforming και τη σύνδεση χρηστών, αντιμετωπίζοντας την υψηλή πολυπλοκότητα και τη δυναμική φύση αυτών των συστημάτων για να βελτιώσει την ποιότητα του σήματος, να μειώσει τις παρεμβολές και να μεγιστοποιήσει τη φασματική αποδοτικότητα.

## Η μη επιβλεπόμενη ML στην κατανομή πόρων

Η ML και ειδικότερα η μη επιβλεπόμενη μάθηση, χρησιμοποιείται για την οπτικοποίηση και βελτιστοποίηση της διανομής πόρων στα δίκτυα 5G. Η σύνθετη και δυναμική δομή αυτών των δικτύων, μαζί με την ποικιλομορφία των συνδεδεμένων συσκευών και τις διακυμάνσεις στις απαιτήσεις δεδομένων, απαιτεί την εφαρμογή σύνθετων μεθόδων. Αυτές οι μέθοδοι πρέπει να είναι σε θέση να προσαρμόζονται και να βελτιστοποιούνται σε πραγματικό χρόνο, προκειμένου να αποκαλύπτουν μοτίβα χωρίς γνώση από προκαθορισμένες ετικέτες, ερευνώντας και ύστερα αποκαλύπτοντας την ύπαρξη συστάδων μεταξύ όμοιων δεδομένων ως προς τα επιλεγμένα χαρακτηριστικά.

Η τεχνική της ομαδοποίησης/συσταδοποίησης αναδεικνύεται ως κύρια μέθοδος μη επιβλεπόμενης μάθησης στην κατανομή πόρων 5G, με τις μεθόδους K-Means, HC (συσσωματωτική ή διαιρετική) και DBSCAN να αποτελούν ενδεικτικά παραδείγματα. Αυτοί οι αλγόριθμοι μπορούν να κατατάξουν χρήστες ή δικτυακούς κόμβους σε ξεχωριστές ομάδες, βάσει των ομοιοτήτων τους σε διάφορα χαρακτηριστικά όπως οι απαιτήσεις τους, τα μοτίβα κινητικότητας ή η ποιότητα της σύνδεσης και ο καθένας εκ των αλγορίθμων αυτών ομαδοποιεί τα δεδομένα χρησιμοποιώντας τον δικό του αλγόριθμο. Η εφαρμογή αυτών των τεχνικών επιτρέπει την εξατομικευμένη προσαρμογή των στρατηγικών κατανομής πόρων, αυξάνοντας την αποδοτικότητα του δικτύου.

Ενδεικτικά, η μέθοδος K-Means μπορεί να χρησιμοποιηθεί για την ταξινόμηση χρηστών με βάση γεωγραφικά κριτήρια ή απαιτήσεις σε πόρους και χρησιμοποιεί σημεία "κέντρα" για την ομαδοποίηση. Σαν αλγόριθμος, μπορεί να χρησιμοποιηθεί για να διευκολύνει την αποδοτική διάθεση δικτυακών πόρων, όπως τα RBs, σε συστήματα MIMO. Η HC προσφέρει αναλύσεις σε διάφορες κλίμακες και μας βοηθά να οπτικοποιήσουμε τον τρόπο με τον οποίο δημιουργούνται οι συστάδες μέσω ενός δενδρογράμματος, ανοίγοντας τον δρόμο για λεπτομερείς και πολυεπίπεδες στρατηγικές ομαδοποίησης που μπορούν να προσαρμόζονται σε όλες τις δικτυακές συνθήκες, ενώ το DBSCAN συσταδοποιεί τα δεδομένα, δηλαδή τους χρήστες, ανάλογα με το πόσο πυκνά βρίσκονται μεταξύ τους.

Η μείωση διαστατικότητας, συμπληρωματική στην ομαδοποίηση, απλοποιεί τα περίπλοκα δεδομένα υψηλής διάστασης σε πιο διαχειρίσιμες μορφές. Τεχνικές όπως η PCA και η t-Distributed Stochastic Neighbor Embedding (t-SNE), αποκαλύπτουν τις βασικές δομές και επιτρέπουν την αναγνώριση των κύριων χαρακτηριστικών των δεδομένων[[6]](#footnote-6).

Η ανίχνευση ανωμαλιών, από την άλλη πλευρά, είναι κρίσιμη για τη διατήρηση της ακεραιότητας και της απόδοσης του δικτύου, με τη μη επιβλεπόμενη μάθηση να προσφέρει έναν ισχυρό μηχανισμό για την έγκαιρη αναγνώριση και αντιμετώπιση τέτοιων προκλήσεων, ομαδοποιώντας τις ανωμαλίες αυτές.

Η μη επιβλεπόμενη μάθηση επίσης, παρέχει μια εκτεταμένη εργαλειοθήκη για την αντιμετώπιση των πολύπλοκων προκλήσεων στα δίκτυα 5G, υποστηρίζοντας ένα πιο αυτόνομο, ευέλικτο και αποδοτικό δίκτυο. Η συνεχής έρευνα και ανάπτυξη σε αυτούς τους τομείς είναι κρίσιμη, αφού σκοπός τους είναι η ενίσχυση της έξυπνης διαχείρισης και η προσαρμοστικότητα των δικτύων 5G.

## Οι προκλήσεις στη χρήση μη επιβλεπόμενης ML στο 5G MIMO

Η μη εποπτευόμενη ML ξεδιπλώνει μια πλειάδα προκλήσεων στη διαχείριση πόρων δικτύων 5G MIMO. Οι εν λόγω προκλήσεις καθίστανται πιο εμφανείς εξαιτίας της εγγενούς πολυπλοκότητας και δυναμικότητας των δικτύων 5G, όπως και λόγω της απαίτησης για υψηλού επιπέδου απόδοσης και αποδοτικότητας. Μερικές από τις προκλήσεις παρουσιάζονται παρακάτω.

Πρωτίστως, η σωστή επιλογή αλγορίθμου μη επιβλεπόμενης ML κρίνεται ζωτικής σημασίας. Αλγόριθμοι όπως η ομαδοποίηση και η μείωση διαστατικότητας είναι καθοριστικοί για την εξαγωγή νοήματος από αταξινόμητα δεδομένα, τα οποία συναντώνται συχνά σε δίκτυα 5G. Ωστόσο, η επιλογή ενός μη κατάλληλου αλγορίθμου μπορεί να οδηγήσει σε λανθασμένες ερμηνείες και αστοχίες στην κατανομή πόρων.

Δεύτερον, η δυναμική ιδιότητα των δικτύων 5G απαιτεί την ευελιξία και τη γρήγορη προσαρμογή των μοντέλων. Η συνεχής ενημέρωση των μοντέλων με τα τελευταία δεδομένα και η ανταπόκριση στις μεταβολές του δικτύου προϋποθέτουν την ύπαρξη αλγορίθμων με γρήγορους χρόνους εκπαίδευσης και εγγυημένη ακρίβεια.

Τρίτον, η ετερογένεια των σεναρίων χρήσης και η ανάγκη για ισορροπημένη κατανομή πόρων κατά μήκος πολλαπλών διαστάσεων επιβάλλουν πρόσθετη πολυπλοκότητα στην εφαρμογή αλγορίθμων μη επιβλεπόμενης ML. Αντιμετωπίζοντας αντίθετα κριτήρια όπως η μέγιστη απόδοση με ταυτόχρονη ελαχιστοποίηση της κατανάλωσης ενέργειας, τα διλήμματα αυτά δεν είναι πάντα εύκολο να επιλυθούν.

Οι προκλήσεις ασφαλείας και ιδιωτικότητας δεν πρέπει να υποτιμηθούν. Η προστασία των δεδομένων και η αποτροπή επιθετικών ενεργειών μέσω των μοντέλων μη εποπτευόμενης ML στα δίκτυα 5G αποτελούν ουσιώδη ζητήματα [27].

Συνολικά, η επιτυχία στην εφαρμογή της μη εποπτευόμενης ML στη διαχείριση πόρων στα συστήματα 5G MIMO απαιτεί λεπτομερή εξέταση των ιδιαιτεροτήτων του δικτύου και συνεχή προσαρμογή και βελτιστοποίηση των αλγορίθμων, ώστε να ανταποκρίνονται στις διάφορες προκλήσεις του περιβάλλοντος.

# 

# Πειραματική διαδικασία

## Εισαγωγή στην πειραματική διαδικασία

Η πειραματική διαδικασία που παρουσιάζεται σε αυτήν την εργασία επικεντρώνεται στην εφαρμογή και αξιολόγηση τεχνικών μη εποπτευόμενης ML για την ομαδοποίηση χρηστών σε δίκτυα 5G, συνδυάζοντας τη με δυναμική κατανομή πόρων. Η ανάγκη για προηγμένες τεχνικές διαχείρισης πόρων είναι έντονη στα δίκτυα 5G, όπου απαιτούνται εξελιγμένες και αποδοτικές λύσεις.

Για την διεξαγωγή των πειραμάτων χρησιμοποιήθηκε γλώσσα προγραμματισμού Python, έκδοση 3.11.9. Η γλώσσα αυτή επιλέχθηκε λόγω του ότι διαθέτει μια μεγάλη βιβλιοθήκη εργαλείων και πακέτων που διευκολύνουν την ανάπτυξη και την εφαρμογή αλγορίθμων ML ενώ ταυτόχρονα προσφέρει απλότητα και ευκολία χρήσης.

Στο πλαίσιο αυτής της εργασίας, χρησιμοποιήθηκε ο αλγόριθμος συσταδοποίησης K-Means αλλά και άλλοι για την ομαδοποίηση χρηστών βάσει χαρακτηριστικών, όπως οι απαιτήσεις τους σε RBs, η μέγιστη αποδεκτή καθυστέρηση της υπηρεσίας που ζητούν, η πιο κοντινή (σε απόσταση) κεραία σε αυτούς, η ποιότητα του σήματος και το SNR. Η ομαδοποίηση αυτή είναι κρίσιμη για την αναγνώριση των ομάδων χρηστών που πιθανώς θα υποστούν υποβάθμιση στην ποιότητα της υπηρεσίας στην περίπτωση που άλλοι χρήστες χρειαστούν πόρους που οι δύο κεραίες δεν διαθέτουν. Είναι όμως χρήσιμη και για την πιο σωστή και εύκολη διαχείριση των χρηστών ως data points. Οι αλγόριθμοι ύστερα βελτιστοποιούνται με τις κατάλληλες διαδικασίες και τα αποτελέσματα των αλγορίθμων συγκρίνονται για να διαπιστωθεί εκείνος που δίνει την καλύτερη ομαδοποίηση. Επίσης, εξετάζεται η απόδοση της δυναμικής κατανομής των πόρων μέσω πειραμάτων που μετρούν την ικανότητα του δικτύου να διαχειρίζεται το φορτίο αποδοτικά και αυτή η απόδοση συγκρίνεται με την απλή μέθοδο κατανομής πόρων που χρησιμοποιεί μόνο το βέλτιστο SNR για να κατανείμει τους χρήστες στις πιο κοντινές τους κεραίες.

Από τα αποτελέσματα των πειραμάτων εξάγονται σημαντικά συμπεράσματα σχετικά με τη βελτίωση των τεχνικών κατανομής των πόρων αλλά και σχετικά με την εφαρμογή των μεθοδολογιών μη επιβλεπόμενης ML στα σύγχρονα δίκτυα 5G. Το κεφάλαιο αυτό δείχνει το πώς οι στρατηγικές μη επιβλεπόμενης ML μπορούν να συμβάλλουν στην ενίσχυση της απόδοσης του δικτύου, βελτιστοποιώντας την κατανομή των πόρων σε μια περιοχή όπου οι χρήστες έχουν διαφορετικές απαιτήσεις, κατανομές και ανάγκες.

## Φυσική σημασία των χαρακτηριστικών

Παρακάτω, παρουσιάζεται η ανάγκη και η φυσική σημασία των χαρακτηριστικών που χρησιμοποιήθηκαν στην πειραματική διαδικασία :

* **Η καθυστέρηση** αντιπροσωπεύει τον χρόνο που απαιτείται για τα δεδομένα να ταξιδέψουν από την πηγή τους στον προορισμό τους και πίσω, συχνά αναφερόμενο ως "ping" ή "lag". Στον τομέα του 5G, όπου το URLLC είναι κύριο συστατικό, η χαμηλή καθυστέρηση είναι συνώνυμη με σχεδόν άμεσους χρόνους ανταπόκρισης. Στο σενάριο αυτό, η χαμηλή καθυστέρηση είναι ζωτικής σημασίας για την επίτευξη της απαιτούμενης ανταπόκρισης, διασφαλίζοντας ότι οι υπηρεσίες που είναι ευαίσθητες στην καθυστέρηση δέχονται την απαιτούμενη προσοχή όσον αφορά την καθυστέρηση. Η σημασία της καθυστέρησης επεκτείνεται και στην εμπειρία του χρήστη, όπου ακόμη και μικρές βελτιώσεις μπορούν να ενισχύσουν σημαντικά την απόδοση των εφαρμογών.
* **Η απαίτηση εύρους ζώνης** αναφέρεται στο πλάτος του καναλιού επικοινωνίας, το οποίο καθορίζει τον μέγιστο ρυθμό δεδομένων που μπορεί να μεταδοθεί μέσω του δικτύου. Μετριέται σε bits per second (bps) και, σε πρακτικούς όρους, αφορά το πόσα δεδομένα μπορούν να σταλούν ή να ληφθούν από έναν χρήστη σε οποιαδήποτε στιγμή. Σε ένα σενάριο 5G MIMO, η επαρκής κατανομή εύρους ζώνης είναι ουσιαστική για την υποστήριξη των υψηλών ρυθμών δεδομένων που απαιτούνται από προηγμένες εφαρμογές, όπως η ροή βίντεο υψηλής ευκρίνειας, οι εκτεταμένες αναπτύξεις IoT και οι ολοκληρωμένες υπηρεσίες cloud. Οι απαιτήσεις εύρους ζώνης διαφέρουν ανά εφαρμογή, με τις πιο απαιτητικές υπηρεσίες να χρειάζονται υψηλότερο εύρος ζώνης για να λειτουργήσουν αποτελεσματικά. Στην εργασία αυτή, η ικανοποίηση των απαιτήσεων εύρους ζώνης είναι κλειδί για την επίδειξη του πώς τα RBs μπορούν να κατανεμηθούν αποτελεσματικά για να ικανοποιηθούν οι διάφορες ανάγκες μιας ευρείας γκάμας υπηρεσιών και εφαρμογών εντός του δικτύου 5G.
* **Η απόσταση από έναν σταθμό** είναι κρίσιμη για την κατανόηση της δυναμικότητας ενός δικτύου 5G. Στα περιβάλλοντα MIMO, η απόσταση επηρεάζει άμεσα την ποιότητα του σήματος που λαμβάνεται, καθώς τα σήματα που διανύουν μεγαλύτερες αποστάσεις είναι πιο επιρρεπή σε απώλειες λόγω διάδοσης, ενδιάμεσων εμποδίων και παραμόρφωσης.
* **Η απώλεια διαδρομής, δηλαδή το pathloss**, είναι μια ζωτική παράμετρος που περιγράφει την εξασθένιση της ισχύος του σήματος καθώς αυτό διανύει την απόσταση από τον σταθμό βάσης προς τον χρήστη. Σε περιβάλλοντα 5G MIMO, η κατανόηση της απώλειας διαδρομής είναι κρίσιμη για τον σχεδιασμό του δικτύου και τη βελτιστοποίηση της κάλυψης και της χωρητικότητας [28].
* **Η ισχύς του σήματος** αποτελεί μια κρίσιμη μετρική που καθορίζει την ικανότητα του χρήστη να λαμβάνει δεδομένα. Υψηλότερες τιμές ισχύος σηματοδοτούν ισχυρότερη και πιο αξιόπιστη σύνδεση, ενώ χαμηλές τιμές μπορεί να οδηγήσουν σε διακοπές και χαμηλή ποιότητα επικοινωνίας.
* **Το SNR** είναι ένας δείκτης της ποιότητας του σήματος σε σχέση με το υπόβαθρο θορύβου. Υψηλότερες τιμές SNR σημαίνουν ότι το σήμα είναι πιο «καθαρό» από θορύβους, παρέχοντας βελτιωμένη ποιότητα επικοινωνίας και αυξημένη αξιοπιστία.
* **Η ποιότητα σήματος** σε ένα περιβάλλον 5G MIMO είναι ενδεικτική της ακεραιότητας και της σαφήνειας του επικοινωνιακού σήματος καθώς αυτό λαμβάνεται από τις συσκευές των χρηστών. Αποτελεί συνδυασμό αρκετών παραγόντων, συμπεριλαμβανομένων της ισχύος του σήματος και του λόγου σήματος προς θορύβου, μετρικές που επηρεάζουν την ικανότητα του χρήστη να διατηρήσει μια αξιόπιστη και σταθερή σύνδεση. Η καλή ποιότητα σήματος είναι κρίσιμη για την ελαχιστοποίηση της απώλειας πακέτων, τη μείωση της καθυστέρησης και τη μεγιστοποίηση των ρυθμών μετάδοσης δεδομένων, επηρεάζοντας άμεσα την αποδοτικότητα της κατανομής RBs. Για τους χρήστες, αυτό μεταφράζεται σε πιο καθαρές φωνητικές κλήσεις, ομαλότερη ροή βίντεο και πιο αξιόπιστη συνδεσιμότητα στο διαδίκτυο.
* **Οι απαιτήσεις σε Megahertz (MHz) και KHz** αναφέρονται στο πλάτος ζώνης που απαιτείται για την υποστήριξη διαφόρων εφαρμογών και υπηρεσιών σε ένα δίκτυο 5G.
* **Τα RBs** χρησιμοποιούνται για την κατανομή των πόρων στους χρήστες, επηρεάζοντας άμεσα την απόδοση του δικτύου και την εμπειρία του χρήστη.

## Οι υπηρεσίες και τα χαρακτηριστικά τους

Επιπλέον, οι υπηρεσίες και οι απαιτήσεις τους είναι οι ακόλουθες:

* **Οι κλήσεις VoIP** αποτελούν μια βασική υπηρεσία στον κόσμο των τηλεπικοινωνιών, επιτρέποντας τη φωνητική επικοινωνία μέσω του διαδικτύου. Για την υποστήριξη αυτής της υπηρεσίας απαιτείται εύρος ζώνης 1 Mbps, προκειμένου να εξασφαλιστεί καθαρή και συνεχής μετάδοση φωνής χωρίς διακοπές ή απώλειες ποιότητας. Οι χαμηλότερης ποιότητας κλήσεις VoIP χρειάζονται 0.1 Mbps.
* **Η ακρόαση podcasts**, η οποία έχει γίνει εξαιρετικά δημοφιλής, απαιτεί εύρος ζώνης 2 Mbps για να παρέχει στους χρήστες τη δυνατότητα να απολαύσουν υψηλής ποιότητας ηχητικό περιεχόμενο. Αυτό το εύρος ζώνης επιτρέπει την ομαλή ροή δεδομένων για την ακρόαση εκπομπών χωρίς προβλήματα φόρτωσης ή διακοπής. Τα podcasts μπορούν να χρειάζονται και 128Kbps «θυσιάζοντας» την υψηλή ποιότητα.
* **Για την περιήγηση στον ιστό**, 5 Mbps είναι το απαιτούμενο εύρος ζώνης για να φορτώνουν γρήγορα οι ιστοσελίδες και να διασφαλιστεί μια ομαλή εμπειρία χρήστη. Αλλά η περιήγηση στον ιστό μπορεί να γίνει και με 1 Mbps σε χαμηλότερη ποιότητα. Αυτό το εύρος ζώνης υποστηρίζει την πλοήγηση σε περιεχόμενο και διαδραστικές εφαρμογές χωρίς καθυστερήσεις.
* **Η προβολή περιεχομένου HDTV** απαιτεί εύρος ζώνης 16 Mbps για να παρέχει στους χρήστες υψηλής ποιότητας εικόνα σε υψηλή ευκρίνεια. Αυτή η απαίτηση εξασφαλίζει ότι τα βίντεο και οι ταινίες μεταδίδονται ομαλά, προσφέροντας μια μοναδική εμπειρία παρακολούθησης. Για χαμηλότερη ποιότητα χρειάζονται 4 Mbps ώστε να υπάρχει ανάλυση 720p.
* **Για τη ροή περιεχομένου σε 4K** ανάλυση, απαιτούνται 25 Mbps. Αυτό το εύρος ζώνης είναι απαραίτητο για να υποστηριχθεί η μετάδοση υψηλότερης ποιότητας εικόνας με περισσότερες λεπτομέρειες, προσφέροντας μια εξαιρετικά πλούσια οπτική εμπειρία. Για χαμηλότερης ποιότητας streaming, χρειάζονται 8 Mbps.
* **Οι κλήσεις βίντεο** απαιτούν εύρος ζώνης 3-5 Mbps (στο σενάριο επιλέχθηκαν τα 4Mbps απαίτηση) για να εξασφαλιστεί ότι η επικοινωνία είναι ομαλή, με καθαρή εικόνα και συγχρονισμένο ήχο. Αυτή η απαίτηση εύρους ζώνης είναι ζωτικής σημασίας για την παροχή μιας ευχάριστης και απρόσκοπτης εμπειρίας κατά τη διάρκεια τηλεδιασκέψεων ή προσωπικών κλήσεων. Για χαμηλότερης ποιότητας video conference απαιτείται 1 Mbps.

Οι τυπικές μέγιστες αποδεκτές καθυστερήσεις για τις εφαρμογές που αναφέρθηκαν και χρησιμοποιήθηκαν ήταν :

* Κάτω από 150 ms είναι επιθυμητό για κλήσεις VoIP καλής ποιότητας. Καθυστερήσεις κάτω από 100 ms θεωρούνται εξαιρετικές, παρέχοντας σχεδόν πραγματικού χρόνου επικοινωνία χωρίς αισθητή καθυστέρηση. Επιλέχθηκε 150 ms.
* Παρόμοια με τις κλήσεις VoIP, οι κλήσεις βίντεο απαιτούν καθυστερήσεις κάτω από 150 ms για να διατηρηθεί μια ομαλή και φυσική ροή συνομιλίας. Όσο χαμηλότερο, τόσο το καλύτερο, με καθυστερήσεις κάτω από 100 ms να είναι ιδανικές για τον συγχρονισμό του βίντεο με τον ήχο αποτελεσματικά. Επιλέχθηκε 150 ms.
* Η περιήγηση στον ιστό είναι περισσότερο ανεκτική στην καθυστέρηση. Καθυστερήσεις έως 200-300 ms μπορούν να είναι αποδεκτές, αν και χαμηλότερες καθυστερήσεις (κάτω από 100 ms) βελτιώνουν σημαντικά την εμπειρία του χρήστη κάνοντας τη φόρτωση των σελίδων να φαίνεται ακαριαία. Επιλέχθηκε 300 ms.
* Καθώς τα podcasts περιλαμβάνουν ροή ηχητικού περιεχομένου που δεν είναι ζωντανό, καθυστερήσεις έως 300-500 ms είναι γενικά αποδεκτές. Επιλέχθηκε 500 ms.
* Για τη ροή βίντεο περιεχομένου όπως HDTV και ιδιαίτερα 4K, αρχικές καθυστερήσεις buffering έως 1-2 δευτερόλεπτα (1000-2000 ms) μπορούν να είναι αποδεκτές (γενικά κάτω από 5sec), καθώς το περιεχόμενο δεν είναι ζωντανό και η έμφαση βρίσκεται στη συνεχή αναπαραγωγή χωρίς διακοπές. Χαμηλότερη καθυστέρηση βελτιώνει την εμπειρία κατά την έναρξη της ροής ή την αναζήτηση. Επιλέχθηκε 1000 ms (1sec)[[7]](#footnote-7).

