



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΜΗ ΕΠΟΠΤΕΥΟΜΕΝΗ ΑΝΑΓΝΩΡΙΣΗ ΦΟΡΤΙΩΝ

Διπλωματική Εργασία

ΚΟΥΓΙΟΥΜΗΣ ΓΕΩΡΓΙΟΣ

Επιβλέπων : Μπαργιώτας Δημήτριος

Βόλος 2019



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΜΗ ΕΠΟΠΤΕΥΟΜΕΝΗ ΑΝΑΓΝΩΡΙΣΗ ΦΟΡΤΙΩΝ

Διπλωματική Εργασία

ΚΟΥΓΙΟΥΜΗΣ ΓΕΩΡΓΙΟΣ

Επιβλέπων : Μπαργιώτας Δημήτριος

Βόλος 2019



UNIVERSITY OF THESSALY

SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

UNSUPERVISED LOAD RECOGNITION

Diploma Thesis

KOUGIOUMIS GEORGIOS

Supervisor: Bargiotas Dimitrios

Volos 2019

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να εκφράσω την αμέριστη ευγνωμοσύνη μου σε όλους όσους συνεισέφεραν στην περάτωση αυτής της Διπλωματικής Εργασίας προσφέροντας την βοήθεια τους αλλά και σε όλους τους διδάσκοντες του Τμήματος που καλλιέργησαν το γνωστικό υπόβαθρο για να προκύψει αυτό το αποτέλεσμα κατά την διάρκεια της φοίτησης μου. Ιδιαίτερη μνεία αξίζει στον επιβλέποντα κ. Μπαργιώτα Δημήτριο, που από την πρώτη στιγμή στάθηκε αρωγός και ενίσχυσε την προσπάθεια μου για υλοποίηση του αντικειμένου της παρούσας εργασίας, δείχνοντας πρακτικό ενδιαφέρον καθοδηγώντας το κάθε βήμα. Επιπλέον ένας από τους αφανείς ήρωες που βοήθησαν στην διεκπεραίωση αυτής της διπλωματικής εργασίας είναι ο συμβασιούχος διδάσκων κ. Μπουχουράς Άγγελος, ο οποίος προσέφερε απίστευτη βοήθεια με τις γνώσεις του επί του αντικειμένου, πολύ από τον προσωπικό του χρόνο και μια τεράστια βάση δεδομένων, με πραγματικές μετρήσεις, για την υλοποίηση πειραμάτων. Ακόμα, θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου για την υποστήριξη τους όλα τα χρόνια των σπουδών μου.

**ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ ΠΕΡΙ ΑΚΑΔΗΜΑΪΚΗΣ ΔΕΟΝΤΟΛΟΓΙΑΣ ΚΑΙ ΠΝΕΥΜΑΤΙΚΩΝ
ΔΙΚΑΙΩΜΑΤΩΝ**

«Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ρητά ότι η παρούσα διπλωματική εργασία, καθώς και τα ηλεκτρονικά αρχεία και πηγαίοι κώδικες που αναπτύχθηκαν ή τροποποιήθηκαν στα πλαίσια αυτής της εργασίας, αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο, αρχεία ή/και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή. Αναλαμβάνω πλήρως, ατομικά και προσωπικά, όλες τις νομικές και διοικητικές συνέπειες που δύναται να προκύψουν στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής».