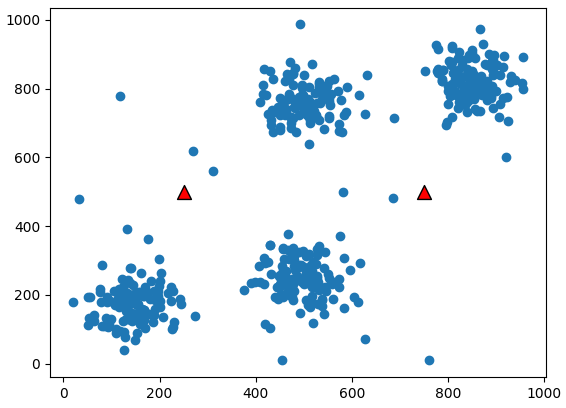
## Περιγραφή του περιβάλλοντος και η μηχανική των χαρακτηριστικών

Η συχνότητα λειτουργίας που χρησιμοποιήθηκε ήταν Μεσαίου Εύρους Φάσμα, 2.5GHz, καθώς οι συχνότητες μεσαίου φάσματος στα δίκτυα 5G προσφέρουν μια εξαιρετική ισορροπία μεταξύ εύρους κάλυψης και ικανότητας διείσδυσης σε φυσικά και ανθρωπογενή εμπόδια, όπως κτίρια και δέντρα. Αυτή η ισορροπία καθιστά τις μεσαίες συχνότητες ιδανικές για την παροχή σταθερής και αξιόπιστης κάλυψης σε διάφορα περιβάλλοντα, ενισχύοντας τη συνδεσιμότητα.

Στον τομέα της χωρητικότητας και της ταχύτητας, το μεσαίο φάσμα υποστηρίζει υψηλότερους ρυθμούς δεδομένων σε σύγκριση με τις χαμηλότερες συχνότητες, ικανοποιώντας ευρύ φάσμα υπηρεσιών από απλή περιήγηση στο διαδίκτυο έως ροή βίντεο υψηλής ευκρίνειας. Αυτό εγγυάται ότι το δίκτυο μπορεί να παρέχει υψηλής ποιότητας υπηρεσίες, ανταποκρινόμενο στις απαιτήσεις σύγχρονων εφαρμογών.

Όσον αφορά την αποδοτικότητα του δικτύου, το μεσαίο φάσμα διασφαλίζει την αποδοτική χρήση του φάσματος, προσφέροντας επαρκή ποσότητα διαθέσιμων RBs για δυναμική κατανομή στους χρήστες. Αυτό επιτρέπει στο δίκτυο να διαχειρίζεται αποτελεσματικά την κίνηση και να παρέχει ομαλή πρόσβαση σε πολλαπλές υπηρεσίες ταυτόχρονα, βελτιώνοντας τη συνολική εμπειρία των χρηστών.

Η καθυστέρηση στο μεσαίο φάσμα είναι επίσης σχετικά χαμηλή, κρίσιμη για υπηρεσίες πραγματικού χρόνου όπως VoIP ή κλήσεις βίντεο.



**Εικόνα 3:** Γράφημα της περιοχής του πειράματος, των κεραιών και των χρηστών

Το περιβάλλον του πειράματος όπως φαίνεται στην Εικόνα 3, (όπου τα κόκκινα σημάδια αναπαριστούν τις κεραίες και τα μπλε τους χρήστες) απαρτίζεται από δύο κεραίες τύπου macrocell με χωρητικότητα 400MHz η καθεμία και ισχύς μετάδοσης 45 dBm, καθώς και 500 χρήστες τοποθετημένους σε μια περιοχή 1000m x 1000m. Τα αρχικά χαρακτηριστικά των χρηστών του πειράματος είναι τα αναγνωριστικά των χρηστών (User Ids), οι γεωγραφικές τους συντεταγμένες, η υπηρεσία που ζητούν και η μέγιστη αποδεκτή καθυστέρηση της υπηρεσίας. Oι γεωγραφικές συντεταγμένες των χρηστών, που αναπαρίστανται ως (x, y), είναι απαραίτητες για τη χωρική ανάλυση και παρέχουν μια βάση για την κατανόηση της φυσικής διάταξης των χρηστών εντός του δικτύου. Αυτά τα χωρικά δεδομένα είναι κρίσιμα για την κατανομή των RBs άρα και τη βελτιστοποίηση της ισορροπίας φόρτου του δικτύου καθώς βοηθούν στον υπολογισμό νέων χαρακτηριστικών των χρηστών όπως το SNR. Οι τιμές των συντεταγμένων x και y είναι τιμές γεωγραφικού πλάτους και μήκους αντίστοιχα. Οι περισσότεροι χρήστες ακολουθούν κανονική κατανομή γύρω από 4 “hotspots” ή αλλιώς σημεία συγκέντρωσης ενώ οι υπόλοιποι ακολουθούν ομοιόμορφη κατανομή σε όλη την περιοχή. Οργανώνοντας τους χρήστες με αυτόν τον τρόπο, επιτυγχάνεται η προσομοίωση ενός συνηθισμένου περιβάλλοντος αστικής περιοχής με μερικά σημεία συγκέντρωσης αλλά και χρήστες που βρίσκονται μακριά από αυτά.

Η απαίτηση των χρηστών σε Mbps προέρχεται από την υπηρεσία που ζητά ο χρήστης και αντικατοπτρίζει άμεσα τον ρυθμό δεδομένων που απαιτείται από κάθε χρήστη για την πρόσβαση σε διάφορες υπηρεσίες, από εφαρμογές χαμηλού εύρους ζώνης όπως το email μέχρι υπηρεσίες που απαιτούν πολύ μεγαλύτερο εύρος ζώνης όπως η ροή βίντεο 4K. Αυτό το χαρακτηριστικό είναι κρίσιμο για την κατανόηση των απαιτήσεων που τίθενται στο δίκτυο, καθοδηγώντας τα μοντέλα μη επιβλεπόμενης μάθησης να ομαδοποιήσουν τους χρήστες αποτελεσματικά.

Η μέγιστη αποδεκτή καθυστέρηση, είναι ένα κρίσιμο στοιχείο απόδοσης στα δίκτυα 5G, ιδιαίτερα για εφαρμογές πραγματικού χρόνου. Στη διπλωματική αυτή εργασία, οι καθυστερήσεις σε ms είναι αναγκαίες για την αξιολόγηση του πώς επηρεάζεται η εμπειρία του χρήστη. Η προσθήκη των καθυστερήσεων των υπηρεσιών των χρηστών ως χαρακτηριστικό είναι επίσης χρήσιμο και για να αναδειχθεί το πώς η μη επιβλεπόμενη μάθηση και οι αλγόριθμοι κατανομής πόρων μπορούν να τις χειριστούν.

Στην έρευνα αυτή, τα User IDs δεν χρησιμοποιούνται στον αλγόριθμο συσταδοποίησης, αλλά υπάρχουν για να παρακολουθείται ο κάθε χρήστης "ξεχωριστά" και λειτουργούν ως απλά αναγνωριστικά των χρηστών. Ενσωματώνοντας το User ID, η εργασία αναγνωρίζει τη σημασία των ατομικών εμπειριών χρήστη στη συνολική αξιολόγηση της απόδοσης δικτύου σε περιβάλλοντα 5G MIMO.

Μέσω της προσεκτικής επιλογής, δημιουργίας και επεξεργασίας χαρακτηριστικών, μπορούν να εξαχθούν πολύτιμες πληροφορίες από τα δεδομένα, οδηγώντας σε πιο στοχευμένες και αποδοτικές λύσεις. Στην περίπτωση της ανάλυσης δικτύων 5G, η σωστή μηχανική των χαρακτηριστικών, το σωστό feature engineering δηλαδή, επιτρέπει την ακριβή κατανομή πόρων, βελτιστοποιώντας την απόδοση του δικτύου και την εμπειρία των χρηστών.

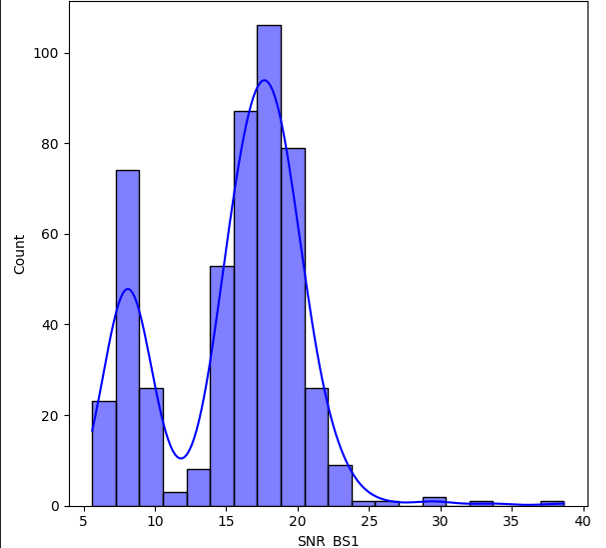
Τα βήματα που ακολουθούνται στη φάση της μηχανικής των χαρακτηριστικών (feature engineering) για τον υπολογισμό των νέων χαρακτηριστικών των χρηστών είναι τα παρακάτω:

* Χρησιμοποιείται το μοντέλο Hata-Okumura για αστικές περιοχές, για να εκτιμηθεί η απώλεια διαδρομής (Pathloss) για να κατανοηθεί το πώς η ισχύς του σήματος μειώνεται λόγω της απόστασης και των περιβαλλοντικών παραγόντων μεταξύ του βασικού σταθμού και κάθε χρήστη. Αυτό το μοντέλο είναι ιδιαίτερα κατάλληλο για την πρόβλεψη της μέσης απώλειας διαδρομής σε αστικά περιβάλλοντα, λαμβάνοντας υπόψη το ύψος των κεραιών του σταθμού βάσης (50 μέτρα) και των χρηστών (1.5 μέτρα), καθώς και τη συχνότητα λειτουργίας (2500 MHz / 2.5 GHz). Ο τύπος του μοντέλου Hata-Okumura για αστικές περιοχές παρουσιάζεται παρακάτω :

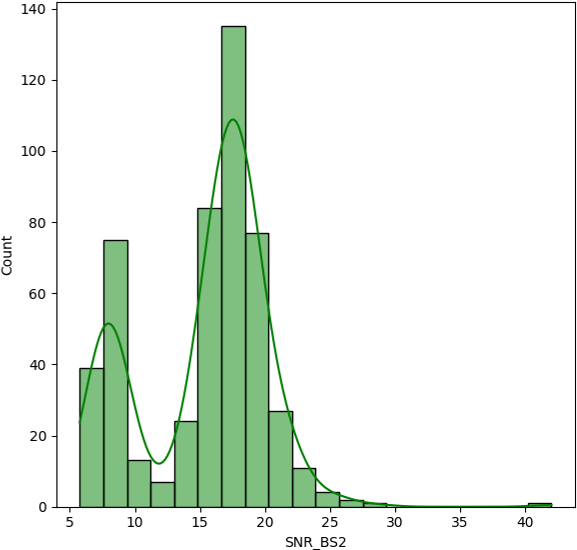
* Αφού καθοριστεί η απώλεια διαδρομής, το επόμενο βήμα είναι να υπολογιστεί η ισχύς του σήματος που λαμβάνεται από κάθε χρήστη. Αυτό επιτυγχάνεται εφαρμόζοντας τον τύπο Pr = Pt – PL + Gains − Losses. Η ισχύς μετάδοσης (Pt) του βασικού σταθμού ορίζεται στα 45 dBm, αντανακλώντας την τυπική ισχύ για σταθμούς macrocell. Ως Gains θεωρούνται τα κέρδη του σταθμού (21 dbi ως macrocell) και των συσκευών του χρήστη (συνήθως 0 μέχρι 5 dbi , επιλέχθηκε 2 dbi) και ως losses οι «ήττα» του σήματος λόγω καλωδίων (συνήθως 0 με 5 dbi , επιλέχθηκε 2 dbi) και λόγω διείσδυσης των κτηρίων (συνήθως 10 με 20dbi, επιλέχθηκε 15 dbi).
* Υπολογισμός θορύβου : Ο θόρυβος είναι ένας θεμελιώδης δείκτης για την αξιολόγηση της ποιότητας ασύρματης επικοινωνίας, υπολογίζεται με τον τύπο Noise = − 174 + 10\*log10(Bandwidth). Για ένα κανάλι πλάτους ζώνης 400 MHz. Ο υπολογισμός του θορύβου παρέχει μια βάση για την αξιολόγηση του θορύβου στο δίκτυο, ουσιώδης για την εκτίμηση του SNR.
* Ο SNR υπολογίζεται αφαιρώντας το επίπεδο θορύβου από την ισχύ του ληφθέντος σήματος: SNR = Pr − Noise. Αυτός ο λόγος είναι καθοριστικός για την καθορισμό της ποιότητας της επικοινωνιακής σύνδεσης, όπου υψηλότερο SNR δηλώνει καθαρότερο και πιο αξιόπιστο σήμα.
* Η ποιότητα του σήματος για κάθε χρήστη κατηγοριοποιείται βάσει προκαθορισμένων ορίων SNR: “Excellent” για SNR > 40 dB, έως “Subpar” για SNR < 10 dB. Αυτά τα όρια βοηθούν στην ποσοτικοποίηση της εμπειρίας του χρήστη, μεταφράζοντας τις διάφορες μετρήσεις σε κατανοητά επίπεδα QoS.
* Χρησιμοποιώντας τον τύπο του Shannon για την χωρητικότητα του καναλιού, όπου το SNR από db πρέπει να μετατραπεί σε αριθμό με τον τύπο , το C μετριέται σε Mbps και το B σε MHz, υπολογίζονται τα MHz που χρειάζεται ο κάθε χρήστης με βάση το SNR του και την απαίτηση του σε Mbps. Ύστερα, από MHz υπολογίζονται τα KHz που απαιτεί ο χρήστης και τέλος με τη διαίρεση αυτού με το μέγεθος ενός RB (12 subcarriers \* 30KHz = 360KHz) για να υπολογιστεί η απαίτηση του κάθε χρήστη σε RB.

Στο σενάριο, διατίθενται επίσης δύο σταθμοί βάσης, των οποίων τα χαρακτηριστικά περιλαμβάνουν την τοποθεσία και το διαθέσιμο εύρος ζώνης σε KHz. Ένα βασικό σενάριο χρήσης για αυτούς τους σταθμούς είναι η χρήση της τοποθεσίας τους σε συνδυασμό με τις τοποθεσίες των χρηστών, για τον υπολογισμό της απώλειας διαδρομής (pathloss) κάθε χρήστη προς τους σταθμούς. Η κατανόηση αυτής της απώλειας είναι απαραίτητη για την εκτίμηση της ποιότητας του σήματος που θα λάβει κάθε χρήστης και, συνεπώς, για την προσαρμοσμένη κατανομή των πόρων στις ανάγκες του χρήστη. Η στρατηγική τοποθέτηση των σταθμών σε ένα αστικό περιβάλλον μπορεί να βελτιστοποιήσει την κάλυψη και να ελαχιστοποιήσει τις περιοχές με χαμηλή ποιότητα σήματος, βελτιώνοντας την εμπειρία των χρηστών και την ολική απόδοση του δικτύου.

Άλλη μια σημαντική χρήση των σταθμών αφορά τη χρήση του διαθέσιμου εύρους ζώνης για τον υπολογισμό των υπολειπόμενων RBs της βάσης, διαιρώντας το διαθέσιμο εύρος ζώνης της με το μέγεθος του RB. Αυτό επιτρέπει στο δίκτυο να διαχειρίζεται αποδοτικά τους πόρους του, διασφαλίζοντας ότι οι χρήστες λαμβάνουν την κατάλληλη ποσότητα εύρους ζώνης βάσει των αναγκών τους.



**Εικόνα 4:** Κατανομή του SNR των χρηστών της κεραίας 1



**Εικόνα 5:** Κατανομή του SNR των χρηστών της κεραίας 2

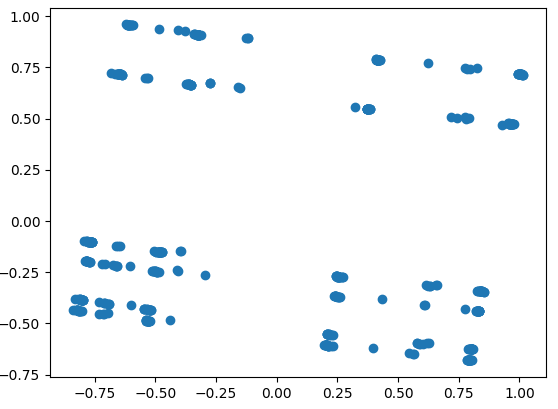
Στις εικόνες 4 και 5 παρουσιάζεται η κατανομή του SNR των χρηστών που είναι συνδεδεμένοι στην κεραία 1 (Εικόνα 4) και στην κεραία 2 (Εικόνα 5). Φαίνεται πως οι περισσότεροι χρήστες έχουν SNR μεταξύ 15 και 20 και στις δύο κεραίες, αρκετοί είχαν μεταξύ 5 και 10 και λίγοι είχαν περισσότερο από 20.

## Προεπεξεργασία των χαρακτηριστικών

Η προεπεξεργασία των χαρακτηριστικών είναι ζωτικής σημασίας για την επιτυχία οποιουδήποτε έργου ML. Μέσω της κανονικοποίησης, της κωδικοποίησης και της μείωσης της διαστατικότητας, μπορούμε να βελτιώσουμε την ακρίβεια και την αποδοτικότητα των αλγορίθμων.

Η πρώτη φάση του κώδικα αφορά την κωδικοποίηση των κατηγορικών δεδομένων. Οι τιμές ποιότητας σήματος (Signal\_Quality\_BS1 και Signal\_Quality\_BS2) είναι κατηγορικές και πρέπει να μετατραπούν σε αριθμητικές για να μπορούν να επεξεργαστούν. Αυτό επιτυγχάνεται με τη χρήση του Label Encoder. Ο Label Encoder μετατρέπει κάθε μοναδική τιμή σε έναν μοναδικό αριθμό, επιτρέποντας τη συνέχιση της ανάλυσης με αριθμητικά δεδομένα.

Μετά την κωδικοποίηση, τα δεδομένα κλιμακώνονται χρησιμοποιώντας τον MinMax Scaler. Ο MinMax Scaler μετασχηματίζει τις τιμές των χαρακτηριστικών στο εύρος [0, 1]. Επειδή ο αλγόριθμος ML μπορεί να θεωρήσει πως χαρακτηριστικά με μεγαλύτερες τιμές από άλλα, έχουν και μεγαλύτερο «βάρος» σαν χαρακτηριστικά, κάτι που επηρεάζει αρνητικά την απόδοση του αλγορίθμου αν δεν ισχύει (όπως στη συγκεκριμένη περίπτωση), η διαδικασία αυτής της μετατροπής είναι σημαντική για να διασφαλιστεί ότι όλα τα χαρακτηριστικά έχουν την ίδια βαρύτητα κατά τη διάρκεια της ανάλυσης και της εκπαίδευσης των μοντέλων.

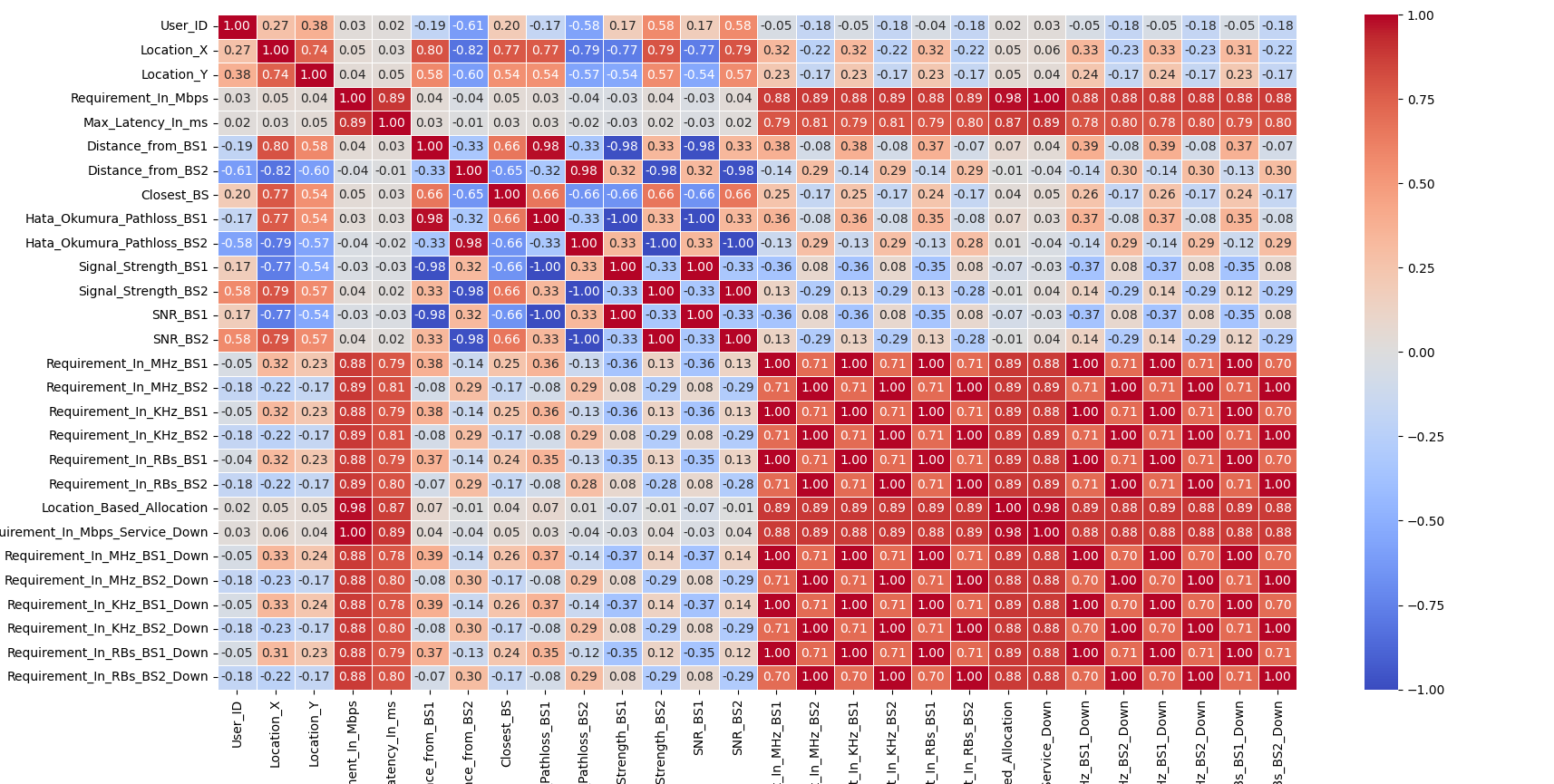


**Εικόνα 6:** Γράφημα (scatterplot) των στοιχείων των χρηστών μετά το PCA 2 διαστάσεων

Η τελευταία φάση του κώδικα αφορά τη μείωση των διαστάσεων των δεδομένων χρησιμοποιώντας την PCA και το αποτέλεσμά της στα δεδομένα φαίνεται στην Εικόνα 6. Η PCA μετασχηματίζει τα δεδομένα σε ένα νέο σύνολο δύο κύριων συνιστωσών (δύο διαστάσεων), που εξηγούν την περισσότερη δυνατή διακύμανση στα δεδομένα. Η μείωση των διαστάσεων διευκολύνει την οπτικοποίηση των δεδομένων (εφόσον θα έχουν λιγότερες διαστάσεις/χαρακτηριστικά) και μπορεί να βελτιώσει την αποδοτικότητα των αλγορίθμων ML.