Ο Δηλών

Κουγιούμης Γεώργιος

7 Οκτωβρίου 2019

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια η αναβάθμιση των υποδομών στο δίκτυο της ηλεκτρικής ενέργειας, με την είσοδο σε αυτό τεχνολογιών πληροφορικής και τηλεπικοινωνιών έχει οδηγήσει σε ραγδαία εισχώρηση της πληροφορικής στον κλάδο της ενέργειας. Όλα αυτά συμβαίνουν με σκοπό την μετατροπή του τελικού χρηστή από ένα παθητικό καταναλωτή ηλεκτρικής ενέργειας σε ενεργό, για να συμβεί αυτό θα πρέπει ο καταναλωτής να έχει πλήρη γνώση του προφίλ κατανάλωσης του. Με τον τρόπο αυτό είναι εφικτή η προσαρμογή των αναγκών του χρηστή έτσι ώστε η κατανάλωση ηλεκτρικής ενέργειας να γίνεται με τρόπο τόσο οικονομικότερο αλλά και αποδοτικότερο. Βασική προϋπόθεση για όλα τα παραπάνω είναι η καταγραφή όλων των ενεργών συσκευών του χρηστή και η απεικόνιση αυτών. Τα τελευταία χρόνια οι έξυπνες συσκευές έχουν κερδίσει μεγάλο κομμάτι της αγοράς, δεδομένο ότι παρέχουν τις παραπάνω δυνατότητες, όμως η αντικατάσταση όλων των συσκευών με «έξυπνες» αποδεικνύεται ιδιαίτερα δαπανηρή. Έτσι τα τελευταία χρόνια άρχισαν να εμφανίζονται στη βιβλιογραφία τεχνικές που απαιτούν την εγκατάσταση ενός και μόνο μετρητή. Αυτό σε συνδυασμό με την διαφαινόμενη αντικατάσταση των παραδοσιακών αναλογικών μετρητών της ΔΕΗ με νέους ψηφιακούς, δημιουργεί ένα πολύ προσοδοφόρο ερευνητικό τομέα στη χώρα και όχι μόνο.

ABSTRACT

During the past years the power grid's infostructure have passed through shiver upgrades, with the introduction of technologies like telecommunications and computers that have ultimately led to a computerized power grid. All this occur in order to reshape the final user's behavior and convert him from a passive energy consumer to into an active one. In order for this to happen the consumer should have an all-around information about his consumption profile. In this way the adjustment of the user's needs in energy is possible, this way the energy consumption occurs in more efficient and less "expensive" ways. Vital condition to all this, is the mapping of all active devices and the representation of them. Recently smart devices have earned a large part of the market as they provide the above capabilities, though the replacement of all the residential devices with "smart" ones, is proved to be notably expensive. Thus, during the recent years mapping techniques that require the installation of one smart meter have start to appear. This combined with the imminent replacement, by the power provider company, of all the old analog meters to new digital smart meters have created a lucrative research field in the country and beyond.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ	vi
ABSTRACT	vii
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	viii
ΚΕΦΑΛΑΙΟ 1	10
ΕΙΣΑΓΩΓΗ	10
1.1 Αντικείμενο της Διπλωματικής Εργασίας	10
1.2 Δομή της Διπλωματικής Εργασίας	11
ΚΕΦΑΛΑΙΟ 2	12
ΘΕΩΡΗΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ	12
2.1 Παρακολούθηση φορτίου συσκευής (ALM)	12
2.2 Μη-Παρεμβατική Αναγνώριση Φορτίου (NILM)	13
2.3 Υπογραφές φορτίων (LS)	24
2.4 Αρμόνικες συχνότητες	26
2.5 Επιστήμη των υπολογιστών και μη εποπτευομένη μάθηση.	31
2.6 Συσταδοποίηση.	33
2.6.1 Εισαγωγή	33
2.6.2 Ο αλγόριθμος DBSCAN	35
2.7 Η προσέγγιση της παρούσας εργασίας	39
ΚΕΦΑΛΑΙΟ 3	40
ΠΡΑΚΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ	40
3.1 Εισαγωγή	40
3.2 Μετρήσεις	40
3.2.1 Ρύθμιση μέτρησης	40
3.2.2 Μετρούμενες συσκευές και συνδυασμοί	42
3.2.3 Μορφή ακατέργαστων δεδομένων	43
3.3 Μεθοδολογίες	45
3.3.1 Ο αλγόριθμος του Hart	45
3.3.2 Αναγνώριση συμβάντων (ανίχνευση ακμών)	46
3.3.3 Μετά – επεξεργασία δεδομένων	49
3.4 Παραδοχές	53
ΑΠΟΤΕΛΕΣΜΑΤΑ	53
4.1 Πίνακες συσχέτισης	53
4.2 Εξαγωγή υπογραφών με χρήση παρεμβατικές μεθόδους	56
4.3 Χρόνος εκτέλεσης	58