## Η σχέση των χαρακτηριστικών

Ύστερα από τη μηχανική των χαρακτηριστικών, χρειάστηκε να βρεθεί η συσχέτιση των χαρακτηριστικών και να βρεθούν τα χαρακτηριστικά που έχουν πολύ μεγάλη ή πολύ μικρή συσχέτιση μεταξύ τους, ώστε να αποφασιστούν τα καλύτερα χαρακτηριστικά για τον αλγόριθμο συσταδοποίησης.



**Εικόνα 7:** Γράφημα (heatmap) συσχετίσεων των χαρακτηριστικών των χρηστών

Οι συσχετίσεις φαίνονται στην Εικόνα 7 (heatmap), με τη μέγιστη δυνατή θετική συσχέτιση να αντιπροσωπεύεται με το «1» (κόκκινο) και η μέγιστη δυνατή αρνητική συσχέτιση με το «-1» (μπλε). Χαρακτηριστικά που προήλθαν από άλλα όπως το SNR (που προήλθε από το pathloss), δεν χρειάζεται να συμπεριληφθούν στη διαδικασία συσταδοποίησης (μπορούμε να εξαιρέσουμε οποιοδήποτε από τα δύο).

Επίσης, χαρακτηριστικά όπως το User ID, οι συντεταγμένες και η απόσταση από τις κεραίες δε θα συμπεριληφθούν στην διαδικασία ομαδοποίησης καθώς οι συστάδες που θα δημιουργηθούν θα πρέπει να βασίζονται στην ποιότητα του σήματος, στις απαιτήσεις και στην καθυστέρηση εφόσον αυτά είναι τα πιο σημαντικά για την βελτιστοποίηση της κατανομής πόρων στο πείραμα.

Τελικά, τα χαρακτηριστικά που επιλέχθηκαν ήταν η απαίτηση του χρήστη σε Mbps (οι απαιτήσεις σε MHz, KHz και RBs είχαν μεγάλη συσχέτιση με την απαίτηση σε Mbps καθώς ήταν παράγωγά της), η μέγιστη αποδεκτή καθυστέρηση, η πιο κοντινή κεραία στον χρήστη, το SNR του χρήστη και στις δύο κεραίες και η ποιότητα του σήματος του χρήστη και στις δύο κεραίες (η οποία κωδικοποιήθηκε από τον Label Encoder με το αποτέλεσμα να είναι αριθμοί από το 0 μέχρι και το 4 για τις καταστάσεις “excellent”,”great”,”good”,”ok”,”subpar” αντίστοιχα). Όπως προαναφέρθηκε, στα χαρακτηριστικά που επιλέχθηκαν έγινε αλλαγή κλίμακας (scaling) ώστε κανένα από αυτά να μην υπερεκτιμηθεί στη διαδικασία ομαδοποίησης. Οι νέες τιμές μετά το scaling ήταν στο εύρος [0,1]. Μαζί με το PCA, μειώθηκαν οι διαστάσεις σε 2 κρατώντας αρκετή από την αρχική πληροφορία (των 7 διαστάσεων) κάτι που βοηθάει τους αλγορίθμους αργότερα να είναι πιο αποδοτικοί υπολογιστικά καθώς και να αποφύγουν την υπερεκπαίδευση καθώς ο στόχος των αλγορίθμων ML είναι να «καταλάβουν» τις σχέσεις των δεδομένων ώστε να μπορούν να γενικεύσουν σε νέα δεδομένα και όχι να αποκτήσουν μεγάλη πολυπλοκότητα μαθαίνοντας με ακρίβεια τα δεδομένα εκπαίδευσης κάτι που οδηγεί σε υπερεκπαίδευση και χαμηλή απόδοση σε νέα δεδομένα.

## Οι αλγόριθμοι που χρησιμοποιήθηκαν

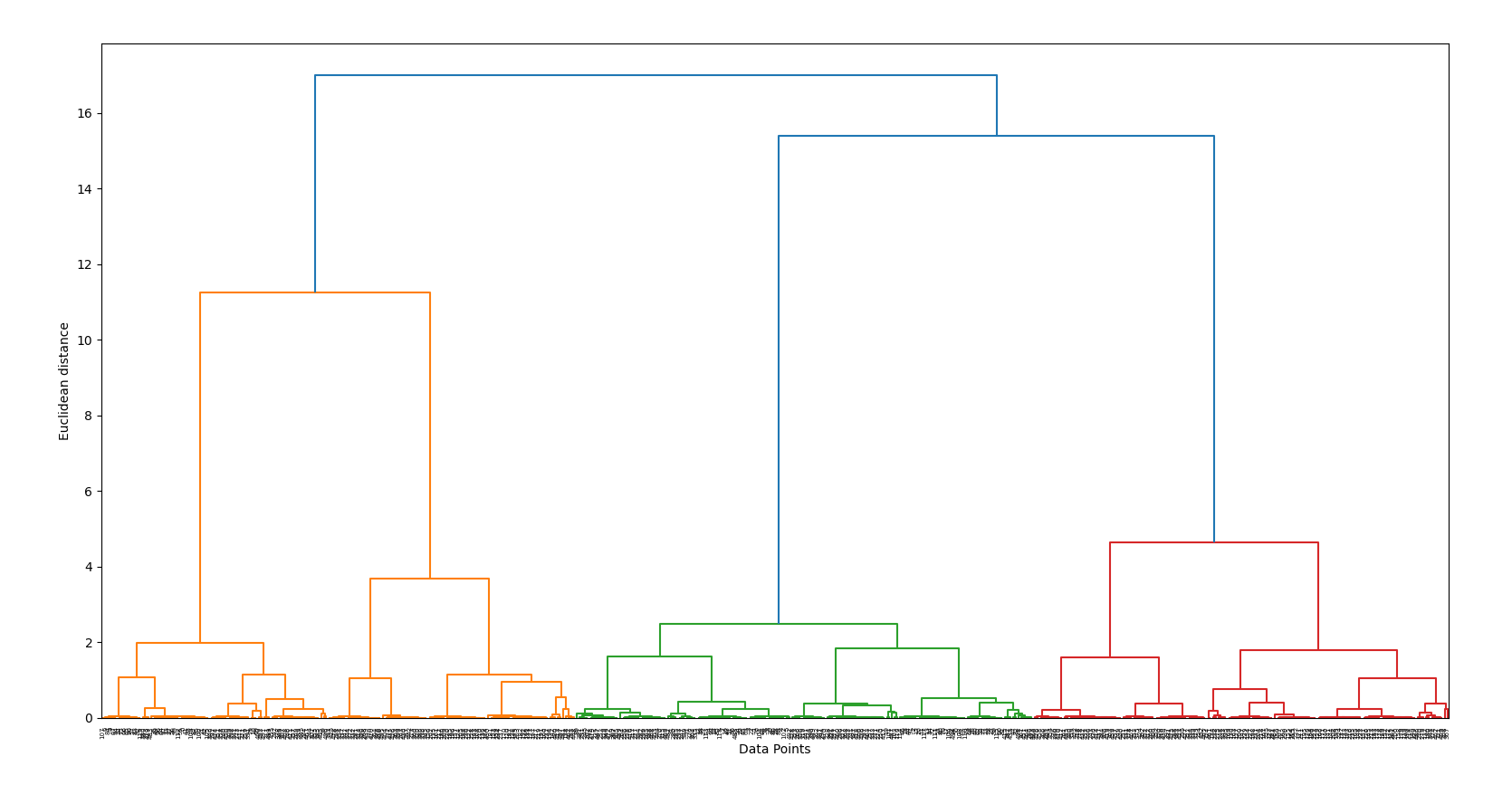
* **Η συσταδοποίηση** **K-Means** το 1957 αναφέρθηκε ως τεχνική από τον Stuart Lloyd, αλλά πρώτη φορά δημοσιεύθηκε από τον Edward W. Forgy το 1965 και αναφέρεται και μερικές φορές ως Lloyd-Forgy αλγόριθμος. Έχει μέση πολυπλοκότητα O(k n T) όπου k ο αριθμός των συστάδων, n ο αριθμός των δειγμάτων και T ο αριθμός επανάληψης στην οποία τελείωσε ο αλγόριθμος. Η χειρότερη δυνατή πολυπλοκότητα είναι O(n^(k+2/p)) όπου p ο αριθμός των χαρακτηριστικών των δειγμάτων. Αποτελεί έναν από τους πλέον διαδεδομένους αλγορίθμους διαμερισμού, προσφέροντας αξιόλογη αποδοτικότητα στην τμηματοποίηση δεδομένων σε k προκαθορισμένες συστάδες, με στόχο τη μεγιστοποίηση της ομοιομορφίας εντός των συστάδων. Ξεκινώντας με τυχαία επιλεγμένα κεντροειδή, σύμφωνα με την παράμετρο init με πιθανές τιμές ‘random’ (που αρχικοποιεί τα κεντροειδή τυχαία) και ‘kmeans++’ (που αρχικοποιεί τα κέντρα με σκοπό την καλύτερη δυνατή συσταδοποίηση και την πιο γρήγορη σύγκλιση), ο αλγόριθμος εφαρμόζει επαναληπτικά τη λειτουργία της εκχώρησης δεδομένων στα κεντροειδή με βάση την Ευκλείδεια απόσταση και τη λειτουργία ανανέωσης των κεντροειδών έτσι ώστε αντικατοπτρίζουν τη μέση θέση των δεδομένων που έχουν εκχωρηθεί στην ομάδα τους. Στα πλεονεκτήματά του η απλότητα και η αποδοτικότητα για πολλά δεδομένα, ενώ ως μειονεκτήματα έχει την απαίτηση καθορισμού του αριθμού των συστάδων και την ευαισθησία στην αρχική επιλογή των κέντρων. Στα δίκτυα 5G, ο K-Means μπορεί να χρησιμοποιηθεί για την κατανομή πόρων, ομαδοποιώντας χρήστες ή συσκευές με παρόμοια πρότυπα χρήσης ή απαιτήσεις σε κοντινά clusters, βελτιστοποιώντας έτσι την κατανομή πόρων και την απόδοση του δικτύου. Ο αλγόριθμος αυτός επιλέγεται και σε άλλες εργασίες λόγω της απλότητάς του και της ικανότητάς του να χειρίζεται μεγάλες ποσότητες δεδομένων γρήγορα[[8]](#footnote-8) [29].
* **Η** **HC** και στην περίπτωση της εργασίας **η Ιεραρχική Συσσωματωτική Συσταδοποίηση** (Hierarchical Agglomerative Clustering - HAC), με πολυπλοκότητα Ο(kn2), όπου k ο αριθμός των clusters και n ο αριθμός των χρηστών (πολυπλοκότητα χειρότερης περίπτωσης Ο(n3), μπορεί να χαρακτηριστεί και άπληστος αλγόριθμος σύμφωνα με τους Horowitz και Sahni, 1979. Το paper των Robert R. Sokal and Peter H. A. Sneath in 1963, με τίτλο "Principles of Numerical Taxonomy" ήταν η αρχή του αλγορίθμου ο οποίος στη συνέχεια χρησιμοποίησε και άλλες μεθόδους εκτός του Ward και συνέχισε να βελτιώνεται. Ο αλγόριθμος αυτός δεν έχει ανάγκη προκαθορισμού του αριθμού των συστάδων και επιτρέπει την ανάλυση σε διαφορετικά επίπεδα λεπτομέρειας, εφόσον παρέχει ένα δενδρόγραμμα που βοηθά στην οπτικοποίηση της ιεραρχίας δεδομένων το οποίο μπορεί να «κοπεί» στον επιθυμητό αριθμό συστάδων. Ως αλγόριθμος, η HC προσφέρει μια δυναμική μεθοδολογία για τη συνένωση δεδομένων σε ολοένα μεγαλύτερες συστάδες, ξεκινώντας από τα επιμέρους στοιχεία μέχρι την τελική ενοποίηση όλων των δεδομένων υπό ένα σύνολο, βάσει ειδικών κριτηρίων σύνδεσης. Τα κριτήρια αυτά καθορίζουν την προσέγγιση των συστάδων με βάση την ελάχιστη, τη μέγιστη ή τη μέση απόσταση μεταξύ των μελών τους. Η HC είναι χρήσιμη για την κατανομή πόρων στα δίκτυα 5G καθώς επιτρέπει την ανάλυση της ιεραρχικής δομής των δεδομένων χρήσης, διευκολύνοντας την κατανομή πόρων σε διαφορετικά επίπεδα και κλίμακες. Σε άλλες εργασίες, επιλέγεται αυτή η μέθοδος λόγω της ικανότητάς της να παρέχει μια πλήρη εικόνα των σχέσεων μεταξύ των δεδομένων [30]. Πιο συγκεκριμένα, οι αλγόριθμοι (όλοι με πολυπλοκότητα Ο(n2logn)) που χρησιμοποιούνται από την HC είναι :
  + ‘Ward’, ο οποίος στοχεύει στην ελάχιστη διασπορά όταν ενώνει δύο clusters.
  + ‘Average’, ο οποίος χρησιμοποιεί την μέση απόσταση των δειγμάτων μίας συστάδας από τις άλλες συστάδες για να αποφασίσει ποια clusters θα ενώσει.
  + ‘Complete’ ή ‘Maximum’ χρησιμοποιεί την μέγιστη απόσταση που μπορεί να έχει μία συστάδα (ένα δείγμα της) από τις άλλες συστάδες για να αποφασίσει ποια clusters θα ενώσει[[9]](#footnote-9).
  + ‘Single’ χρησιμοποιεί την ελάχιστη απόσταση που μπορεί να έχει μία συστάδα (ένα δείγμα της) από τις άλλες συστάδες για να επιλέξει τα clusters που θα ενώσει.
* **Το** **DBSCAN**, προτάθηκε για πρώτη φορά ως αλγόριθμος το 1996 και αποτελεί μια ανθεκτική μέθοδο συσταδοποίησης που εντοπίζει συστάδες με βάση τη χωρική πυκνότητα. Η χειρότερη δυνατή πολυπλοκότητα είναι O(n^2). Σαν αλγόριθμος, το DBSCAN δεν χρειάζεται η ανάγκη προκαθορισμού των κέντρων και είναι εξαιρετικό στο να διαχωρίζει συστάδες με διαφορετικές πυκνότητες. Οι κύριες παράμετροί του των οποίων η επιλογή μπορεί να επηρεάσει έντονα το αποτέλεσμα της συσταδοποίησης, είναι τα eps (ένας κύκλος ακτίνας eps γύρω από κάθε σημείο δεδομένων) και min\_samples (ο απαιτούμενος αριθμός σημείων εντός του ϵ για τον ορισμό μιας πυκνής περιοχής) και επιτρέπουν τη δημιουργία συστάδων που συνδέουν περιοχές υψηλής πυκνότητας. Οι συστάδες δημιουργούνται με βάση την παράμετρο eps. Στην αρχή της διαδικασίας υπολογίζεται για κάθε δεδομένο το πόσα άλλα δεδομένα έχει γύρω του. Ύστερα βρίσκονται τα core points (τα δεδομένα δηλαδή που έχουν τουλάχιστον έναν συγκεκριμένο αριθμό δεδομένων γύρω τους) και μετά η συσταδοποίηση ξεκινά εντάσσοντας ένα τυχαίο core point στο 1ο cluster και συνεχίζει με την ένταξη των core points που βρίσκονται στη «γειτονιά» του προηγούμενου point, στο ίδιο cluster. Για κάθε δεδομένο στη «γειτονιά» το δεδομένου που μόλις εντάχθηκε, βρίσκονται τα core points γύρω του, μπαίνουν στο ίδιο cluster και η διαδικασία αυτή επαναλαμβάνεται μέχρι να μη βρεθούν άλλα core points γύρω. Τα non-core points που βρέθηκαν κατά τη διάρκεια αυτής της αναζήτησης, εντάσσονται τελικά στο 1ο cluster. Εάν έχουν μείνει core points που δεν έχουν βρεθεί, τότε επιλέγεται ένα τυχαίο σημείο από αυτά και όλη η προαναφερθείσα διαδικασία γίνεται ξανά δημιουργώντας το 2ο cluster, ενώ η όλη διαδικασία δημιουργίας συστάδων τελειώνει όταν δεν υπάρχουν άλλα δεδομένα με τα δεδομένα που είναι non-core και δεν βρίσκονται σε καμία ομάδα να ονομάζονται ακραίες περιπτώσεις τιμών (outliers). Το DBSCAN διαχειρίζεται αποτελεσματικά το θόρυβο και τα ακραία σημεία δεδομένων, κάτι που το καθιστά ιδανικό για τα δίκτυα 5G MIMO, όπου η πυκνότητα των χρηστών μεταβάλλεται δραματικά, εξασφαλίζοντας ομοιογενή κατανομή πόρων. Επίσης, ο αλγόριθμος αυτός μπορεί να βοηθήσει στην ανίχνευση ανωμαλιών και στη διαχείριση περιοχών με υψηλή πυκνότητα χρηστών, εξασφαλίζοντας αποτελεσματικότερη κατανομή των πόρων στις πυκνοκατοικημένες περιοχές. Σε άλλες εργασίες το DBSCAN επιλέγεται κυρίως λόγω της αντοχής του σε θορυβώδη δεδομένα και της ικανότητάς του να ανιχνεύει αυθαίρετα διαμορφωμένα clusters[[10]](#footnote-10) [31].
* **Το** **GMM**, αναφέρθηκε για πρώτη φορά από τον Karl Pearson το 1894. Η πολυπλοκότητά του είναι O(NKD3) όπου n ο αριθμός των δεδομένων (data points) μας, k ο αριθμός των clusters (ή components όπως αναφέρονται στις παραμέτρους του αλγορίθμου) και d η διάσταση του προβλήματος ( ο αριθμός των χαρακτηριστικών των δεδομένων μας). Το GMM, μπορεί να διαχειριστεί συστάδες που έχουν διαφορετικά μεγέθη, αλλά είναι πιο υπολογιστικά απαιτητικό από αλγορίθμους που βασίζονται σε κεντροειδή όπως το K-Means. Το μοντέλο αυτό, προσφέρει ένα πιθανοτικό πλαίσιο για την ανάλυση συσταδοποίησης, βασιζόμενο στην υπόθεση ότι τα δεδομένα προέρχονται από πολλαπλές Γκαουσιανές κατανομές. Χρησιμοποιώντας τον αλγόριθμο Προσδοκίας-Μεγιστοποίησης (Expectation Maximization - EM), το GMM αξιολογεί τις παραμέτρους και καθορίζει την πιθανότητα κάθε δεδομένου να ανήκει σε μια συγκεκριμένη συστάδα. Αυτό επιτρέπει μια λεπτομερέστερη κατανόηση και ανάλυση των δεδομένων, κάτι κρίσιμο στα δίκτυα 5G όπου η μεταβλητότητα και η αβεβαιότητα των απαιτήσεων υπηρεσιών και της ποιότητας του σήματος είναι υψηλή. Η εφαρμογή του GMM μπορεί να αναδείξει τη δυνατότητα των δικτύων 5G MIMO να διαχειρίζονται αποδοτικά τις υπηρεσίες ανάλογα με τις συγκεκριμένες ανάγκες των χρηστών, παρά τις υπάρχουσες αβεβαιότητες. Ωστόσο, είναι κρίσιμο να αναγνωριστεί ότι εφόσον το GMM λειτουργεί υπό τη θεμελιώδη υπόθεση ότι τα δεδομένα που αναλύει παράγονται από Γκαουσιανές κατανομές (μια προϋπόθεση που μπορεί να μην συμβαδίζει πάντα με την πολυεπίπεδη και απρόβλεπτη φύση των πραγματικών δεδομένων, ειδικά στο πλαίσιο των δικτύων 5G όπου οι συμπεριφορές των χρηστών και η διάδοση του σήματος μπορεί να εκδηλώσουν μη-Γκαουσιανά χαρακτηριστικά), στην διαδικασία της συσταδοποίησης προσπαθεί να μεγιστοποιήσει την πιθανότητα του κάθε cluster να αντιπροσωπεύει μία Γκαουσιανή Κατανομή. Έτσι, μπορεί να περιοριστεί η αποδοτικότητα και η εφαρμογή του σε ορισμένα πραγματικά σενάρια. Παρ’ όλ’ αυτά, η ανθεκτικότητα του μοντέλου στη διαχείριση επικαλυπτόμενων συστάδων και η ικανότητά του να ενσωματώνει την αβεβαιότητα στις αναθέσεις συστάδων το καθιστούν ένα ισχυρό εργαλείο για την κατανομή πόρων σε περιβάλλοντα 5G MIMO. Το GMM συνήθως επιλέγεται λόγω της ικανότητάς του να διαχειρίζεται περίπλοκες κατανομές και να προσφέρει λεπτομερή ανάλυση των δεδομένων[[11]](#footnote-11) [32].
* **Η** **φασματική συσταδοποίηση (spectral clustering)**, «γεννημένη» από το spectral graph partitioning το 1972-1973 και με πολυπλοκότητα Ο(n^3) (όπου n ο αριθμός των δειγμάτων), μας βοηθά να ξεπεράσουμε δύο μεγάλα προβλήματα όσον αφορά τη συσταδοποίηση. Το πρώτο είναι ο καθορισμός των κέντρων ( ή κεντροειδών) ώστε να αρχίσει ο αλγόριθμος και το δεύτερο είναι το ότι αρκετοί αλγόριθμοι συσταδοποίησης έχουν ως βασική υπόθεση ότι οι ομάδες που θα καθοριστούν θα είναι του ίδιου σχήματος. Σαν αλγόριθμος, είναι υπολογιστικά δαπανηρός αλλά προσφέρει εξαιρετική απόδοση όταν χρειάζεται να διαχειριστούμε δεδομένα υψηλού αριθμού διαστάσεων. Η φασματική συσταδοποίηση λοιπόν, αποτελεί μια εξελιγμένη τεχνική, ιδιαίτερα χρήσιμη για την ομαδοποίηση σημείων δεδομένων με βάση την ομοιότητά τους, όπως εφαρμόζεται στα συστήματα 5G MIMO. Αντίθετα με παραδοσιακές μεθόδους, η φασματική συσταδοποίηση αναλύει τα δεδομένα μέσω των ιδιοτιμών του πίνακα ομοιότητας, δημιουργώντας ένα γράφημα ομοιότητας όπου οι κόμβοι αντιπροσωπεύουν τα δεδομένα και οι ακμές την ομοιότητα μεταξύ τους. Μέσω του πίνακα Laplace του γραφήματος, χρησιμοποιούνται ιδιοτιμές και ιδιοδιανύσματα για την μετατροπή των δεδομένων σε χαμηλότερη διάσταση, επιτρέποντας την εφαρμογή τεχνικών όπως η K-Means για αποτελεσματική συσταδοποίηση. Στον τομέα των 5G MIMO, η φασματική συσταδοποίηση μπορεί να αναδείξει σημαντικές ομαδοποιήσεις, βελτιώνοντας την απόδοση και την αξιοπιστία του δικτύου. Η τεχνική αυτή προσφέρει ένα ισχυρό εργαλείο για τον σχεδιασμό συστημάτων και την επεξεργασία σήματος, καθιστώντας την απαραίτητη για την επόμενη γενιά της ασύρματης επικοινωνίας. Ο αλγόριθμος αυτός έχει δείξει σε προηγούμενες εργασίες πως το προτέρημά του βρίσκεται στη διαχείριση μη γραμμικών σχέσεων στα δεδομένα[[12]](#footnote-12) [33].
* **Το** **PCA** εφευρέθηκε σαν τεχνική το 1901 από τον Karl Pearson και δημιουργήθηκε ωςαλγόριθμος από τον Harold Hoteling τη δεκαετία του ’30 και αποτελεί μία κορυφαία στατιστική τεχνική, ιδιαίτερα χρήσιμη για τη μείωση της διαστατικότητας στα πλαίσια της ανάλυσης δεδομένων και της ML. Μέσω της μετατροπής ενός εκτεταμένου συνόλου μεταβλητών σε ένα περιορισμένο σύνολο (μείωση της διαστατικότητας των δεδομένων), η PCA διατηρεί την ουσιαστική πληροφορία των δεδομένων, μειώνοντας ταυτόχρονα την πολυπλοκότητά τους. Αυτό επιτυγχάνεται μέσω του εντοπισμού κατευθύνσεων υψηλής μεταβλητότητας, τα λεγόμενα κύρια συστατικά, κάτι που είναι ιδιαίτερα σημαντικό σε συστήματα 5G MIMO που χαρακτηρίζονται από μεγάλο αριθμό παραμέτρων και κεραιών και σύνθετες αλληλεπιδράσεις σημάτων. Στο περιβάλλον των 5G MIMO, η PCA βοηθά κυρίως στην ανάλυση και οπτικοποίηση των συσχετίσεων και των μοτίβων σήματος, βασικά στοιχεία για την βελτιστοποίηση της χωρητικότητας του καναλιού και της ακεραιότητας του σήματος. Η δυνατότητα μείωσης της διαστατικότητας των δεδομένων επιτρέπει όχι μόνο την ενίσχυση της υπολογιστικής αποδοτικότητας αλλά και τη βελτίωση των μοντέλων ML, προσφέροντας πιο αποτελεσματικές τεχνικές μείωσης θορύβου και βελτιωμένη ανάλυση κατάστασης καναλιού. Επιπλέον, η PCA στα δίκτυα 5G διευκολύνει τη διαχείριση συνθετότερων σεναρίων πολλαπλών διαδρομών και παρεμβολών σημάτων, καθιστώντας την κρίσιμη για την εξαγωγή χαρακτηριστικών και την αναγνώριση διαφορετικών τύπων παρεμβολών. Αυτή η ικανότητα της PCA να απλοποιεί και να βελτιώνει τις στρατηγικές επεξεργασίας σημάτων προάγει την αξιοπιστία και την αποδοτικότητα των συστημάτων 5G, παρέχοντας μια ισχυρή βάση για την προώθηση της επόμενης γενιάς τηλεπικοινωνιακών τεχνολογιών. Το PCA έχει πολυπλοκότητα Ο(p^3 + p^2n) (όπου p τα χαρακτηριστικά των δειγμάτων και n ο αριθμός τους), εφόσον απαρτίζεται από δύο διαδικασίες. Τον υπολογισμό του πίνακα συνδιασπορών (με πολυπλοκότητα Ο(p^2n)) και την ανάλυση σε ιδιοτιμές (με πολυπλοκότητα Ο(p^3)). Στο συγκεκριμένο πείραμα, ο αριθμός των διαστάσεων μετά το PCA ήταν δύο (όπως παρουσιάζεται και στην Εικόνα 6)
* **Η απλή μέθοδος κατανομής RBs,** βασισμένη στην ανάθεση πόρων σε χρήστες βάσει του πλησιέστερου σταθμού και με προτεραιότητα στον καλύτερο δείκτη SNR. Ο σκοπός του αλγορίθμου είναι η διανομή των πόρων του δικτύου, διασφαλίζοντας ότι οι χρήστες με την υψηλότερη ποιότητα σήματος λαμβάνουν προτεραιότητα στην πρόσβαση σε αυτούς τους πόρους. Ο αλγόριθμος αυτός αποτελεί τη βάση για περαιτέρω βελτιστοποίηση, πάνω στην οποία μπορεί να αναπτυχθεί μια πιο σύνθετη στρατηγική διαχείρισης πόρων, ανταποκρινόμενη δυναμικά στις αλλαγές του φορτίου και των απαιτήσεων του δικτύου.
* **Η σύνθετη μεθοδολογία κατανομής των RBs στο πείραμα της εργασίας,** χρησιμοποιεί έναν συνδυασμό ομαδοποίησης χρηστών βάσει των διαφορών SNR και των απαιτήσεων καθυστέρησης, ακολουθούμενο από έναν πολύπλοκο αλγόριθμο κατανομής πόρων που λαμβάνει υπόψη την εγγύτητα του χρήστη στους βασικούς σταθμούς και την τρέχουσα διαθεσιμότητα πόρων. Αυτή η προσέγγιση όχι μόνο ενισχύει την αποδοτικότητα χρήσης των πόρων σε όλο το δίκτυο αλλά εξασφαλίζει επίσης ότι η εμπειρία του χρήστη διατηρείται σε βέλτιστο επίπεδο, παρά τις διακυμάνσεις των συνθηκών του δικτύου και τις απαιτήσεις των άλλων χρηστών. Μέσω αυτής της στρατηγικής, η εργασία αποδεικνύει ένα προσαρμόσιμο πλαίσιο για τη διαχείριση πόρων σε ασύρματα δίκτυα 5G. Τα βήματα που ακολουθεί παρουσιάζονται παρακάτω.
  + Αρχικά, οι χρήστες ομαδοποιούνται βάσει των ταυτοτήτων των συστάδων τους, οι οποίες στη συγκεκριμένη περίπτωση προέρχονται από το αποτέλεσμα του K-Means. Στις συστάδες με τους χρήστες με τις μεγαλύτερες αποδεκτές καθυστερήσεις, δίνονται τα απαιτούμενα RBs. Αυτό γίνεται ώστε ύστερα αυτοί οι χρήστες να μπορούν να δώσουν σε άλλους χρήστες (με μικρότερες αποδεκτές καθυστερήσεις) τα ζητούμενα RBs εάν καμία κεραία δεν τα διαθέτει. Άρα, προτιμώνται πρώτα χρήστες που εάν χρειαστεί να δώσουν RBs, να μπορούν και να αντέξουν τις καθυστερήσεις μέχρι να αποδεσμευτούν RBs από άλλους χρήστες όταν αυτοί βγουν από το δίκτυο. Ύστερα, για κάθε συστάδα, υπολογίζουμε την απόλυτη διαφορά SNR μεταξύ των δύο σταθμών (δηλαδή τη διαφορά SNR στο BS1 και SNR στο BS2). Αυτός ο δείκτης είναι ζωτικής σημασίας καθώς βοηθά στην αναγνώριση των χρηστών που θα υποστούν τη μεγαλύτερη υποβάθμιση στην ποιότητα της υπηρεσίας, αν δεν διατεθούν επαρκείς πόροι, κυρίως λόγω της γεωγραφικής τους θέσης σε σχέση με τους σταθμούς βάσης.
  + Μετά την ομαδοποίηση, οι συστάδες ταξινομούνται με βάση την υπολογισμένη απόλυτη διαφορά SNR. Όπως προαναφέρθηκε, η λογική πίσω από αυτή την ταξινόμηση είναι ότι οι χρήστες με μεγαλύτερη διαφορά SNR μεταξύ δύο βασικών σταθμών πιθανότατα θα βιώσουν μεγαλύτερες ανισότητες στην ποιότητα της υπηρεσίας, επομένως η απόδοση προτεραιοτήτων στους χρήστες εξασφαλίζει ότι οι πόροι διατίθενται όπου χρειάζονται περισσότερο για να διατηρηθεί η απόδοση του δικτύου και η ικανοποίηση των χρηστών.
  + Στη συνέχεια, πριν την επαναληπτική διαδικασία κατανομής πόρων οι χρήστες εντός των συστάδων ταξινομούνται με αύξουσα σειρά με βάση την αποδεκτή καθυστέρηση. Αυτή η ταξινόμηση βοηθά στο να δοθεί προτεραιότητα στους χρήστες που δεν μπορούν να ανεχτούν καθυστερήσεις, επιτρέποντας στο δίκτυο να εξυπηρετήσει αυτούς τους χρήστες πρώτα. Αυτό το βήμα είναι κρίσιμο σε σενάρια όπως αυτό, όπου οι πόροι του δικτύου είναι περιορισμένοι και πρέπει να διαχειριστούμε δυναμικά τις ποιότητες υπηρεσίας και την κατανομής πόρων.
  + Ο πυρήνας της μεθόδου περιλαμβάνει έναν λεπτομερή αλγόριθμο κατανομής πόρων. Κάθε χρήστης εντός των ταξινομημένων συστάδων αξιολογείται για τον καθορισμό του βέλτιστου σταθμού βάσης, βάσει της εγγύτητάς του και του τρέχοντος φορτίου στον σταθμό. Οι χρήστες αρχικά δοκιμάζονται από τον πλησιέστερο σταθμό (όπως καθορίζεται από τα απαιτούμενα RBs ή από την απόσταση του χρήστη από τους σταθμούς). Εάν η κατανομή στον πλησιέστερο σταθμό δεν είναι δυνατή, λόγω περιορισμένων πόρων ο αλγόριθμος προσπαθεί να κατανείμει τα RBs από τον άλλο σταθμό βάσης.
  + Σε περιπτώσεις που οι σταθμοί βάσης είναι κοντά στην πλήρη χωρητικότητα, γίνεται αναγκαία η έξυπνη αναδιανομή πόρων στους χρήστες που ζητούν σύνδεση. Ο προτεινόμενος αλγόριθμος περιλαμβάνει έναν κύκλο/λούπα κατά τον οποίο RBs φεύγουν από τους χρήστες με τη μέγιστη αποδεκτή καθυστέρηση μέχρις ότου εκείνοι να φτάσουν μία πιο χαμηλή QoS και δίνονται στους νέους χρήστες που προσπαθούν να μπουν μέχρις ότου οι νέοι χρήστες να επιτύχουν την υπηρεσία που ζητούν αλλά σε χαμηλότερη ποιότητα. Αν σε κάποιο σημείο αυτής της επανάληψης η αναδιανομή προκαλέσει υπερβολική κατανομή RBs (δηλαδή το άθροισμα των δοθέντων RBs υπερβαίνει τη ζήτηση του χρήστη), το πλεόνασμα επιστρέφεται στον τελευταίο χρήστη από τον οποίο αφαιρέθηκαν RBs. Αυτή η δυναμική αναδιανομή των RBs, βοηθά στη διατήρηση ισορροπίας στη διανομή των πόρων σε όλο το δίκτυο, εξασφαλίζοντας δικαιοσύνη και βελτιστοποιώντας την συνολική απόδοση του δικτύου.