4.4 Εξαχθείσες υπογραφές φορτίου	60
4.5 Αποτελεσματικότητα διαφόρων δομών υπογραφής	62
ΚΕΦΑΛΑΙΟ 5	65
ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΣΘΗΚΕΣ	65
ΒΙΒΛΙΟΓΡΑΦΙΑ	66
ΠΑΡΑΡΤΗΜΑ Α	68
ΚΩΔΙΚΑΣ	68
ΠΑΡΑΡΤΗΜΑ Β	72
ΓΡΑΦΙΚΕΣ ΠΑΡΑΣΤΑΣΕΙΣ ΠΙΝΑΚΩΝ ΣΥΣΧΕΤΗΣΗΣ	72
ΠΑΡΑΡΤΗΜΑ Γ	78
ΓΡΑΦΙΚΕΣ ΠΑΡΑΣΤΑΣΕΙΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΑΛΓΟΡΙΘΜΟΥ	78

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

1.1 Αντικείμενο της Διπλωματικής Εργασίας

Στο παρόν κεφάλαιο πραγματοποιείται μια συνοπτική παρουσίαση του αντικειμένου της εργασίας, της δομής και του στόχου που εξυπηρετεί. Ακόμα, γίνεται σύντομη αναφορά στα κεφάλαια που ακολουθούν για πληροφόρηση του αναγνώστη.

Αντικείμενο της Διπλωματικής Εργασίας αποτελεί ο σχεδιασμός και η ανάπτυξη ενός αλγορίθμου μέσω του οποίου θα είναι δυνατή η αναγνώριση ανά πάσα στιγμή των συσκευών οι οποίες καταναλώνουν ενέργεια σε ένα οικιακό σύστημα. Στόχος είναι η απόκτηση με αυτόν τον τρόπο γνώσεων για το προφίλ κατανάλωσης του εκάστοτε νοικοκυριού και μέσω των γνώσεων αυτών που προσφέρονται, η μείωση της κατανάλωσης. Σε ανώτερο επίπεδο, δηλαδή αυτό της εταιρίας παροχής ηλεκτρικής ενέργειας, η γνώση των καταναλωτικών προφίλ θα μπορούσε να βοηθήσει τόσο στη βελτίωση του δικτύου, να δώσει πληροφορίες για ελαττωματικές συσκευές καθώς και να προσφέρει πληροφορίες με σκοπό στη διερεύνηση της συμπεριφοράς λειτουργίας των συσκευών.

Το θεωρητικό πρόβλημα που καλείται η εργασία να επιλύσει στην βασική του μορφή απαντάται στην βιβλιογραφία ως παρακολούθηση φορτίου εφαρμογών (Appliance Load Monitoring). Στο παραπάνω πρόβλημα υπάρχουν δύο διαφορετικές προσεγγίσεις οι οποίες αφορούν κυρίως τον τρόπο με τον οποίο καταλήγουμε στα αποτελέσματα. Για την υλοποίηση του συστήματος χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python και έγιναν πειράματα με διάφορους αλγορίθμους συσταδοποίησης, όλα αυτά πάνω σε βάσεις δεδομένων πραγματικών μετρήσεων από δύο οικίες.

1.2 Δομή της Διπλωματικής Εργασίας

Στο Κεφάλαιο 2 γίνεται μια ανάλυση από την θεωρητική οπτική του προβλήματος που επιλύουμε επεξηγώντας όρους που έχουν να κάνουν με αυτό.

Στο Κεφάλαιο 3 παρουσιάζονται οι μετρήσεις και ο τρόπος με τον οποίο αποκτήθηκαν, οι διάφορες μεθοδολογίες οι οποίες ακολουθήθηκαν για την υλοποίηση του αλγορίθμου καθώς και βήμα-βήμα η προγραμματιστική υλοποίηση μερών του συστήματος επιδεικνύοντας πώς αναγνωρίζουμε ένα φορτίο ανά πασα στιγμή σε ένα νοικοκυριό.