## PCA, scaling και βελτιστοποίηση των αλγορίθμων

Όσον αφορά τη βελτιστοποίηση των αλγορίθμων που χρησιμοποιήθηκαν, αρκετές τεχνικές εφαρμόστηκαν με σκοπό τον εντοπισμό του βέλτιστου αριθμού κέντρων του K-Means του GMM και του spectral clustering, τον εντοπισμό των βέλτιστων παραμέτρων του DBSCAN και τη δημιουργία του δενδρογράμματος του HΑC για την αρχική εκτίμηση των συστάδων. Οι τεχνικές που χρησιμοποιήθηκαν και τα αποτελέσματά τους είναι τα παρακάτω :

### Αρχική εκτίμηση συστάδων με χρήση δενδρογράμματος και HAC

Για την αρχική εκτίμηση των συστάδων χρησιμοποιήθηκε το δενδρόγραμμα με τη μέθοδο ‘Ward’, ώστε να παρουσιαστούν οι σχέσεις μεταξύ των όλων των clusters. Έτσι θα μπορέσει να υπάρξει ένα αρχικό εύρος (μία εκτίμηση) όσον αφορά τον αριθμό των συστάδων που θα πρέπει να χρησιμοποιηθούν. ‘Κόβοντας’ το γράφημα αυτό οριζόντια σε διαφορετικά επίπεδα συναντώνται κάθετες γραμμές. Όσο πιο χαμηλά χωριστεί το γράφημα αυτό, τόσο πιο λεπτομερές είναι το clustering (τόσες περισσότερες κάθετες γραμμές συναντώνται). Επίσης, όσο πιο μεγάλες είναι οι δύο κάθετες γραμμές που ενώνονται (και δημιουργούν μια οριζόντια γραμμή), τόσο μεγαλύτερες διαφορές είχαν οι δύο συστάδες που ενώθηκαν.



**Εικόνα 8:** Δενδρόγραμμα HAC

Χωρίζοντας λοιπόν σε διαφορετικά επίπεδα, προκύπτει το ιδανικό εύρος συστάδων από 3 μέχρι και 8. Όπως φαίνεται και στην Εικόνα 8 τα points αρχικά ομαδοποιούνται με όμοιά τους για να σχηματίσουν μεγαλύτερα clusters και όσο «προχωράμε» προς τα επάνω οι ομοιότητες γίνονται όλο και μικρότερες.

### Βελτιστοποίηση του αλγορίθμου K-Means

Έχοντας λοιπόν ιδανικό εύρος από 3 μέχρι και 8 συστάδες, χρειάζεται να υπολογιστεί ο βέλτιστος αριθμός κέντρων για τον αλγόριθμο K-Means. Αυτό γίνεται μέσω των τεχνικών του elbow point και του silhouette score.

### *Άθροισμα τετραγώνων εντός συστάδας (Within Cluster Sum of Squares - WCSS)*

Η τεχνική του σημείου του αγκώνα για το K-Means, βασίζεται πρώτα στον υπολογισμό της τετραγωνικής απόστασης των σημείων εντός ενός cluster από το κέντρο του cluster για κάθε cluster και ονομάζεται WCSS. O μαθηματικός τύπος για το WCSS φαίνεται παρακάτω :

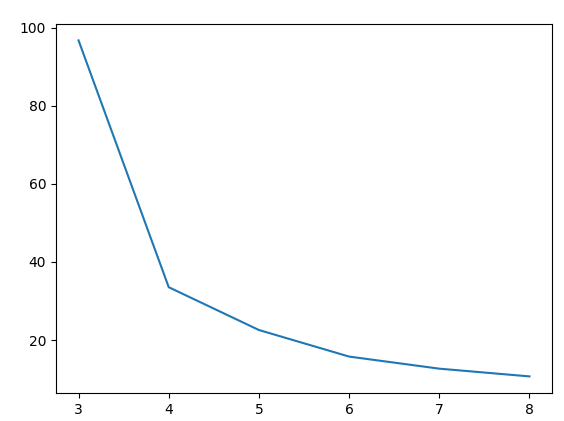
Όπου Ν είναι το πλήθος των συστάδων, και το j κάθε φορά υποδηλώνει το cluster στο οποίο γίνεται ο υπολογισμός (επιλέγοντας το κέντρο του Cj), ενώ το Χ κάθε φορά αντιπροσωπεύει τα σημεία εντός της συστάδας (όπου CLj η συστάδα j). Πρέπει ύστερα να μπορέσει να αποφασιστεί ποιος αριθμός clusters δίνει το χαμηλότερο WCSS, άρα θα πρέπει να εκπαιδευτεί ο αλγόριθμος επαναληπτικά για 3 μέχρι και 8 συστάδες και να βρεθεί το WCSS για κάθε επανάληψη. Όσο περισσότερα clusters χρησιμοποιούνται τόσο μικρότερo WCSS προκύπτει, άρα πρέπει να χρησιμοποιηθεί η τεχνική του σημείου “αγκώνα” (elbow point) καθώς υπάρχει ένα trade-off στον αριθμό των clusters και στο WCSS, εφόσον δεν επιθυμούνται πολλά clusters λόγω του ότι θα αυξηθεί η πολυπλοκότητα αλλά ταυτόχρονα επιθυμείται μικρό WCSS.

### *Elbow point*

Υπολογίζοντας το WCSS για τον αλγόριθμο K-Means για 3 μέχρι 8 κέντρα, προκύπτει το παρακάτω γράφημα όπου στον άξονα των x βρίσκεται ο αριθμός των κέντρων και στον άξονα των y το “σκορ” (αναλυτικά τα αποτελέσματα παρουσιάζονται και στον Πίνακα). Το elbow point είναι το σημείο όπου από εκεί και μετά η μείωση του WCSS γίνεται λιγότερο δραστική. Το σημείο αυτό δείχνει δηλαδή τον αριθμό συστάδων ύστερα από τον οποίο δεν βελτιώνεται σημαντικά την ομοιογένεια των συστάδων και έχει επιτευχθεί μία καλή ισορροπία μεταξύ πολυπλοκότητας του αλγορίθμου και πυκνότητας των συστάδων[[13]](#footnote-13) [34].

**Πίνακας 1:** Αποτελέσματα WCSS αλγορίθμου K-Means

|  |  |
| --- | --- |
| **Number of Clusters** | **K-Means Result** |
| 3 | 96.786042 |
| 4 | 33.533191 |
| 5 | 22.565698 |
| 6 | 15.754578 |
| 7 | 12.658495 |
| 8 | 10.675843 |

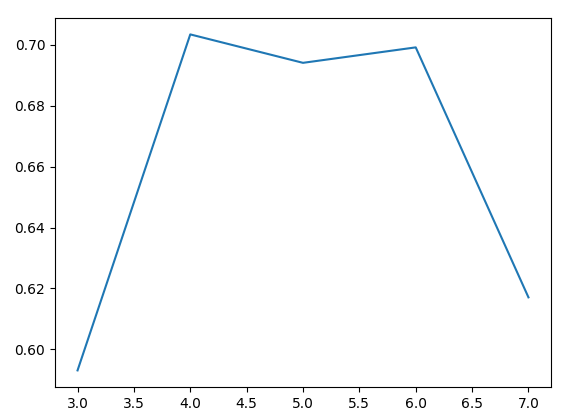


**Εικόνα 9:** WCSS score για αριθμό συστάδων 3 με 8

Στον Πίνακα 1 και στην Εικόνα 9 παρατηρείται πως το σημείο του αγκώνα βρίσκεται στις 4 συστάδες. Η διαδικασία βελτιστοποίησης συνεχίζει με τον υπολογισμό του silhouette score για αριθμό συστάδων «γύρω» από το elbow point.

### *Silhouette score*

Για την περαιτέρω βελτιστοποίηση του K-Means αλγορίθμου χρησιμοποιείται το silhouette score, το οποίο υπολογίζει και δείχνει το πόσο καλά “ταιριάζει” το κάθε data point στο cluster στο οποίο έχει τοποθετηθεί. Δηλαδή το πόσο μοιάζει με τα υπόλοιπα δεδομένα του ίδιου cluster, μια και αυτός είναι ο σκοπός της συσταδοποίησης. Όσο υψηλότερο silhouette score, τόσο περισσότερο μοιάζουν τα δεδομένα του ίδιου cluster μεταξύ τους. Εφόσον η τεχνική elbow point έδειξε πως ένα αρκετά χαμηλό WCSS επιτυγχάνεται με 4 ομάδες, η αναζήτηση για το βέλτιστο silhouette score γίνεται από το elbow point -1 μέχρι το elbow point +3.



**Εικόνα 10:** Silhouette score K-Means για αριθμό συστάδων 3 με 7

**Πίνακας 2:** Αποτελέσματα silhouette score αλγορίθμου K-Means

|  |  |
| --- | --- |
| **Number of Clusters** | **K-Means Result** |
| 3 | 0.593069 |
| 4 | 0.703475 |
| 5 | 0.694131 |
| 6 | 0.699213 |
| 7 | 0.703475 |

Στην Εικόνα 10 και στον Πίνακα 2 φαίνεται πως τα 4 cluster δίνουν το καλύτερο σκορ άρα και επιλέγονται ως τα n\_clusters του K-Means.