Στο Κεφάλαιο 4 γίνεται παρουσίαση των αποτελεσμάτων, η αξιοπιστία τους και οι τρόποι με τους οποίους μπορούμε να τα χρησιμοποιήσουμε.

Στο Κεφάλαιο 5 αναφέρονται μελλοντικές προοπτικές στη χρήση του αλγορίθμου με την υιοθέτηση από τεχνολογικά εργαλεία καθώς και περαιτέρω ερευνητικές ενέργειες που μπορούν να βελτιώσουν την απόδοσή του.

ΚΕΦΑΛΑΙΟ 2

ΘΕΩΡΗΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

2.1 Παρακολούθηση φορτίου συσκευής (ALM)

Στο κεφάλαιο αυτό εξηγούνται έννοιες που χρησιμοποιούνται στα προβλήματα αναγνώρισης φορτίων, η λογική των υπογραφών φορτίου και τέλος η προτεινόμενη οπτική υλοποίησης με την χρήση των αρμονικών συχνοτήτων.

Ο στόχος ενός μοντέλου παρακολούθησης φορτίου συσκευής - Appliance Load Monitoring (ALM) είναι η αναλυτική αναγνώριση των φορτίων ενός καταναλωτή και η παροχή πληροφοριών για το προφίλ κατανάλωσης ενέργειας. Η ανάλυση αυτή σε συνδυασμό με τη χρήση ενός αυτοματοποιημένου συστήματος ενεργειακής διαχείρισης καθιστά αρκετά πιο εύκολη τη διαδικασία βελτιστοποίησης του προφίλ κατανάλωσης, μετατοπίζοντας μεγάλα φορτία σε χρονικές στιγμές “εκτός φόρτου” του δικτύου με τεράστια οφέλη τόσο για τον καταναλωτή όσο και για τον πάροχο αλλά εν τέλει και για τον διαχειριστή του δικτύου. Η ιδέα αυτή δεν είναι καινούργια όμως για πολλά χρόνια, λόγω του μεγάλου κόστους των απαραίτητων υλικών, είχε βρεθεί στο περιθώριο. Σήμερα, με την τεράστια τεχνολογική εξέλιξη που βιώνουμε σε τεχνολογίες όπως αυτές των αισθητήρων, των συσκευών επικοινωνίας, των δικτύων, την τεχνητή νοημοσύνη αλλά και τη μηχανική μάθηση, η δημιουργία τέτοιων συστημάτων έχει επανέλθει στο προσκήνιο. Τέλος δεν θα έπρεπε να θεωρήσουμε ότι τα συστήματα τέτοιου τύπου βρίσκουν εφαρμογή αποκλειστικά και μόνο σε οικιακό επίπεδο καθώς η δυνατότητα αναγνώρισης σφαλμάτων, που παρέχουν, αποτελεί τεράστιο πλεονέκτημα και για τον βιομηχανικό τομέα, ιδίως αν κάποιος αναλογιστεί ότι μερικά από αυτά τα συστήματα μπορούν να λειτουργήσουν με την προσθήκη ενός και μόνο μετρητή [1].

Υπάρχουν δύο βασικές προσεγγίσεις για τέτοιου είδους συστήματα, η παρεμβατική (Intrusive Load Monitoring – ILM) και η μη-παρεμβατική (Non-Intrusive Load Monitoring – NILM). Στη βιβλιογραφία, η παρεμβατική παρακολούθηση φορτίου συσκευής και η μη-παρεμβατική παρακολούθηση φορτίου συσκευής οι οποίες αναφέρονται εναλλακτικά ως κατανεμημένες