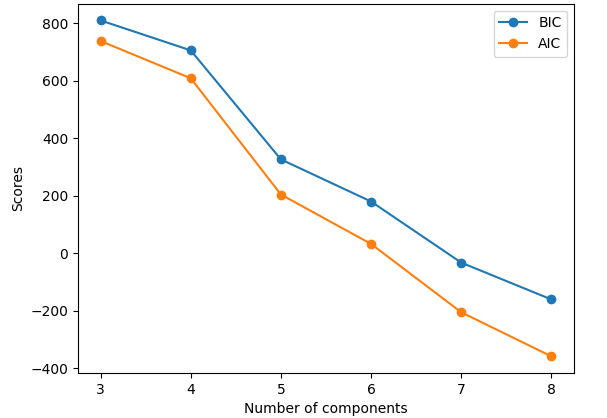
### Βελτιστοποίηση του αλγορίθμου GMM

Οι δείκτες BIC και AIC μετρούν την ποιότητα του μοντέλου GMM. Οι χαμηλότερες τιμές Κριτηρίου Πληροφορίας του Bayes (Bayesian Information Criterion - BIC) και Κριτηρίου Πληροφορίας του Akaike (Akaike Information Criterion - AIC) υποδεικνύουν καλύτερη επίδοση του μοντέλου και η ιδέα στην οποία βασίζονται και τα δύο είναι το να προσθέτουν “ποινή” σε ένα μοντέλο με μεγάλη πολυπλοκότητα (το BIC δίνει μεγαλύτερη ποινή από το AIC) [35]. Άρα περισσότερες παράμετροι στο μοντέλο ζημιώνουν (αυξάνουν) αυτά τα δύο σκορ. Οι μαθηματικοί τύποι των δύο αυτών σκορ είναι :

Όπου k ο αριθμός των παραμέτρων, n ο αριθμός των παρατηρήσεων (δεδομένων) και L η μέγιστη πιθανοφάνεια του μοντέλου, δηλαδή το πόσο πιθανό είναι να έχουμε τα αποτελέσματα που παρουσιάστηκαν, δεδομένου πως έχουμε το συγκεκριμένο μοντέλο και τις συγκεκριμένες παραμέτρους.

Στις γραφικές παραστάσεις των BIC και AIC στη συνέχεια, αναζητούμε ξανά ένα σημείο "αγκώνα”. Αυτό το σημείο μας υποδεικνύει τον βέλτιστο αριθμό συστάδων για το μοντέλο έτσι ώστε να αποφευχθεί η υπερεκπαίδευση αλλά ταυτόχρονα να υπάρχει επαρκής πολυπλοκότητα για την επιτυχή προσαρμογή στα δεδομένα που δίνονται.

### *BIC και AIC scores*



**Εικόνα 11:** BIC Και AIC scores του GMM για αριθμό συστάδων 3 με 8

**Πίνακας 3:** Αποτελέσματα BIC και AIC score αλγορίθμου GMM

|  |  |  |
| --- | --- | --- |
| **Number of Clusters** | **BIC Score** | **AIC Score** |
| 3 | 810.753971 | 739.105633 |
| 4 | 706.463817 | 609.527830 |
| 5 | 327.191442 | 204.967807 |
| 6 | 180.522132 | 33.010848 |
| 7 | -32.087277 | -204.886209 |
| 8 | -160.177930 | -358.264510 |

Από την Εικόνα 11 και τον Πίνακα 3 φαίνεται πως ο βέλτιστος αριθμός clusters για το GMM είναι 5 (ο οποίος είναι, όπως και πριν, το elbow point). Εφόσον με 5 clusters, το GMM δεν απέδωσε τόσο καλά όσο ο αλγόριθμος ΚMeans με 4, «αποθηκεύονται» τα αποτελέσματα του K-Means ως βέλτιστα.

### Βελτιστοποίηση παραμέτρων του DBSCAN

O αλγόριθμος DBSCAN, βασίζεται στην πυκνότητα των δεδομένων και ομαδοποιεί δεδομένα εντός μίας μέγιστης απόστασης. Αυτή η απόσταση είναι και η πιο σημαντική παράμετρος του DBSCAN και ονομάζεται eps και αν δεν αλλαχθεί χειροκίνητα, ισούται με 0.5. Για την βελτιστοποίηση του αλγορίθμου, πρέπει να γίνει επαναληπτική εκπαίδευσή του για διαφορετικά eps (επιλέχθηκε 0.15 μέχρι και 1 με βήμα 0.05).

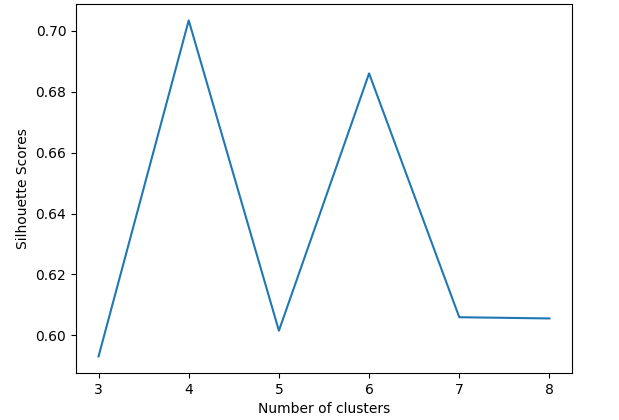
**Πίνακας 4:** Αποτελέσματα αλγορίθμου DBSCAN για διαφορετικές τιμές του eps

|  |  |
| --- | --- |
| **Eps Value** | **DBSCAN Silhouette Score** |
| 0.15 | 0.6590240898617581 |
| 0.2 | 0.6657135586955105 |
| 0.25 | 0.7034750787882275 |
| 0.30 | 0.7034750787882275 |
| 0.35 | 0.7034750787882275 |
| 0.40 | 0.7034750787882275 |
| 0.45 | 0.7034750787882275 |
| 0.50 | 0.5930689520144629 |
| 0.55 | 0.4575803694263533 |
| 0.60 | 0.4575803694263533 |
| 0.65 | 0.4575803694263533 |
| 0.70 | 0.4575803694263533 |
| 0.75 | 0.4575803694263533 |
| 0.80 | 0.4575803694263533 |
| 0.85 | Only 1 cluster detected |
| 0.90 | Only 1 cluster detected |
| 0.95 | Only 1 cluster detected |
| 1.00 | Only 1 cluster detected |

Όπως φαίνεται στον Πίνακα 4, για eps τιμές μεταξύ 0.25 και 0.45, οι δείκτες silhouette είναι υψηλοί (0.703475), υποδεικνύοντας ότι η ομαδοποίηση είναι αποτελεσματική. Για eps τιμές μεγαλύτερες από 0.5, οι δείκτες silhouette μειώνονται σημαντικά, και για τιμές πάνω από 0.85, ανιχνεύεται μόνο μία συστάδα, δείχνοντας ότι το eps είναι πολύ μεγάλο για να διαχωρίσει αποτελεσματικά τα δεδομένα σε διαφορετικές συστάδες. Αυτά τα αποτελέσματα δείχνουν ότι οι μικρότερες τιμές eps είναι πιο αποτελεσματικές για την κατανομή πόρων στο συγκεκριμένο πρόβλημα. Επιλέχθηκε eps ίσο με 0.25.

### Silhouette score για το spectral clustering

Το φασματικό clustering, ενώ διαθέτει αρκετές παραμέτρους, οι προκαθορισμένες τιμές για αυτές τις παραμέτρους δίνουν καλό silhouette score και καλή πολυπλοκότητα στα περισσότερα datasets.



**Εικόνα 12:** Silhouette score spectral clustering για αριθμό συστάδων 3 με 8

**Πίνακας 5:** Αποτελέσματα silhouette score αλγορίθμου spectral clustering

|  |  |
| --- | --- |
| **Number of Clusters** | **K-Means Result** |
| 3 | 0.5930689520144629 |
| 4 | 0.7034750787882275 |
| 5 | 0.6015117222452537 |
| 6 | 0.6860534806047354 |
| 7 | 0.6059406400675139 |

Με εκπαίδευση από 3 μέχρι και 8 clusters έχουμε τα silhouette score που παρουσιάζονται στον Πίνακα 5 ή στην Εικόνα 12 (σαν σχήμα) και διαπιστώνεται πως τα 4 clusters είναι ιδανικά και για αυτό τον αλγόριθμο. Αν και όπως και στον K-Means, τα 6 clusters έδωσαν επίσης καλό score αλλά όχι καλύτερο απ’ ότι τα 4.

### Δείκτης CH

Ο Δείκτης CH αποτελεί μια στατιστική μέτρηση που αξιολογεί την ικανότητα διαχωρισμού των συστάδων (clusters) σε ένα πρόβλημα συσταδοποίησης δεδομένων. Συγκεκριμένα, μετρά το λόγο της διακύμανσης μεταξύ των συστάδων προς τη διακύμανση εντός κάθε συστάδας. Υψηλές τιμές CH υποδηλώνουν συστάδες με μεγάλη διακύμανση μεταξύ τους και χαμηλή διακύμανση εντός τους, κάτι που είναι επιθυμητό στη συσταδοποίηση καθώς δηλώνει ξεκάθαρη διαχωριστικότητα και συνοχή. Η μετρική αυτή λοιπόν δείχνει το πόσο καλά είναι διαχωρισμένες οι συστάδες αλλά και πόσο πυκνά είναι τα δείγματα εντός των συστάδων. Ακολουθεί τον παρακάτω τύπο όπου nk ο αριθμός των points εντός της συστάδας k, ck το κέντρο της συστάδας, c το ολικό κέντρο, Ν ο αριθμός των data points συνολικά και dk το συγκεκριμένο data point που ελέγχεται κάθε φορά και Κ ο αριθμός των clusters[[14]](#footnote-14) [36]:

A close-up of a number

Description automatically generated

**Εικόνα 13:** Μαθηματικός τύπος για τον δείκτη CH

**Πίνακας 6:** Αποτελέσματα δείκτη CH για κάθε αλγόριθμο (μετά τη βελτιστοποίηση)

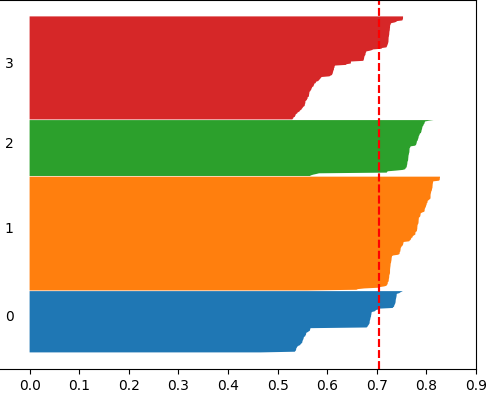
|  |  |
| --- | --- |
| **Algorithm** | **CH Result** |
| K-Means | 1606.4847814510365 |
| DBSCAN | 1606.4847814510365 |
| Hierarchical (4 clusters) | 1606.4847814510365 |
| GMM | 1831.4887517765555 |
| Spectral | 1606.4847814510365 |

### Αποτέλεσμα CH score

Όπως φαίνεται στην Εικόνα 13 και στον Πίνακα 6, με τα βέλτιστα clusters για εκείνον (4), ο αλγόριθμος K-Means επιτυγχάνει σκορ CH 1606,48. Με τον ίδιο αριθμό clusters, ο DBSCAN και η HC και το φασματικό clustering παρουσιάζουν αντίστοιχο σκορ. Με τα βέλτιστα clusters για εκείνον (5) ο αλγόριθμος GMM καταγράφει το υψηλότερο σκορ της μετρικής CH, φτάνοντας το 1831,49. Αυτό υποδηλώνει ότι, παρά την χαμηλότερη απόδοση στη μετρική Silhouete, ο GMM έχει τον καλύτερο λόγο διαχωρισμού των clusters προς συνοχή των data points ανά cluster. Παρά τα ευρήματα αυτά, ο K-Means επιλέχθηκε ως ο καταλληλότερος αλγόριθμος για το συγκεκριμένο dataset λόγω του ότι είναι ερμηνεύσιμος, απλός και γρήγορος αλλά και είναι κατάλληλος για συστάδες περίπου ίδιου μεγέθους όπως στο σενάριο που μελετάται.

### Τελική επιλογή αλγορίθμου συσταδοποίησης

Συμπερασματικά, οι περισσότεροι αλγόριθμοι επέλεξαν τα 4 clusters ως τα βέλτιστα για εκείνους. Εφόσον σε ένα περιβάλλον πραγματικού χρόνου πρέπει να βρεθεί μια ισορροπία μεταξύ της λεπτομέρειας του αποτελέσματος συσταδοποίησης και του υπολογιστικού φορτίου, οι αλγόριθμοι που παρέχουν γρήγορη σύγκλιση με λιγότερους πόρους προτιμώνται αφού συμβάλλουν σε μια πιο απλή αλλά και προσαρμόσιμη υποδομή δικτύου. Έτσι, το πείραμα προχωρά στην φάση κατανομής των πόρων χρησιμοποιώντας τα αποτελέσματα (ετικέτες) του K-Means για 4 συστάδες.



**Εικόνα 14:** Αναλυτικό point-by-point silhouette score για τον αλγόριθμο K-Means

Στην Εικόνα 14 παρουσιάζεται και ένα πιο αναλυτικό Silhouette Score για τον αλγόριθμο που επιλέχθηκε (με τον βέλτιστο για εκείνον αριθμό από συστάδες). Με την κόκκινη διακεκομμένη γραμμή φαίνεται το μέσο σκορ των σημείων, με διαφορετικό χρώμα φαίνονται τα διαφορετικά clusters (και στον άξονα y οι ετικέτες τους) και καθεμία από τις οριζόντιες μπάρες αντιστοιχεί στο σκορ που πετυχαίνει ένα σημείο δεδομένων ή αλλιώς ένας χρήστης.

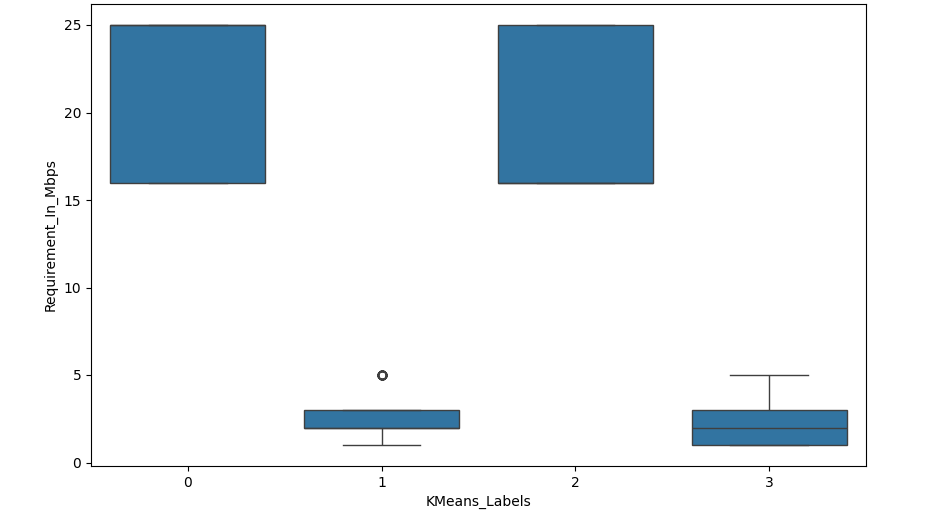
# 

# Αποτελέσματα των πειραμάτων

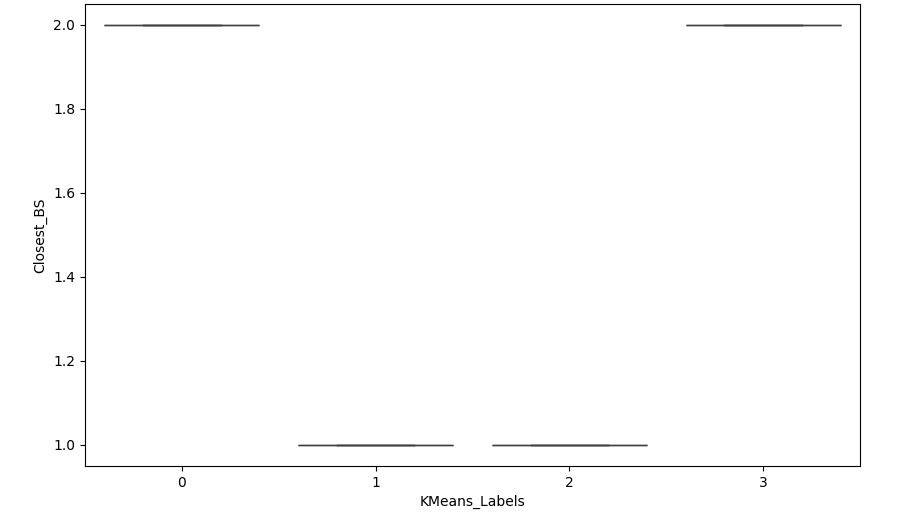
## Τα συμπεράσματα της συσταδοποίησης

Σε αυτή την ενότητα παρουσιάζονται τα συμπεράσματα που μπορούν να εξαχθούν από τους αλγορίθμους συσταδοποίησης για τους χρήστες του πειράματος και τις συστάδες που δημιουργήθηκαν.

Στην Εικόνα 15 και 16 παρουσιάζονται τα αποτελέσματα του K-Means Clustering όσον αφορά τις απαιτήσεις (Εικόνα 15) και την κοντινότερη κεραία (Εικόνα 16) ανάλογα με το cluster στο οποίο βρίσκονται. Οι συστάδες 0 και 2 έχουν μεγάλες απαιτήσεις σε Mbps και οι 1 και 3 μικρές. Οι συστάδες 0 και 3 βρίσκονται πιο κοντά στην κεραία 1 και οι συστάδες 1 και 2 πιο κοντά στην κεραία 2. Συμπεραίνουμε επίσης πως οι συστάδες 0 και 2 αποτελούνται κυρίως από υπηρεσίες 4k streaming, HDTV, podcast ενώ οι 1 και 3 κυρίως από web browsing και video/voice calls.

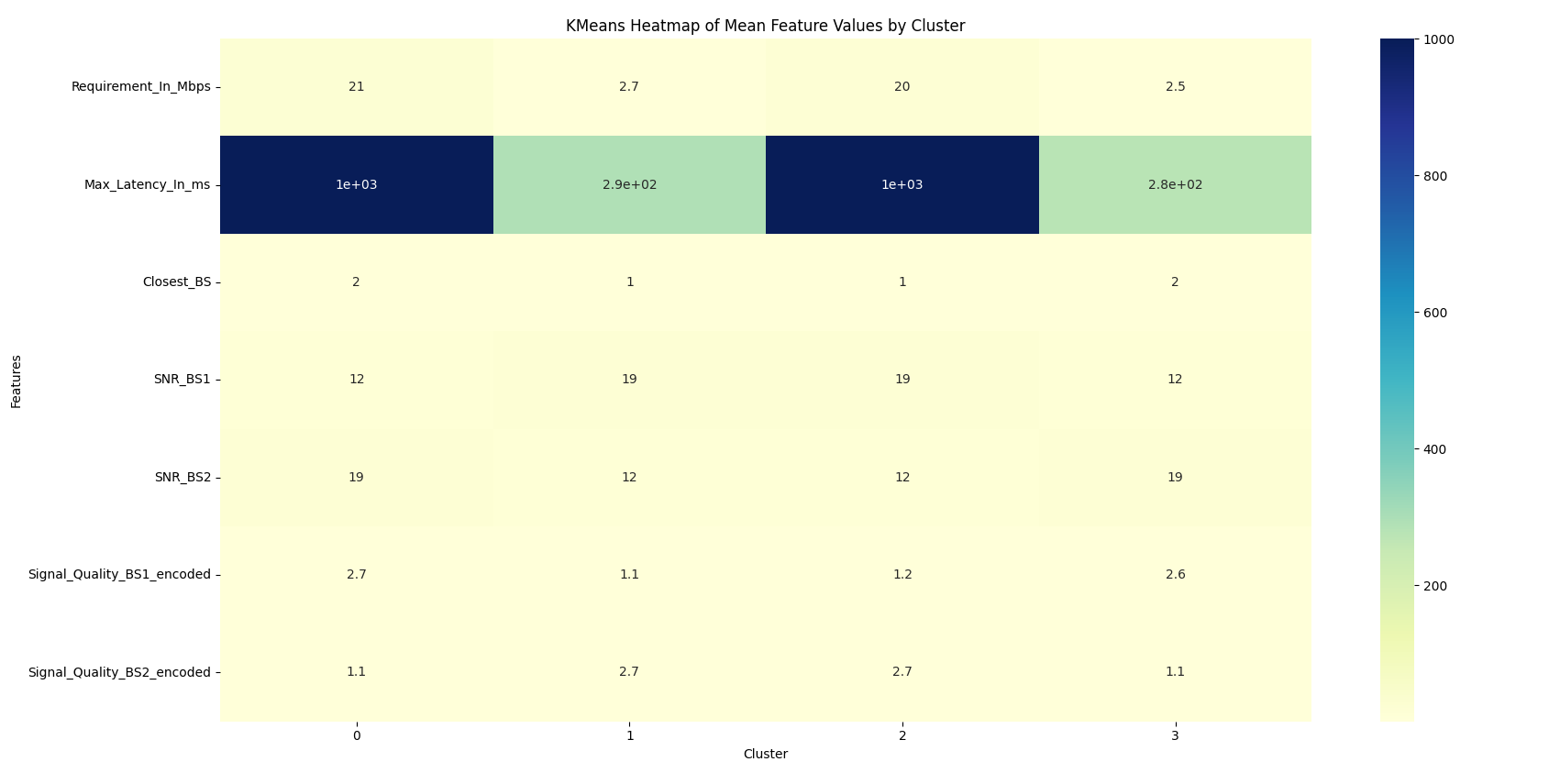


**Εικόνα 15:** Γράφημα (boxplot) απαιτήσεων των χρηστών ανάλογα με το cluster τους

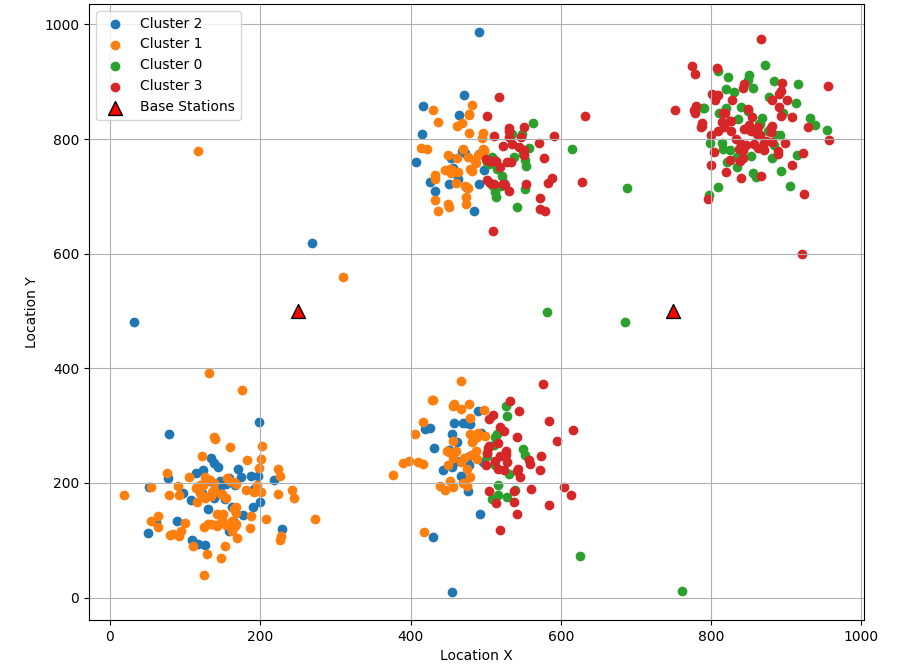


**Εικόνα 16:** Γράφημα κοντινότερων κεραιών των χρηστών ανάλογα με το cluster τους

Στο heatmap της Εικόνας 17, παρουσιάζονται αναλυτικά τα χαρακτηριστικά των χρηστών του κάθε cluster και τέλος στην Εικόνα 18 φαίνεται ο «χάρτης» με τις κεραίες και τους χρήστες (στους άξονες χ και ψ είναι οι συντεταγμένες), με τον κάθε χρήστη να είναι «χρωματισμένος» με το χρώμα του cluster του (πράσινο για το 0, κίτρινο για το 1, μπλε για το 2 και κόκκινο για το 3).



**Εικόνα 17:** Γράφημα (heatmap) των μέσων τιμών των χαρακτηριστικών των χρηστών ανά cluster

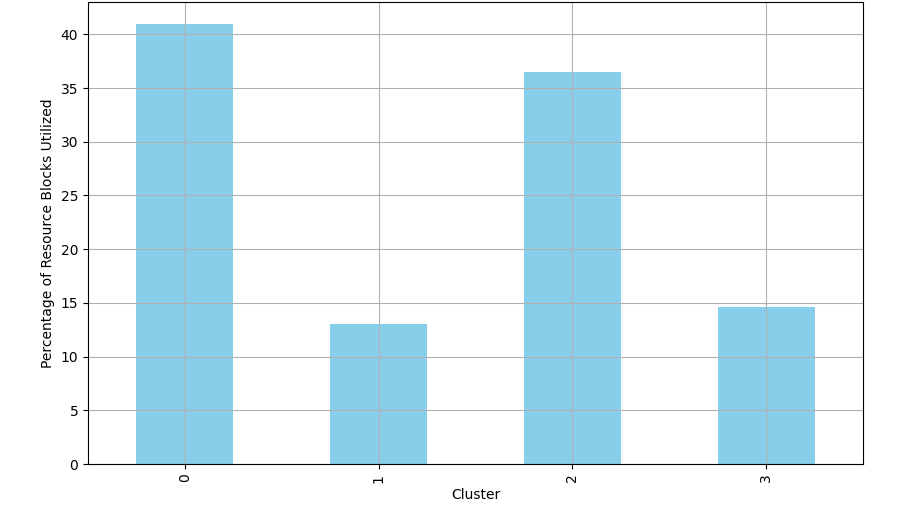


**Εικόνα 18:** Γράφημα των τοποθεσιών των χρηστών ανάλογα με το cluster τους

Σύμφωνα λοιπόν με τις εικόνες αυτές, παρατηρείται πως οι 4 συστάδες χωρίστηκαν ως εξής:

* **Συστάδα 0,** που περιέχει τους χρήστες με υπηρεσίες 4k streaming, HDTV, podcast και βρίσκονται πιο κοντά στην κεραία 2 και έχουν μεγάλη απαίτηση, μεγάλη μέγιστη αποδεκτή καθυστέρηση, καλύτερη ποιότητα σήματος και SNR στην κεραία 2.
* **Συστάδα 1,** που περιέχει τους χρήστες με υπηρεσίες web browsing και video/voice calls που βρίσκονται πιο κοντά στην κεραία 1 και έχουν μικρή απαίτηση και μικρή μέγιστη αποδεκτή καθυστέρηση, καλύτερη ποιότητα σήματος και καλύτερο SNR στην κεραία 1.
* **Συστάδα 2,** που περιέχει τους χρήστες με υπηρεσίες 4k streaming, HDTV, podcast που βρίσκονται πιο κοντά στην κεραία 1 και έχουν μεγάλη απαίτηση και μεγάλη μέγιστη αποδεκτή καθυστέρηση, καλύτερη ποιότητα σήματος και SNR στην κεραία 1.
* **Συστάδα 3,** που περιέχει τους χρήστες με υπηρεσίες web browsing και video/voice calls που βρίσκονται πιο κοντά στην κεραία 2 και έχουν μικρή απαίτηση και μικρή μέγιστη αποδεκτή καθυστέρηση, καλύτερη ποιότητα σήματος και καλύτερο SNR στην κεραία 2.

Παρατηρώντας ξανά την Εικόνα 10, γίνεται αντιληπτό πως ο επόμενος «καλύτερος» αριθμός συστάδων είναι 6. Σε αυτή την περίπτωση θα είχαμε πιο λεπτομερή ομαδοποίηση των χρηστών και διαχωρισμούς όπως χρήστες που έχουν μεγάλη απαίτηση και βρίσκονται πιο κοντά στην κεραία 1 αλλά είναι κοντά και στην κεραία 2, χρήστες με μεγάλη απαίτηση που βρίσκονται πιο κοντά στην κεραία 1 και είναι μακριά από την κεραία 2 κ.α. Συμπεραίνουμε λοιπόν πως ενώ τα score που δείχνουν τη συνοχή/πυκνότητα των συστάδων και αποσκοπούν στην μεγαλύτερη δυνατή απόσταση μεταξύ συστάδων, μπορούν να μας δώσουν τις βέλτιστες παραμέτρους και το βέλτιστο clustering, θα πρέπει να δοκιμάζονται διαφορετικά επίπεδα λεπτομέρειας αν η επιθυμία μας είναι να ανακαλυφθούν πιο περίπλοκες σχέσεις μεταξύ των δεδομένων. Στον συγκεκριμένο αλγόριθμο όμως, εντοπίζεται και χρησιμοποιείται αυτόματα ο αριθμός των συστάδων που θα επιστρέψει το καλύτερο score έτσι ώστε να υπάρχει ευκολότερη εξήγηση των πειραμάτων και μεγαλύτερη σταθερότητα και πυκνότητα των συστάδων.



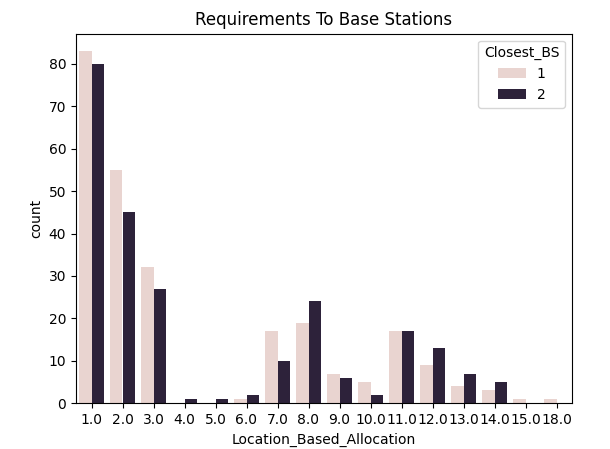
**Εικόνα 19:** Ποσοστό RBs που χρησιμοποιήθηκαν από κάθε cluster στον δυναμικό αλγόριθμο

Στην Εικόνα 19 παρουσιάζεται το ποσοστό των RBs που χρησιμοποιήθηκαν από το κάθε cluster στον δυναμικό αλγόριθμο κατανομής. Χρησιμοποιήθηκε το 100% των RBs αποδοτικά από τον αλγόριθμο αυτόν και όλοι οι χρήστες μπόρεσαν να συνδεθούν και να ικανοποιήσουν τις απαιτήσεις των υπηρεσιών τους, ακόμα και όταν χρειάστηκε να «δώσουν» RBs σε άλλους χρήστες. Περισσότερα RBs χρησιμοποιήθηκαν από τις συστάδες 0 και 2 που είχαν και τις μεγαλύτερες απαιτήσεις και ύστερα από την 3 και τέλος την 1. Οι λόγοι που η συστάδα 0 είχε μεγαλύτερη κατανάλωση από τη 2 και η 3 από την 1 είναι η ποιότητα σήματος των συστάδων και ο αριθμός των χρηστών μέσα στις συστάδες.

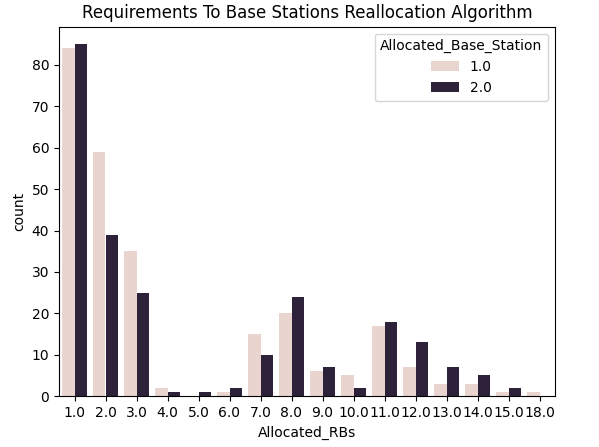
## Αποτελέσματα και σύγκριση των αλγορίθμων κατανομής πόρων

Σε αυτή την ενότητα, παρουσιάζονται τα αποτελέσματα των αλγορίθμων και η σύγκρισή τους.

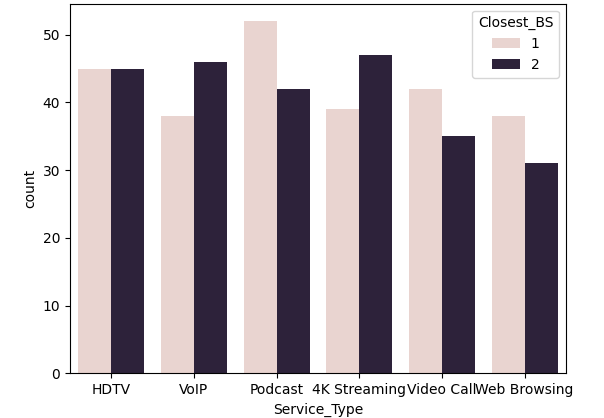
Στις εικόνες 20 και 21 , παρατηρείται η κατανομή των RBs στους δύο αλγορίθμους ενώ στις 22 και 23 φαίνεται η κατανομή των διάφορων τύπων υπηρεσιών στις κεραίες. Τα συμπεράσματα που μπορούν να εξαχθούν, είναι πως με τον δυναμικό αλγόριθμο η κατανάλωση RB από τους χρήστες γίνεται πιο αποδοτική εφόσον όλοι οι χρήστες κατάφεραν να επιτύχουν σύνδεση ακόμα και σε χαμηλότερη ποιότητα, εφόσον φαίνεται πως σε αντίθεση με τον απλό αλγόριθμο υπήρχε λιγότερη κατανάλωση 8,9,11 RBs, και περισσότερη σε 1,2,3 και 4 RBs. Επιπλέον, για τις υπηρεσίες γίνεται αντιληπτό πως περισσότερες απαιτήσεις υπηρεσιών κατάφεραν να ικανοποιηθούν από τις κεραίες στο δεύτερο σενάριο (δυναμικός αλγόριθμος).



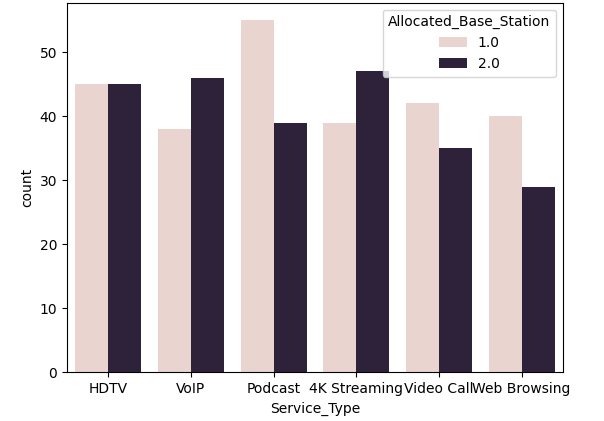
**Εικόνα 20:** Κατανομή των απαιτήσεων των χρηστών στις κεραίες στον απλό αλγόριθμο



**Εικόνα 21:** Κατανομή των απαιτήσεων των χρηστών στις κεραίες στον δυναμικό αλγόριθμο

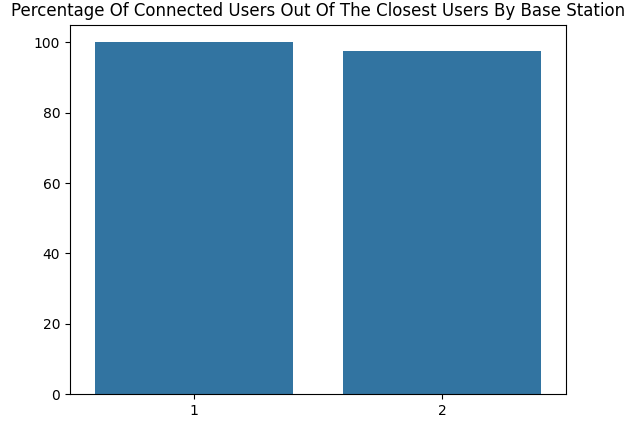


**Εικόνα 22:** Κατανομή των υπηρεσιών των χρηστών στις κεραίες στον απλό αλγόριθμο



**Εικόνα 23:** Κατανομή των υπηρεσιών των χρηστών στις κεραίες στον δυναμικό αλγόριθμο

Τέλος, στην Εικόνα 24 παρουσιάζεται το ποσοστό των χρηστών που μπόρεσαν να εισαχθούν σε κάποια κεραία και να έχουν ικανοποιημένο αίτημα (αριστερά ο δυναμικός αλγόριθμος και δεξιά ο απλός/στατικός). 5 χρήστες από τους 500 δεν μπόρεσαν να ικανοποιηθούν από τον απλό αλγόριθμο και 11 RB δεν δόθηκαν πουθενά. Σε αντίθεση με τον στατικό, ο δυναμικός αλγόριθμος κατάφερε να τοποθετήσει όλους τους χρήστες σε κεραίες και μόνο 6 από τους 500 χρειάστηκε να χαμηλώσουν την ποιότητα της υπηρεσίας τους με σκοπό να μπορέσουν να διατεθούν τα απαραίτητα RBs για όλους.



**Εικόνα 24:** Ποσοστό χρηστών που εξυπηρετήθηκαν σε κάθε αλγόριθμο (δυναμικός αλγόριθμος - απλός αλγόριθμος)

## Συμπεράσματα πειραμάτων

Στα πλαίσια της παρούσας διπλωματικής εργασίας, αναπτύχθηκε ένα σύστημα αλγορίθμων μη επιβλεπόμενης ML με σκοπό την εξήγηση των δεδομένων των χρηστών και την παρουσίαση όμοιων χρηστών σε ομάδες, καθώς και ένας αλγόριθμος κατανομής πόρων για δίκτυα 5G MIMO με τη χρήση των αποτελεσμάτων των τεχνικών μη εποπτευόμενης ML. Η κύρια καινοτομία του προτεινόμενου αυτού αλγορίθμου είναι η δυνατότητα του να εξυπηρετεί όλους τους χρήστες, θυσιάζοντας ελαφρώς την ποιότητα της υπηρεσίας ορισμένων χρηστών, με στόχο την καλύτερη συνολική απόδοση του δικτύου.

Η έρευνα αυτή αποσκοπούσε στην επίδειξη των αλγορίθμων συσταδοποίησης, στη βελτιστοποίηση τους και στην επίδειξη των αποτελεσμάτων και των συμπερασμάτων τους επάνω σε δεδομένα χρηστών 5G καθώς και στη βελτιστοποίηση της κατανομής πόρων στα δίκτυα 5G MIMO χρησιμοποιώντας τα αποτελέσματα και τα συμπεράσματα αυτά.

Αποδείχθηκε ότι οι μέθοδοι μη επιβλεπόμενης ML όπως οι K-Means, DBSCAN, spectral clustering και GMM μπορούν να ομαδοποιήσουν αποτελεσματικά τους χρήστες με βάση τις απαιτήσεις τους και τα χαρακτηριστικά του σήματός τους και πως η HC μπορεί να χρησιμέψει ως βάση για τους άλλους αλγορίθμους. Οι ομάδες που προέκυψαν αντικατόπτριζαν σημαντικές διαφορές στις απαιτήσεις των χρηστών γεγονός που μπορεί να αξιοποιηθεί στην κατανομή πόρων. Ο προτεινόμενος αλγόριθμος εστίαζε στην ικανοποίηση όλων των χρηστών μέσω της στρατηγικής διαχείρισης των πόρων (σε RBs) και της απόδοσης προτεραιοτήτων και διαχείρισης των χρηστών στις διάφορες ομάδες/συστάδες. Τα ευρήματα δείχνουν ότι ο αλγόριθμος κατάφερε να κατανείμει πόρους σε όλους τους χρήστες, αυξάνοντας την απόδοση του δικτύου και την ικανοποίηση των χρηστών, καθώς και να ισορροπήσει αποτελεσματικά το δεδομένο εύρος ζώνης σύμφωνα με την κίνηση του δικτύου. Εξασφαλίστηκε λοιπόν μια ολοκληρωμένη κατανομή των πόρων σε όλους τους χρήστες, επιτρέποντας σε ορισμένους χρήστες να αποδεχθούν χαμηλότερα επίπεδα QoS, βελτιστοποιώντας τη συνολική απόδοση του δικτύου. Συγκεκριμένα, έξι χρήστες αποδέχθηκαν μείωση στην ποιότητα της υπηρεσίας τους ώστε να επιτρέψουν σε πέντε άλλους χρήστες να κατανεμηθούν στις επιθυμητές βάσεις σταθμών τους, αν και με χαμηλότερο επίπεδο QoS.

Τέλος, η συγκριτική ανάλυση με τις παραδοσιακές μεθόδους κατανομής πόρων έδειξε ότι ο προτεινόμενος αλγόριθμος υπερέχει σημαντικά όσον αφορά την ικανοποίηση των χρηστών και την αξιοποίηση των πόρων. Μπόρεσε να διαχειριστεί υποβαθμίσεις υπηρεσιών με ευελιξία, εξασφαλίζοντας ότι οι κρίσιμες υπηρεσίες (π.χ. VoIP) διατηρούν την ποιότητα τους, ενώ άλλες υπηρεσίες προσαρμόστηκαν ανάλογα με τις ανάγκες. Στον απλό αλγόριθμο, όπου η κατανομή των πόρων εστίασε απλά στο καλύτερο SNR, πέντε χρήστες δεν κατανεμήθηκαν πουθενά και έντεκα RBs από τον σταθμό βάσης 2 έμειναν αχρησιμοποίητα. Αντίθετα, ο προτεινόμενος αλγόριθμος εξασφάλισε ότι και οι 500 χρήστες συνδέθηκαν και όλα τα RBs χρησιμοποιήθηκαν πλήρως. Αν και όπως αναφέρθηκε, έξι χρήστες έλαβαν μείωση στην ποιότητα των υπηρεσιών τους για να κατανεμηθούν πέντε άλλοι χρήστες, η αντιστάθμιση που έγινε επέτρεψε στο δίκτυο να βελτιστοποιήσει την κατανομή των πόρων, να χρησιμοποιήσει πλήρως τα RBs και να διαχειριστεί πιο αποτελεσματικά και δίκαια ένα ευρύ φάσμα απαιτήσεων των χρηστών.

## Περιπτώσεις χρήσης των συμπερασμάτων της συσταδοποίησης στο 5G

Οι επιπτώσεις αυτής της έρευνας είναι σημαντικές για τον σχεδιασμό και τη διαχείριση των μελλοντικών δικτύων 5G. Χρησιμοποιώντας τεχνικές μη επιβλεπόμενης μάθησης, αναδείχθηκε μία προσαρμοστική προσέγγιση για την κατανομή των πόρων που μπορεί να ανταποκρίνεται δυναμικά στις μεταβαλλόμενες συνθήκες του δικτύου και τις απαιτήσεις των χρηστών. Αυτό συμβάλλει στον ευρύτερο τομέα των τηλεπικοινωνιών, παρέχοντας ένα ανθεκτικό πλαίσιο που βελτιώνει την αποδοτικότητα και τη δικαιοσύνη της κατανομής πόρων σε πολύπλοκα περιβάλλοντα δικτύων. Τα ευρήματα υποδηλώνουν ότι οι στρατηγικές αντισταθμίσεις QoS μπορούν να βελτιώσουν σημαντικά την συνολική απόδοση του δικτύου και την ικανοποίηση των χρηστών.

Σε μία ευρύτερη κλίμακα, οι αλγόριθμοι μη επιβλεπόμενης ML μπορούν να βοηθήσουν τα δίκτυα 5G και με περισσότερους τρόπους. Για παράδειγμα, η ομαδοποίηση που γίνεται με τους αλγορίθμους αυτούς, μπορεί να βοηθήσει στην διαχείριση πόρων με βάση τις μεμονωμένες ανάγκες κάθε συστάδας, καθώς και στον ευκολότερο χειρισμό των χρηστών εφόσον χρησιμοποιώντας τις ετικέτες που δίνονται, υπάρχει ευκολότερη και γρηγορότερη πρόσβαση σε όμοιους χρήστες. Επιπλέον, τα ευρήματα από την συσταδοποίηση μπορεί να βοηθήσουν σε διάφορες περιπτώσεις όπως:

* **Προβλεπτική Μοντελοποίηση** για την ενίσχυση της διαχείρισης των πόρων δικτύου σε περιβάλλοντα 5G. Ενσωματώνοντας τα αποτελέσματα της συσταδοποίησης ως χαρακτηριστικό, μπορούν να αναπτυχθούν μοντέλα όπως τα Random Forests ή τα Gradient Boosting Machines για την πρόβλεψη μελλοντικών αναγκών του δικτύου. Αυτή η προσέγγιση αξιοποιεί τα μοτίβα και τις σχέσεις που ανακαλύφθηκαν μέσω της συσταδοποίησης για να προβλέψει με μεγαλύτερη ακρίβεια τη συμπεριφορά των χρηστών και το φορτίο του δικτύου. Εκπαιδεύοντας μοντέλα σε ιστορικά δεδομένα που περιλαμβάνουν ετικέτες συστάδων, το μοντέλο μαθαίνει να συσχετίζει συγκεκριμένες συμπεριφορές συστάδων με μελλοντικές ανάγκες του δικτύου. Αυτή η προβλεπτική ικανότητα θα επιτρέψει στους διαχειριστές δικτύου να προσαρμόζουν προληπτικά την ανάθεση πόρων και να βελτιστοποιούν την απόδοση του δικτύου πριν από τις περιόδους αιχμής, διασφαλίζοντας έτσι την αποδοτική χρήση της υποδομής και την διατήρηση υψηλών επιπέδων QoS.
* **Προσαρμοστικές στρατηγικές ανάθεσης πόρων,** όπου οι πόροι του δικτύου ανατίθενται δυναμικά σύμφωνα με τις προσδιορισμένες συστάδες, εξασφαλίζοντας την βέλτιστη χρήση και απόδοση. Μέσω της παρούσας εργασίας, φαίνεται πως η συσταδοποίηση αποτελεί καίριο αλγόριθμο στη δημιουργία προηγμένων μοντέλων ML που μπορούν να κατηγοριοποιήσουν τους χρήστες δικτύου 5G με βάση διάφορα κριτήρια, όπως ο τύπος υπηρεσίας, οι απαιτήσεις εύρους ζώνης και οι ευαισθησίες στην καθυστέρηση. Χρησιμοποιώντας αλγορίθμους μη εποπτευόμενης ML, μπορεί να εξερευνηθεί περαιτέρω η πολυδιάστατη φύση της συμπεριφοράς των χρηστών και η αλληλεπίδρασή τους με το δίκτυο. Αυτή η προσέγγιση, σε συνδυασμό και με άλλους αλγορίθμους ML και περισσότερα δεδομένα, θα μπορέσει να ξεπεράσει τις στατικές τεχνικές ανάθεσης εισάγοντας μια μεθοδολογία που βασίζεται στα δεδομένα και προσαρμόζεται στη συμπεριφορά των χρηστών και στις καταστάσεις του δικτύου σε πραγματικό χρόνο.
* **Ενίσχυση της αποδοτικότητας του δικτύου και της QoS**. Η έρευνα που περιγράφεται σε αυτή την εργασία τονίζει την ανάγκη για σημαντικές βελτιώσεις στην QoS με την προσαρμογή των πόρων του δικτύου ώστε να ταιριάζουν στις απαιτήσεις των χρηστών. Τα αποτελέσματα της συσταδοποίησης έχουν προσφέρει πολύτιμες πληροφορίες για το πώς μπορούν να διαχειριστούν και να κατανεμηθούν πιο αποδοτικά οι πόροι, επιτρέποντας την βέλτιστη κατανομή πόρων (χωρίς την ανάθεση παραπάνω RBs από τα βέλτιστα) στους χρήστες με τη σωστή σειρά προτεραιότητας. Μια τέτοια στοχευμένη κατανομή εξασφαλίζει ότι οι χρήστες λαμβάνουν το απαραίτητο εύρος ζώνης και δεν ξεπερνάται η μέγιστη καθυστέρηση για τις εφαρμογές τους, κάτι που είναι ιδιαίτερα κρίσιμο για υπηρεσίες όπως η VoIP και η ροή βίντεο που είναι ευαίσθητες στην απόδοση του δικτύου.
* **Τοπολογικές βελτιώσεις του δικτύου**, υποδεικνύοντας περιοχές όπου η ενίσχυση της κάλυψης ή η αναβάθμιση της υποδομής θα μπορούσε να οδηγήσει σε καλύτερη QoS των χρηστών χρησιμοποιώντας τις γεωγραφικές πληροφορίες που προκύπτουν από τις αναλύσεις συσταδοποίησης με βάση την τοποθεσία και την απαίτηση. Αυτές οι πληροφορίες μπορούν να επισημάνουν συγκεκριμένες τοποθεσίες όπου οι χρήστες βιώνουν υποβάθμιση υπηρεσίας ή όπου η ζήτηση υπερβαίνει συνεχώς την προσφορά, καθοδηγώντας βελτιώσεις της υποδομής, όπως η τοποθέτηση πρόσθετων base stations. Επιπλέον, αυτός ο γεωγραφικός προσδιορισμός βοηθά στην κατανόηση της χωρικής κατανομής διαφόρων τύπων υπηρεσιών, που είναι ανεκτίμητη για τον σχεδιασμό του δικτύου ή την επέκτασή του σε περιοχές που δεν εξυπηρετούνται επαρκώς, αντιμετωπίζοντας τυχόν σημεία συμφόρησης σε περιόδους αιχμής διασφαλίζοντας παράλληλα τη δίκαιη κατανομή των υπηρεσιών στις κεραίες και την ισορροπία του φόρτου.

## Μελλοντική εργασία

Η έρευνα που πραγματοποιήθηκε ανοίγει τον δρόμο για διάφορες προεκτάσεις και βελτιώσεις. Στο μέλλον, θα μπορούσαν να διερευνηθούν οι ακόλουθες κατευθύνσεις:

* **Η Ενσωμάτωση Περισσότερων Αλγορίθμων Μάθησης**, όπως της ενισχυτικής μάθησης, που μπορεί να προσφέρει νέες ευκαιρίες για περαιτέρω βελτιστοποίηση της κατανομής πόρων. Οι αλγόριθμοι ενισχυτικής μάθησης έχουν τη δυνατότητα να προσαρμόζονται δυναμικά σε αλλαγές στο περιβάλλον του δικτύου και να βελτιώνουν τις αποφάσεις κατανομής σε πραγματικό χρόνο, αυξάνοντας έτσι την αποδοτικότητα και την προσαρμοστικότητα του δικτύου. Επίσης μπορούν να χρησιμοποιηθούν προβλεπτικοί αλγόριθμοι μαζί με τα αποτελέσματα της συσταδοποίησης με σκοπό ο κώδικας να μπορεί να χειριστεί σενάρια με μεγαλύτερη πολυπλοκότητα όπως χρήστες με μεταβλητό τύπο κίνησης.
* **Εφόσον η κατανάλωση ενέργειας αποτελεί ένα κρίσιμο ζήτημα στα σύγχρονα δίκτυα 5G**, η μελλοντική έρευνα θα μπορούσε να επικεντρωθεί στη διαχείριση της ενεργειακής απόδοσης των σταθμών βάσης και των χρηστών, αναπτύσσοντας αλγόριθμους που βελτιστοποιούν την κατανομή πόρων με βάση την κατανάλωση ενέργειας. Αυτό θα μπορούσε να οδηγήσει σε πιο βιώσιμες λύσεις, μειώνοντας το συνολικό ενεργειακό αποτύπωμα του δικτύου και ενισχύοντας την περιβαλλοντική βιωσιμότητα.
* **Επέκταση σε Δίκτυα B5G και 6ης Γενιάς (6th Generation - 6G)** αφού είναι εμφανές πως η προσαρμογή και εφαρμογή των προτεινόμενων αλγορίθμων σε μελλοντικά δίκτυα πέρα από το 5G, όπως τα δίκτυα 6G, είναι αναγκαία. Τα δίκτυα 6G αναμένεται να έχουν ακόμα μεγαλύτερες απαιτήσεις σε πόρους και ταχύτητες, και η αποτελεσματική κατανομή των πόρων θα είναι καθοριστική για την επιτυχία τους [37],[38]. Η έρευνα μπορεί να επικεντρωθεί στην ανάπτυξη αλγορίθμων που θα υποστηρίζουν τις αυξημένες απαιτήσεις και την πολυπλοκότητα αυτών των μελλοντικών δικτύων.
* **Πειραματική αξιολόγηση των αλγορίθμων σε πραγματικά περιβάλλοντα δικτύων 5G** η οποία θα προσφέρει πολύτιμες πληροφορίες για την αποτελεσματικότητά τους υπό διαφορετικές συνθήκες και σενάρια χρήσης. Αυτό θα επιτρέψει την αναγνώριση και αντιμετώπιση προβλημάτων που ενδέχεται να προκύψουν σε πραγματικές συνθήκες, και θα βοηθήσει στη βελτίωση της πρακτικής εφαρμογής των αλγορίθμων.
* **Ανάλυση ανθεκτικότητας** **των προτεινόμενων αλγορίθμων** σε συνθήκες σφάλματος ή απώλειας συνδεσιμότητας αποτελεί ένα κρίσιμο ζήτημα για την αξιοπιστία του δικτύου. Η μελέτη της ανθεκτικότητας των αλγορίθμων σε τέτοιες συνθήκες θα διασφαλίσει ότι το δίκτυο μπορεί να συνεχίσει να λειτουργεί αποδοτικά και αξιόπιστα ακόμη και σε περιπτώσεις έκτακτης ανάγκης, προσφέροντας σταθερή και ασφαλή σύνδεση στους χρήστες.

Επιπλέον, η ενσωμάτωση τεχνικών ML με άλλες τεχνολογίες, όπως το IoT και τα αυτόνομα οχήματα, θα μπορούσε να οδηγήσει σε νέες και καινοτόμες εφαρμογές που θα ενισχύσουν την αποτελεσματικότητα και την ευελιξία των δικτύων. Η έρευνα αυτή μπορεί να αποτελέσει τη βάση για τη δημιουργία πιο αποδοτικών, αξιόπιστων και βιώσιμων συστημάτων κατανομής πόρων στα δίκτυα 5G, ανοίγοντας τον δρόμο για την επόμενη γενιά τεχνολογιών και υπηρεσιών.

Με την ενσωμάτωση αυτών των προεκτάσεων, η παρούσα εργασία μπορεί να αποτελέσει ένα ισχυρό θεμέλιο για τη συνεχιζόμενη βελτίωση και ανάπτυξη των δικτύων κινητής τηλεφωνίας, εξασφαλίζοντας την παροχή υψηλής ποιότητας υπηρεσιών στους χρήστες και τη βέλτιστη εκμετάλλευση των διαθέσιμων πόρων.

Βιβλιογραφία- Αναφορές

1. S. Li, L. D. Xu, and S. Zhao, "5G Internet of Things: A survey," Journal of Industrial Information Integration, vol. 10, pp. 1-9, 2018, doi: 10.1016/j.jii.2018.01.005.
2. N. Ferdosian, S. Berri and A. Chorti, "5G New Radio Resource Allocation Optimization for Heterogeneous Services," 2022 International Symposium ELMAR, Zadar, Croatia, 2022, pp. 1-6, doi: 10.1109/ELMAR55880.2022.9899817.
3. A. Patil, S. Iyer, O. L. A. Lopez, R. J. Pandya, K. Pai, A. Kalla, and R. Kallimani, "A Comprehensive Survey on Spectrum Sharing Techniques for 5G/B5G Intelligent Wireless Networks: Opportunities, Challenges and Future Research Directions," arXiv preprint arXiv:2211.08956, 2022.
4. B. Maaz, "Radio resource allocation in 5G wireless networks," Ph.D. dissertation, Université Paris Saclay (COmUE), Mar. 2017. [Online]. Available: <https://theses.hal.science/tel-01558458>
5. Y. Y. Munaye, R. -T. Juang, H. -P. Lin, G. B. Tarekegn, D. -B. Lin and S. -S. Jeng, "Radio Resource Allocation for 5G Networks Using Deep Reinforcement Learning," 2021 30th Wireless and Optical Communications Conference (WOCC), Taipei, Taiwan, 2021, pp. 66-69, doi: 10.1109/WOCC53213.2021.9603111.
6. M. Alsenwi, K. Kim, and C. S. Hong, "Radio Resource Allocation in 5G New Radio: A Neural Networks Approach," Journal of KIISE, vol. 46, no. 9, pp. 961-967, Sep. 2019, doi: 10.5626/jok.2019.46.9.961.
7. C. Bouras, D. Diasakos, A. Gkamas, V. Kokkinos, P. Pouyioutas and N. Prodromos, "Evaluation of User Allocation Techniques in Massive MIMO 5G Networks," *2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Istanbul, Turkiye, 2023, pp. 1-6, doi: 10.1109/WINCOM59760.2023.10322955.
8. Prodromos, N., Gkamas, A., Diasakos, D., Bouras, C., Kokkinos, V., & Pouyioutas, P. (2024). Dynamic Bandwidth Allocation in MIMO 5G Networks. In *The 20th International Wireless Communications & Mobile Computing Conference.*
9. A. Madelkhanova, "Machine Learning-based Virtualized 5G Resource Allocation," 2019.
10. Bouras, C.J., Michos, E., Prokopiou, I. (2022). Applying Machine Learning and Dynamic Resource Allocation Techniques in Fifth Generation Networks. In: Barolli, L., Hussain, F., Enokido, T. (eds) Advanced Information Networking and Applications. AINA 2022. Lecture Notes in Networks and Systems, vol 449. Springer, Cham. <https://doi.org/10.1007/978-3-030-99584-3_57>
11. Ferrús, R. [et al.]. Machine learning-assisted cross-slice radio resource optimization: Implementation framework and algorithmic solution. "ITU journal on future and evolving technologies (ITU J-FET)", Desembre 2020, vol. 1, núm. 1, p. 1-18.
12. A. Margaris, I. Filippas, and K. Tsagkaris, "Hybrid Network–Spatial Clustering for Optimizing 5G Mobile Networks," Applied Sciences, vol. 12, no. 3, p. 1203, 2022, doi: 10.3390/app12031203.
13. Bouras, Christos, and Ioannis Prokopiou. "Performance Analysis of MIMO using Machine Learning in 5G Networks."
14. M. E. Morocho Cayamcela and W. Lim, "Artificial Intelligence in 5G Technology: A Survey," 2018 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2018, pp. 860-865, doi: 10.1109/ICTC.2018.8539642.
15. V. P. Rekkas, S. Sotiroudis, P. Sarigiannidis, G. K. Karagiannidis and S. K. Goudos, "Unsupervised Machine Learning in 6G Networks -State-of-the-art and Future Trends," 2021 10th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece, 2021, pp. 1-4, doi: 10.1109/MOCAST52088.2021.9493388.
16. N. Ferdosian, S. Berri and A. Chorti, "5G New Radio Resource Allocation Optimization for Heterogeneous Services," 2022 International Symposium ELMAR, Zadar, Croatia, 2022, pp. 1-6, doi: 10.1109/ELMAR55880.2022.9899817.
17. A. Patil, S. Iyer, O. L. A. Lopez, R. J. Pandya, K. Pai, A. Kalla, and R. Kallimani, "A Comprehensive Survey on Spectrum Sharing Techniques for 5G/B5G Intelligent Wireless Networks: Opportunities, Challenges and Future Research Directions," arXiv preprint arXiv:2211.08956, 2022.
18. R. I. Rony, E. Lopez-Aguilera and E. Garcia-Villegas, "Dynamic Spectrum Allocation Following Machine Learning-Based Traffic Predictions in 5G," in IEEE Access, vol. 9, pp. 143458-143472, 2021, doi: 10.1109/ACCESS.2021.3122331.
19. J. G. Andrews et al., "What Will 5G Be?," IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1065-1082, June 2014, doi: 10.1109/JSAC.2014.2328098.
20. Alves, Hirley, Taneli Riihonen, and Himal A. Suraweera, eds. Full-Duplex Communications for Future Wireless Networks. Singapore: Springer, 2020.
21. X. Zhou, D. Pan, H. Song and X. Huang, "Socially-Aware D2D Pair Strategy: A Stable Matching Approach," 2020 IEEE 39th International Performance Computing and Communications Conference (IPCCC), Austin, TX, USA, 2020, pp. 1-4, doi: 10.1109/IPCCC50635.2020.9391547.
22. G. Zhu, J. Zan, Y. Yang and X. Qi, "A Supervised Learning Based QoS Assurance Architecture for 5G Networks," in IEEE Access, vol. 7, pp. 43598-43606, 2019, doi: 10.1109/ACCESS.2019.2907142.
23. C. Ssengonzi, O. P. Kogeda, and T. O. Olwal, “A survey of deep reinforcement learning application in 5G and beyond network slicing and virtualization”, Array, vol. 14, p. 100142, 2022, doi: 10.1016/j.array.2022.100142.
24. S. Ardabili, A. Mosavi and I. Felde, "Deep learning for 5G and 6G," 2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI), Timisoara, Romania, 2023, pp. 000711-000720, doi: 10.1109/SACI58269.2023.10158628.
25. Girelli Consolaro, N.; Shinde, S.S.; Naseh, D.; Tarchi, D. Analysis and Performance Evaluation of Transfer Learning Algorithms for 6G Wireless Networks. Electronics 2023, 12, 3327. <https://doi.org/10.3390/electronics12153327>
26. A. Priovolos, D. Lioprasitis, G. Gardikis and S. Costicoglou, "Using Anomaly Detection Techniques for Securing 5G Infrastructure and Applications," 2021 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), Athens, Greece, 2021, pp. 519-524, doi: 10.1109/MeditCom49071.2021.9647668.
27. Fakhouri, H.N.; Alawadi, S.; Awaysheh, F.M.; Hani, I.B.; Alkhalaileh, M.; Hamad, F. A Comprehensive Study on the Role of Machine Learning in 5G Security: Challenges, Technologies, and Solutions. Electronics 2023, 12, 4604. https://doi.org/10.3390/electronics12224604
28. Samad MA, Choi DY, Choi K. Path loss measurement and modeling of 5G network in emergency indoor stairwell at 3.7 and 28 GHz. PLoS One. 2023 Mar 28;18(3):e0282781. doi: 10.1371/journal.pone.0282781. PMID: 36976772; PMCID: PMC10047544.
29. Na, Shi, Liu Xumin, and Guan Yong. "Research on k-means clustering algorithm: An improved k-means clustering algorithm." *2010 Third International Symposium on intelligent information technology and security informatics*. Ieee, 2010.
30. Ross, H. H. (1964). Principles of numerical taxonomy. Systematic Biology, 13(1-4), 106-108. <https://doi.org/10.2307/sysbio/13.1-4.106>
31. Khan, Kamran, et al. "DBSCAN: Past, present and future." *The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014)*. IEEE, 2014.
32. Yang, Miin-Shen, Chien-Yo Lai, and Chih-Ying Lin. "A robust EM clustering algorithm for Gaussian mixture models." *Pattern Recognition* 45.11 (2012): 3950-3961.
33. Jia, Hongjie, et al. "The latest research progress on spectral clustering." Neural Computing and Applications 24 (2014): 1477-1486.
34. Syakur, Muhammad Ali, et al. "Integration k-means clustering method and elbow method for identification of the best customer profile cluster." IOP conference series: materials science and engineering. Vol. 336. IOP Publishing, 2018.
35. Chakrabarti, Arijit, and Jayanta K. Ghosh. "AIC, BIC and recent advances in model selection." *Philosophy of statistics* (2011): 583-605.
36. S. Łukasik, P. A. Kowalski, M. Charytanowicz and P. Kulczycki, "Clustering using flower pollination algorithm and Calinski-Harabasz index," *2016 IEEE Congress on Evolutionary Computation (CEC)*, Vancouver, BC, Canada, 2016, pp. 2724-2728, doi: 10.1109/CEC.2016.7744132.
37. N. H. Mahmood, H. Alves, O. A. López, M. Shehab, D. P. M. Osorio and M. Latva-Aho, "Six Key Features of Machine Type Communication in 6G," *2020 2nd 6G Wireless Summit (6G SUMMIT)*, Levi, Finland, 2020, pp. 1-5, doi: 10.1109/6GSUMMIT49458.2020.9083794.
38. K. Jadhav, "Architectures, Opportunities, and Challenges in Emerging Technologies of 6G," *2023 International Conference on Computer and Applications (ICCA)*, Cairo, Egypt, 2023, pp. 1-4, doi: 10.1109/ICCA59364.2023.10401759.

**Παράρτημα Α: Κώδικας των Πειραμάτων**

**Οι βιβλιοθήκες python που χρησιμοποιήθηκαν είναι οι εξής :**

numpy

pandas

matplotlib.pyplot

seaborn

sklearn.cluster για τις συναρτήσεις KMeans, AgglomerativeClustering, DBSCAN, SpectralClustering

sklearn.mixture για τη συνάρτηση GaussianMixture

sklearn.metrics για τις συναρτήσεις silhouette\_score,silhouette\_samples

scipy.cluster.hierarchy για τις συναρτήσεις dendrogram, linkage

sklearn.preprocessing για τις συναρτήσεις MinMaxScaler,StandardScaler

sklearn.decomposition για τη συνάρτηση PCA

sklearn.preprocessing για τη συνάρτηση LabelEncoder

sklearn.metrics για τη συνάρτηση calinsky-harabasz score

math

**Ο κώδικας για την τοποθέτηση των κεραιών και των χρηστών φαίνεται παρακάτω :**

base\_station\_dict = {'Location\_X': [250.0,750.0], 'Location\_Y':[500.0,500.0]}

hotspots = [(150, 170), (850, 820), (500, 250), (500, 750)]

hot\_X = list()

hot\_Y = list()

for hotspot in hotspots:

    hot\_X.extend(np.random.normal(hotspot[0],50,120))

    hot\_Y.extend(np.random.normal(hotspot[1],50,120))

spreadouts\_X = np.random.uniform(0,1000,20)

spreadouts\_Y = np.random.uniform(0,1000,20)

hot\_X.extend(spreadouts\_X)

hot\_Y.extend(spreadouts\_Y)

users\_dict= {'User\_ID':users,'Location\_X':hot\_X,'Location\_Y':hot\_Y,'Service\_Type': np.random.choice(services,len(users))}

**Ο κώδικας που χρησιμοποιήθηκε για το feature engineering είναι ο εξής:**

**Πιο συγκεκριμένα παρακάτω παρουσιάζεται ο κώδικας για τη δημιουργία και ανάθεση RBs στην κάθε κεραία:**

base\_station\_df['Capacity\_In\_KHz'] = total\_network\_bandwidth\_KHZ

base\_station\_df['Resource\_Blocks\_Available'] = base\_station\_df['Capacity\_In\_KHz'].apply(lambda x : math.ceil(x/resource\_block\_size))

total\_resource\_blocks = base\_station\_df['Resource\_Blocks\_Available'].sum()

base\_station\_df['Resource\_Blocks\_Available\_LBALLOC'] = base\_station\_df['Resource\_Blocks\_Available'].copy()

**Εδώ παρουσιάζεται ο κώδικας που χρησιμοποιεί την τοποθεσία του κάθε χρήστη για να βρεθεί η απόστασή του από κάθε κεραία:**

users\_df['Full\_Location']= users\_df.apply(lambda x : [x['Location\_X'],x['Location\_Y']],axis=1)

users\_df['Distance\_from\_BS1'] = users\_df['Full\_Location'].apply(lambda x : math.dist(x,[base\_station\_df.loc[0,'Location\_X'],base\_station\_df.loc[0,'Location\_Y']]))

users\_df['Distance\_from\_BS2'] = users\_df['Full\_Location'].apply(lambda x : math.dist(x,[base\_station\_df.loc[1,'Location\_X'],base\_station\_df.loc[1,'Location\_Y']]))

def closest\_bs(row):

    if row['Distance\_from\_BS1'] <= row['Distance\_from\_BS2']:

        return 1

    else :

        return 2

users\_df['Closest\_BS'] = users\_df.apply(closest\_bs,axis=1)

**Κώδικας για την απεικόνιση του αριθμού των χρηστών που έχουν πιο κοντά τους την κεραία 1 και των χρηστών που έχουν πιο κοντά τους την κεραία 2:**

plt.bar(users\_df['Closest\_BS'].value\_counts().index,users\_df['Closest\_BS'].value\_counts().values)

plt.show()

**Κώδικας για το Hata Okumura Pathloss :**

def find\_PL1(row):

    return 69.55 + 26.16\*math.log10(transmission\_frequency) - 13.82\*math.log10(height\_of\_antenna) - ((1.1\*math.log10(transmission\_frequency)-0.7)\*height\_of\_user - 1.56\*math.log10(transmission\_frequency)-0.8) + (44.9-6.55\*math.log10(height\_of\_antenna))\*math.log10(row['Distance\_from\_BS1']/1000)

def find\_PL2(row):

    return 69.55 + 26.16\*math.log10(transmission\_frequency) - 13.82\*math.log10(height\_of\_antenna) - ((1.1\*math.log10(transmission\_frequency)-0.7)\*height\_of\_user - 1.56\*math.log10(transmission\_frequency)-0.8) + (44.9-6.55\*math.log10(height\_of\_antenna))\*math.log10(row['Distance\_from\_BS2']/1000)

users\_df['Hata\_Okumura\_Pathloss\_BS1']= users\_df.apply(find\_PL1,axis=1)

users\_df['Hata\_Okumura\_Pathloss\_BS2']= users\_df.apply(find\_PL2,axis=1)

**Κώδικας για την ισχύ του σήματος :**

def find\_ss1(row):

   return transmitted\_power - row['Hata\_Okumura\_Pathloss\_BS1'] + bs\_gain + ue\_gain - pen\_loss - cable\_loss

def find\_ss2(row):

    return transmitted\_power - row['Hata\_Okumura\_Pathloss\_BS2'] + bs\_gain + ue\_gain - pen\_loss - cable\_loss

users\_df['Signal\_Strength\_BS1']= users\_df.apply(find\_ss1,axis=1)

users\_df['Signal\_Strength\_BS2']= users\_df.apply(find\_ss2,axis=1)

**Κώδικας για το SNR από κάθε κεραία:**

users\_df['SNR\_BS1']= users\_df['Signal\_Strength\_BS1'].apply(lambda x : x-Noise)

users\_df['SNR\_BS2']= users\_df['Signal\_Strength\_BS2'].apply(lambda x : x-Noise)

**Κώδικας για την εύρεση της ποιότητας του σήματος, την μετατροπή της απαίτησης των χρηστών από mbps σε mhz, ύστερα σε khz και τέλος σε RBs:**

def find\_quality(snr):

    if snr > 40:

        return 'Excellent'

    elif snr > 25 and snr <= 40:

        return 'Great'

    elif snr > 15 and snr <=25:

        return 'Good'

    elif snr >10 and snr <= 15 :

        return 'Ok'

    else :

        return 'Subpar'

def mbps\_to\_mhz(userdf,bs):

    if bs ==1:

        return userdf['Requirement\_In\_Mbps']/math.log2(1+(math.pow(10,(userdf['SNR\_BS1']/10))))

    else:

        return userdf['Requirement\_In\_Mbps']/math.log2(1+(math.pow(10,(userdf['SNR\_BS2']/10))))

users\_df['Signal\_Quality\_BS1'] = users\_df['SNR\_BS1'].apply(find\_quality)

users\_df['Signal\_Quality\_BS2'] = users\_df['SNR\_BS2'].apply(find\_quality)

users\_df['Requirement\_In\_MHz\_BS1'] = users\_df.apply(mbps\_to\_mhz,args=(1,),axis=1)

users\_df['Requirement\_In\_MHz\_BS2'] = users\_df.apply(mbps\_to\_mhz,args=(2,),axis=1)

users\_df['Requirement\_In\_KHz\_BS1'] = users\_df['Requirement\_In\_MHz\_BS1'].apply(lambda x : x\*1000)

users\_df['Requirement\_In\_KHz\_BS2'] = users\_df['Requirement\_In\_MHz\_BS2'].apply(lambda x : x\*1000)

users\_df['Requirement\_In\_RBs\_BS1'] = users\_df['Requirement\_In\_KHz\_BS1'].apply(lambda x: math.ceil(x/resource\_block\_size))

users\_df['Requirement\_In\_RBs\_BS2'] = users\_df['Requirement\_In\_KHz\_BS2'].apply(lambda x: math.ceil(x/resource\_block\_size))

**Κώδικας του απλού allocation σε σχέση με το SNR και την κοντινότερη κεραία :**

users\_df['Location\_Based\_Allocation'] = np.nan

sorted\_users = users\_df.copy()

sorted\_users['BEST\_SNR'] = sorted\_users.apply(lambda row: max(row['SNR\_BS1'],row['SNR\_BS2']),axis=1)

sorted\_users.sort\_values(by='BEST\_SNR',ascending=False,inplace=True)

def location\_allocation(row):

    if row['Closest\_BS']==1:

        if base\_station\_df.loc[0,'Resource\_Blocks\_Available\_LBALLOC'] - row['Requirement\_In\_RBs\_BS1'] >= 0:

            base\_station\_df.loc[0,'Resource\_Blocks\_Available\_LBALLOC'] -= row['Requirement\_In\_RBs\_BS1']

            return row['Requirement\_In\_RBs\_BS1']

    else:

        if base\_station\_df.loc[1,'Resource\_Blocks\_Available\_LBALLOC'] - row['Requirement\_In\_RBs\_BS2'] >= 0:

            base\_station\_df.loc[1,'Resource\_Blocks\_Available\_LBALLOC'] -= row['Requirement\_In\_RBs\_BS2']

            return row['Requirement\_In\_RBs\_BS2']

sorted\_users['Location\_Based\_Allocation'] = sorted\_users.apply(location\_allocation,axis=1)

users\_df['Location\_Based\_Allocation'] = sorted\_users['Location\_Based\_Allocation']

**Κώδικας για την κωδικοποίηση της ποιότητας του σήματος για ευκολότερη επεξεργασία, την κανονικοποίηση των χαρακτηριστικών και την εφαρμογή του PCA για μείωση της διαστατικότητας:**

lab = LabelEncoder()

qualities = pd.concat([users\_for\_cluster['Signal\_Quality\_BS1'],users\_for\_cluster['Signal\_Quality\_BS2']]).unique()

lab.fit(qualities)

encoded = lab.transform(users\_for\_cluster['Signal\_Quality\_BS1'])

users\_for\_cluster['Signal\_Quality\_BS1'] = encoded

users\_df['Signal\_Quality\_BS1\_encoded'] =encoded

encoded2 = lab.transform(users\_for\_cluster['Signal\_Quality\_BS2'])

users\_for\_cluster['Signal\_Quality\_BS2'] = encoded2

users\_df['Signal\_Quality\_BS2\_encoded'] =encoded2

sc = MinMaxScaler(feature\_range=(0,1))

scaled = sc.fit\_transform(users\_for\_cluster)

users\_for\_cluster\_scaled = pd.DataFrame(scaled, columns=users\_for\_cluster.columns)

pca=PCA(n\_components=2)

pca\_data = pca.fit\_transform(users\_for\_cluster\_scaled)

pca\_dataframe = pd.DataFrame(pca\_data,columns=['Pca1','Pca2'])

**Το dataframe users\_for\_cluster έχει ως χαρακτηριστικά τα χαρακτηριστικά των χρηστών που είναι αναγκαία για να γίνει η συσταδοποίηση. Αυτά είναι η απαίτηση των χρηστών, η μέγιστη αποδεκτή καθυστέρηση σύμφωνα με την υπηρεσία, κοντινότερη κεραία, SNR στην πρώτη κεραία, SNR στην δεύτερη κεραία, ποιότητα σήματος στην πρώτη κεραία, ποιότητα σήματος στην δεύτερη κεραία.**

**Κώδικας για τους αλγορίθμους clustering και τη βελτιστοποίηση:**

**KMeans Clustering :**

pos\_K = list(range(3,9,1))

wcss= list()

silhouettes = list()

best\_sil\_KMeans = 0

best\_K = 2

for clusters in pos\_K:

    kmeans = KMeans(n\_clusters=clusters,init='k-means++',random\_state=42,algorithm='lloyd')

    kmeans.fit(pca\_dataframe)

    wcss.append(kmeans.inertia\_)

elbow\_K = 4

next\_range = range(elbow\_K-1,elbow\_K+4,1)

for cluster in next\_range:

    kmeans = KMeans(n\_clusters=cluster, init='k-means++', random\_state=42,algorithm='lloyd')

    kmeans.fit(pca\_dataframe)

    silhouettes.append(silhouette\_score(pca\_dataframe, kmeans.labels\_))

    if silhouette\_score(pca\_dataframe,kmeans.labels\_)> best\_sil\_KMeans:

        best\_sil\_KMeans = silhouette\_score(pca\_dataframe,kmeans.labels\_)

        best\_K = cluster

kmeans\_final = KMeans(n\_clusters=best\_K,init='k-means++',random\_state=42,algorithm='lloyd')

kmeans\_final.fit(pca\_dataframe)

users\_df['KMeans\_Labels']=kmeans\_final.labels\_

**DBSCAN:**

max\_sil = 0

best\_eps = None

for eps in np.arange(0.15,1.05,0.05):

    db= DBSCAN(eps=eps)

    db.fit(pca\_dataframe)

    if len(set(db.labels\_)) > 1 and np.any(db.labels\_ != -1):

        if silhouette\_score(pca\_dataframe, db.labels\_) > max\_sil:

            max\_sil = silhouette\_score(pca\_dataframe, db.labels\_)

            best\_eps = eps

    else :

        print('Only 1 cluster detected')

dbscan = DBSCAN(eps=best\_eps)

dbscan.fit(pca\_dataframe)

users\_df['DBSCAN\_Labels']=dbscan.labels\_

**Hierarchical Clustering:**

tree = linkage(pca\_dataframe,method='ward')

dendrogram(tree, leaf\_rotation=90)

best\_hierarchical\_cl = 2

best\_hierarchical\_sil = 0

hierarchical\_clusters = list(range(3,9,1))

for cl in hierarchical\_clusters:

Hierarchical = AgglomerativeClustering(n\_clusters=cl)

Hierarchical.fit(pca\_dataframe)

if silhouette\_score(pca\_dataframe,Hierarchical.labels\_) > best\_hierarchical\_sil:

best\_hierarchical\_sil = silhouette\_score(pca\_dataframe,Hierarchical.labels\_)

best\_hierarchical\_cl = cl

best\_hierarchical = AgglomerativeClustering(n\_clusters=best\_hierarchical\_cl)

best\_hierarchical.fit(pca\_dataframe)

users\_df['Hierarchical\_Labels']=best\_hierarchical.labels\_

**Gaussian Mixture Model :**

bic\_scores = list()

aic\_scores = list()

for gmm\_comp in pos\_K:

    gmm = GaussianMixture(n\_components=gmm\_comp)

    gmm.fit(pca\_dataframe)

    bic\_scores.append(gmm.bic(pca\_dataframe))

    aic\_scores.append(gmm.aic(pca\_dataframe))

gmm\_best = GaussianMixture(n\_components=5)

gmm\_best.fit(pca\_dataframe)

users\_df['GMM\_Labels']=gmm\_best.predict(pca\_dataframe)

**Spectral Clustering:**

spectral\_sil = list()

for spec\_clusters in pos\_K:

    spec = SpectralClustering(n\_clusters=spec\_clusters)

    spec.fit(pca\_dataframe)

    spectral\_sil.append(silhouette\_score(pca\_dataframe,spec.labels\_))

spec\_best = SpectralClustering(n\_clusters=4)

spec\_best.fit(pca\_dataframe)

users\_df['Spectral\_Labels']=spec\_best.labels\_

**Αλγόριθμος για την δυναμική ανάθεση RBs στους χρήστες ανάλογα με τα συμπεράσματα του KMeans Clustering, τις υπηρεσίες, τα SNR και τις μέγιστες αποδεκτές καθυστερήσεις τους :**

users\_df['Allocated\_RBs'] = np.nan

users\_df['Allocated\_RBs\_Down'] = np.nan

users\_df['Allocated\_Base\_Station']= np.nan

users\_df.drop(['DBSCAN\_Labels','Hierarchical\_Labels','GMM\_Labels','Spectral\_Labels'],axis=1,inplace=True)

grouped\_users\_by\_cluster = users\_df.groupby('KMeans\_Labels')[['SNR\_BS1','SNR\_BS2','Max\_Latency\_In\_ms']].sum()

grouped\_users\_by\_cluster['SNR\_Abs\_Difference'] = abs(grouped\_users\_by\_cluster['SNR\_BS1']- grouped\_users\_by\_cluster['SNR\_BS2'])

clusters\_ordered\_snr = np.argsort(grouped\_users\_by\_cluster['SNR\_Abs\_Difference'])[::-1]

grouped\_mean\_latency = users\_df.groupby('KMeans\_Labels')['Max\_Latency\_In\_ms'].mean().sort\_values(ascending=False)

userslist\_latency = users\_df[(users\_df['KMeans\_Labels']==grouped\_mean\_latency.index[0]) | (users\_df['KMeans\_Labels']==grouped\_mean\_latency.index[1])].copy()

def give\_rbs(row):

    if row['Closest\_BS']==1:

        return row['Requirement\_In\_RBs\_BS1']

    else :

        return row['Requirement\_In\_RBs\_BS2']

def give\_rbs\_down(row):

    if row['Closest\_BS']==1:

        return row['Requirement\_In\_RBs\_BS1\_Down']

    else :

        return row['Requirement\_In\_RBs\_BS2\_Down']

userslist\_latency['Allocated\_RBs'] = userslist\_latency.apply(give\_rbs,axis=1)

userslist\_latency['Allocated\_RBs\_Down'] = userslist\_latency.apply(give\_rbs\_down,axis=1)

for indlat,userlat in userslist\_latency.iterrows():

    if userlat['Closest\_BS']==1:

        base\_station\_df.loc[0,'Resource\_Blocks\_Available'] -= userlat['Allocated\_RBs']

    else :

        base\_station\_df.loc[1,'Resource\_Blocks\_Available'] -= userlat['Allocated\_RBs']

userslist\_latency['Starting\_RBs'] = userslist\_latency['Allocated\_RBs']

already\_allocated = [uu['User\_ID'] for ii,uu in userslist\_latency.iterrows()]

users\_df.loc[already\_allocated,'Allocated\_RBs'] = userslist\_latency.loc[:,'Allocated\_RBs']

taken\_count\_bs1 = 0

taken\_count\_bs2 = 0

for clusternum in clusters\_ordered\_snr:

    valueslist = list()

    indexlist = list()

    bs\_allocated = list()

  userslist = users\_df[users\_df['KMeans\_Labels']==clusternum].sort\_values(by='Max\_Latency\_In\_ms')

    for row\_index,single\_user in userslist.iterrows():

        closest = single\_user['Closest\_BS']

        if single\_user['User\_ID'] in already\_allocated:

            valueslist.append(users\_df.loc[single\_user['User\_ID'],'Allocated\_RBs'])

            indexlist.append(single\_user['User\_ID'])

            bs\_allocated = closest

            continue

        if closest == 1:

            required\_rbs = single\_user['Requirement\_In\_RBs\_BS1']

            if ((base\_station\_df.loc[0,'Resource\_Blocks\_Available'] - single\_user['Requirement\_In\_RBs\_BS1'])>=0):

                base\_station\_df.loc[0,'Resource\_Blocks\_Available'] -= single\_user['Requirement\_In\_RBs\_BS1']

                valueslist.append(single\_user['Requirement\_In\_RBs\_BS1'])

                indexlist.append(single\_user['User\_ID'])

                bs\_allocated.append(1)

            elif (base\_station\_df.loc[1,'Resource\_Blocks\_Available'] - single\_user['Requirement\_In\_RBs\_BS2']>=0) and (single\_user['Requirement\_In\_RBs\_BS2']>=0):

                base\_station\_df.loc[1,'Resource\_Blocks\_Available'] -= single\_user['Requirement\_In\_RBs\_BS2']

                valueslist.append(single\_user['Requirement\_In\_RBs\_BS2'])

                indexlist.append(single\_user['User\_ID'])

                bs\_allocated.append(2)

            else:

                required\_rbs = single\_user['Requirement\_In\_RBs\_BS1\_Down']

                sum\_to\_give\_1 = 0

                taken\_count\_bs1 +=1

                difference=0

                for ind2, user\_that\_gives in userslist\_latency.iterrows():

                    if (np.isnan(user\_that\_gives['Allocated\_RBs'])==False) and (user\_that\_gives['Allocated\_RBs']>user\_that\_gives['Allocated\_RBs\_Down']) and (user\_that\_gives['Service\_Type']!='VoIP') :

                        rounded = math.ceil(user\_that\_gives['Allocated\_RBs'] - user\_that\_gives['Allocated\_RBs\_Down'])

                        userslist\_latency.loc[ind2,'Allocated\_RBs'] -= rounded

                        users\_df.loc[user\_that\_gives['User\_ID'],'Allocated\_RBs'] -= rounded

                        sum\_to\_give\_1 +=rounded

                        if sum\_to\_give\_1 > required\_rbs:

                            difference = sum\_to\_give\_1 - required\_rbs

                            userslist\_latency.loc[ind2,'Allocated\_RBs'] += difference

                            users\_df.loc[user\_that\_gives['User\_ID'],'Allocated\_RBs'] += difference

                            break

                valueslist.append(math.ceil(sum\_to\_give\_1-difference))

                indexlist.append(single\_user['User\_ID'])

                bs\_allocated.append(1)

        else  :

            required\_rbs = single\_user['Requirement\_In\_RBs\_BS2']

            if ((base\_station\_df.loc[1,'Resource\_Blocks\_Available'] - single\_user['Requirement\_In\_RBs\_BS2'])>=0):

                base\_station\_df.loc[1,'Resource\_Blocks\_Available'] -= single\_user['Requirement\_In\_RBs\_BS2']

                valueslist.append(single\_user['Requirement\_In\_RBs\_BS2'])

                indexlist.append(single\_user['User\_ID'])

                bs\_allocated.append(2)

            elif (base\_station\_df.loc[0,'Resource\_Blocks\_Available'] - single\_user['Requirement\_In\_RBs\_BS1']>=0) and (single\_user['Requirement\_In\_RBs\_BS1']>=0):

                base\_station\_df.loc[0,'Resource\_Blocks\_Available'] -= single\_user['Requirement\_In\_RBs\_BS1']

                valueslist.append(single\_user['Requirement\_In\_RBs\_BS1'])

                indexlist.append(single\_user['User\_ID'])

                bs\_allocated.append(1)

            else:

                required\_rbs = single\_user['Requirement\_In\_RBs\_BS2\_Down']

                sum\_to\_give\_2 = 0

                taken\_count\_bs2 +=1

                difference2=0

                for ind3, user\_that\_gives1 in userslist\_latency.iterrows():

                    if (np.isnan(user\_that\_gives1['Allocated\_RBs'])==False) and (user\_that\_gives1['Allocated\_RBs']>user\_that\_gives1['Allocated\_RBs\_Down']) and (user\_that\_gives1['Service\_Type']!='VoIP'):

                        rounded2 = math.ceil(user\_that\_gives1['Allocated\_RBs'] - user\_that\_gives1['Allocated\_RBs\_Down'])

                        userslist\_latency.loc[ind3,'Allocated\_RBs'] -= rounded2

                        users\_df.loc[user\_that\_gives1['User\_ID'],'Allocated\_RBs'] -= rounded2

                        sum\_to\_give\_2  += rounded2

                        if sum\_to\_give\_2 > required\_rbs:

                            difference2 = sum\_to\_give\_2 - required\_rbs

                            userslist\_latency.loc[ind3,'Allocated\_RBs'] += difference2

                            users\_df.loc[user\_that\_gives1['User\_ID'],'Allocated\_RBs'] += difference2

                            break

                valueslist.append(math.ceil(sum\_to\_give\_2-difference2))

                indexlist.append(single\_user['User\_ID'])

                bs\_allocated.append(2)

    users\_df.loc[indexlist,'Allocated\_RBs'] = valueslist

    users\_df.loc[indexlist,'Allocated\_Base\_Station'] = bs\_allocated

**Τα score calinsky-harabasz και silhouette υπολογίστηκαν με τις αντίστοιχες (έτοιμες) συναρτήσεις calinsky\_harabasz\_score() και silhouette\_score() .**

Σύντομο Βιογραφικό Συγγραφέα



Nikolaos P. Prodromos was born in the city of Patras in 2001 and graduated(with honours) in 2019 from the 13th Lyceum of Patras. He is currently a student (on the 4th year) at the Computer Engineering and Informatics Department of the University of Patras.His Research Interests include 5th generation mobile telecommunications networks, resource allocation and machine learning. He also took part (as a conferrer) in the 3rd Panellenic Conference «Education through games and art in education and culture»(Athens, 5-6 November 2022). The topic of his paper was : «Digital technologies as a tool and learning enviroment in today’s school : The Podcasting case». A paper that was requested for publishing, by the Journal of Modern Education Review(NY,USA) (ISSN 2155-7993).

1. <https://www.samsung.com/global/business/networks/insights/podcasts/0110-what-is-mmwave-technology/> [↑](#footnote-ref-1)
2. <https://techcommunity.microsoft.com/t5/azure-for-operators-blog/what-is-the-5g-access-and-mobility-management-function-amf/ba-p/3707685> [↑](#footnote-ref-2)
3. <https://inseego.com/resources/5g-glossary/what-is-embb/> [↑](#footnote-ref-3)
4. <https://inseego.com/resources/5g-glossary/what-is-urllc/> [↑](#footnote-ref-4)
5. <https://inseego.com/resources/5g-glossary/what-is-mmtc/> [↑](#footnote-ref-5)
6. <https://builtin.com/data-science/step-step-explanation-principal-component-analysis> [↑](#footnote-ref-6)
7. <https://obkio.com/blog/what-is-latency/> [↑](#footnote-ref-7)
8. <https://en.wikipedia.org/wiki/K-means_clustering> [↑](#footnote-ref-8)
9. <https://en.wikipedia.org/wiki/Complete-linkage_clustering> [↑](#footnote-ref-9)
10. <https://en.wikipedia.org/wiki/DBSCAN> [↑](#footnote-ref-10)
11. <https://scikit-learn.org/stable/modules/mixture.html> [↑](#footnote-ref-11)
12. <https://datascientest.com/en/spectral-clustering-definition-operation-use> [↑](#footnote-ref-12)
13. <https://www.geeksforgeeks.org/elbow-method-for-optimal-value-of-k-in-kmeans/> [↑](#footnote-ref-13)
14. <https://www.geeksforgeeks.org/calinski-harabasz-index-cluster-validity-indices-set-3/> [↑](#footnote-ref-14)