μέθοδοι ανίχνευσης και μέθοδοι ανίχνευσης μεμονωμένων σημείων, αντίστοιχα. Αυτό συμβαίνει καθώς οι προσεγγίσεις ενός παρεμβατικού μοντέλου απαιτούν έναν ή περισσότερους αισθητήρες ανά συσκευή για την εκτέλεση της παρακολούθησης φορτίου συσκευής, ενώ το μη-παρεμβατικό μοντέλο απαιτεί ένα μόνο μετρητή ανά σπίτι ή κτίριο που πρόκειται να παρακολουθηθεί. Παρόλο που η παρεμβατική μέθοδος είναι ακριβέστερη στη μέτρηση της κατανάλωσης ενέργειας σε σχέση με το μη-παρεμβατική, τα πρακτικά μειονεκτήματα τα οποία περιλαμβάνουν υψηλό κόστος, πολλαπλές ρυθμίσεις αισθητήρων καθώς και πολυπλοκότητα εγκατάστασης, ευνοούν τη χρήση του μη-παρεμβατικού μοντέλου, ειδικά στην περίπτωση εφαρμογών μεγάλης κλίμακας. Κατά συνέπεια, υπάρχουσες καθώς και νεοσύστατες εταιρείες, μαζί με την ακαδημαϊκή κοινότητα έχουν επικεντρώσει την προσοχή τους στη βελτίωση των προσεγγίσεων μη-παρεμβατικού χαρακτήρα, προκειμένου να καταστεί μια βιώσιμη λύση για ένα ρεαλιστικό περιβάλλον. Με βάση αυτό, παρακάτω θα αναλύσουμε τη λειτουργία ενός τέτοιου συστήματος σχετικά με τις υπογραφές φορτίων των συσκευών καθώς και τους αλγόριθμους ταυτοποίησης φορτίων που χρησιμοποιούνται στη μη-παρεμβατική αναγνώριση φορτίου για την ανίχνευση καταναλωτικών προφίλ [2].

2.2 Μη-Παρεμβατική Αναγνώριση Φορτίου (NILM)

Σε αυτή την ενότητα παρέχουμε μια εισαγωγή σε ένα γενικό πλαίσιο για το σύστημα μη-παρεμβατικής αναγνώρισης φορτίου όπως φαίνεται στο Σχήμα 1. Η έννοια της μη-παρεμβατικής αναγνώρισης φορτίων δεν είναι νέα, καθώς σχεδόν δύο δεκαετίες πίσω ο Hart [3] πρότεινε μια μέθοδο για την αναγνώριση των ηλεκτρικών φορτίων, εξετάζοντας μόνο τις υποδείξεις ειδικής κατανάλωσης ισχύος της συσκευής μέσα στα συγκεντρωτικά δεδομένα φορτίου.

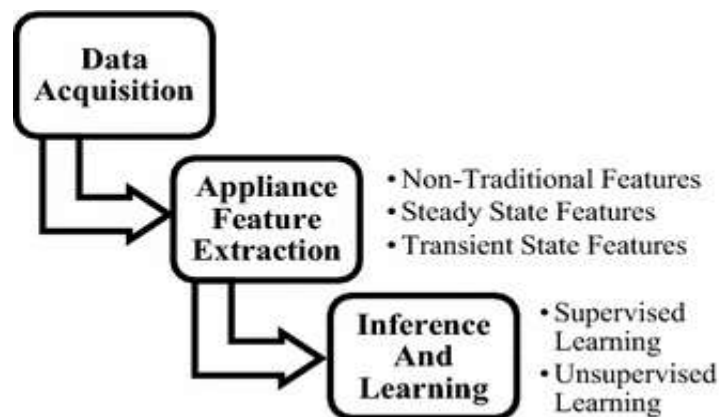
Τα δεδομένα αποκτώνται από τον κύριο ηλεκτρικό πίνακα εκτός του κτιρίου ή της κατοικίας, επομένως θεωρείται ότι δεν είναι ενοχλητικό καθώς η μέθοδος αποφεύγει οποιαδήποτε εγκατάσταση εξοπλισμού εντός της περιουσίας του πελάτη. Ο στόχος είναι να διαχωριστούν τα

στοιχεία ολόκληρης της κατοικίας στα κύρια συστατικά τους. Το πρόβλημα αυτό μπορεί να διατυπωθεί ως εξής:

Τα σήματα ισχύος από το συσσωμάτωμα ενεργών συσκευών στο σημείο εισόδου του μετρητή ως $P(t)$ όπως φαίνεται στο Σχήμα 2, όπου αυτό μπορεί να οριστεί μαθηματικά ως:

$$P(t) = p_1(t) + p_2(t) + \dots + p_n(t) \quad (1)$$

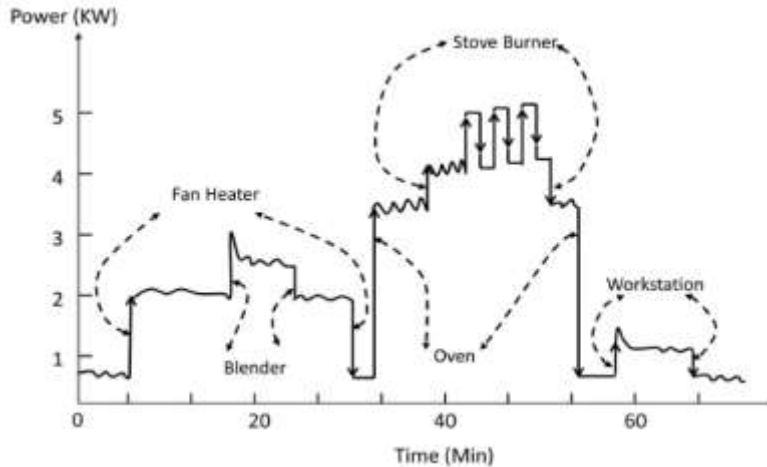
όπου p_i είναι η κατανάλωση ενέργειας των επιμέρους συσκευών που συμβάλλουν στη συνολική μέτρηση και n είναι ο συνολικός αριθμός ενεργών συσκευών εντός της χρονικής περιόδου t . Ο σκοπός της μη-παρεμβατικής αναγνώρισης φορτίων είναι να διεξάγει την αποσύνθεση του $P(t)$ σε σήματα ισχύος, ή και ρεύματος I , εξειδικευμένα ανά συσκευή, προκειμένου να επιτευχθεί ανίχνευση καταναλισκόμενης ενέργειας σε επίπεδο μεμονωμένης συσκευής/φορτίου.



Σχήμα 1 : Γενικό πλαίσιο NILM

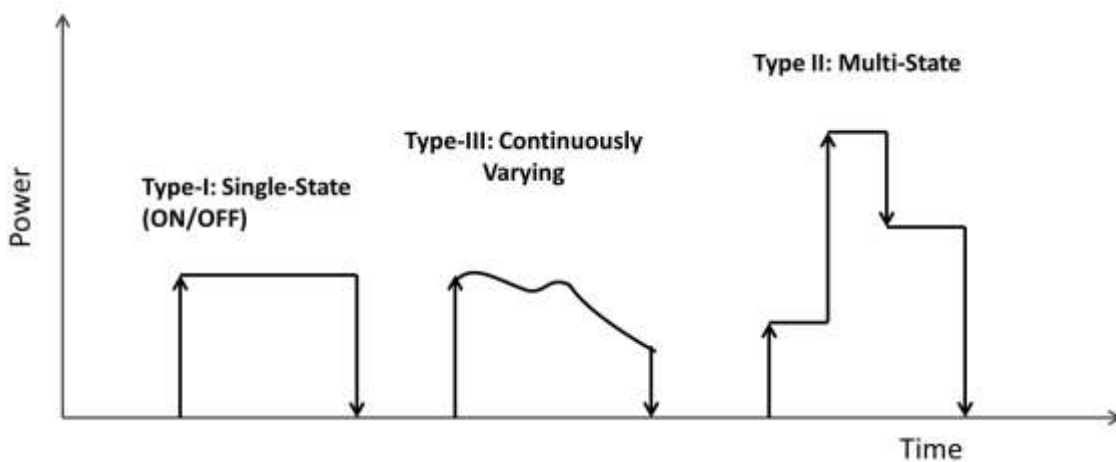
Τα ηλεκτρικά φορτία παρουσιάζουν ένα μοναδικό μοτίβο κατανάλωσης ενέργειας που συχνά ονομάζεται "υπογραφές φορτίου ή συσκευής" – Load Signatures (LS) , που επιτρέπει στους αλγορίθμους διαχωρισμού να διακρίνουν και να αναγνωρίζουν τις λειτουργίες των συσκευών από τις μετρήσεις συνολικού φορτίου. Ο προσδιορισμός της συσκευής εξαρτάται σε μεγάλο βαθμό από τις υπογραφές φορτίου οι οποίες χαρακτηρίζονται περαιτέρω από την κατηγορία της συσκευής [4]. Όπως προτάθηκε, οι καταναλωτικές συσκευές μπορούν να ταξινομηθούν με βάση την κατάσταση λειτουργίας τους ως εξής:

- Τύπος I: Αυτές είναι οι συσκευές με μόνο δύο καταστάσεις λειτουργίας (ενεργοποίηση/απενεργοποίηση). Παραδείγματα τέτοιων συσκευών περιλαμβάνουν λάμπα τραπεζιού, τοστιέρα κ.λπ.
- Τύπος II: Πρόκειται για συσκευές πολλών καταστάσεων, με περιορισμένο όμως αριθμό καταστάσεων λειτουργίας που αναφέρονται επίσης ως μηχανές πεπερασμένων καταστάσεων (FSM). Οι συσκευές που ανήκουν σε αυτή την κατηγορία περιλαμβάνουν πλυντήρια ρούχων, ηλεκτρικές κουζίνες κ.λπ. Ο τρόπος μεταγωγής των συσκευών αυτών είναι επίσης επαναλαμβανόμενος, γεγονός που διευκολύνει τον αλγόριθμο διαχωρισμού να αναγνωρίσει τη λειτουργία τους.
- Τύπος III: Οι συσκευές που ανήκουν στην κατηγορία αυτή είναι επίσης γνωστές ως συνεχώς μεταβλητές συσκευές (CVD) λόγω των χαρακτηριστικών τους για την άντληση μεταβλητής ισχύος χωρίς σταθερό αριθμό καταστάσεων. Το τρυπάνι και οι ηλεκτρονικοί υπολογιστές είναι παραδείγματα CVD χωρίς επαναληψιμότητα στα χαρακτηριστικά τους όσον αφορά την κατανάλωση ενέργειας. Ως εκ τούτου, είναι πολύ δύσκολο για τις μη-παρεμβατικές μεθόδους να διαχωρίσουν αυτόν τον τύπο συσκευής από τις συγκεντρωτικές μετρήσεις φορτίου.
- Τύπος IV: Τέλος έχει επισημανθεί μια επιπλέον κατηγορία συσκευών που παραμένουν ενεργά καθ' όλη τη διάρκεια των εβδομάδων ή των ημερών καταναλώνοντας ενέργεια με σταθερό ρυθμό και, ως εκ τούτου, αναφέρονται ως "υπερπλήρεις καταναλωτικές συσκευές". Συσκευές όπως ο ενσύρματος ανιχνευτής καπνού, οι τηλεφωνικές συσκευές, οι δέκτες καλωδιακής τηλεόρασης συγκαταλέγονται μεταξύ των συσκευών που ανήκουν στην κατηγορία αυτή.



Σχήμα 2 : Πρότυπα κατανάλωσης ισχύος διάφορων συσκευών

Το πρότυπο κατανάλωσης ενέργειας διαφορετικών τύπων φορτίων απεικονίζεται στο Σχήμα 3, το οποίο μεταφράζεται περαιτέρω ως χαρακτηριστικό συσκευής για τη διάκριση μεταξύ διαφορετικών κατηγοριών συσκευών. Η μέχρι σήμερα έρευνα είχε την τάση να επικεντρώνεται στον καθορισμό των υπογραφών φορτίου προσαρμοσμένων στις κατηγορίες συσκευών που αναφέρονται ανωτέρω, ώστε να τις χαρακτηρίζει με τον καλύτερο δυνατό τρόπο για την αναγνώριση. Ωστόσο, πρόσφατα [2] ο συντάκτης υποστήριξε ότι οι συσκευές μπορούν να έχουν ένα μοντέλο πολλαπλής λειτουργίας βασισμένο στην προσαρμογή του χρήστη και τα συλ λειτουργίας, γεγονός το οποίο πρέπει να λαμβάνεται υπόψη στη διαδικασία εξαγωγής χαρακτηριστικών.



Σχήμα 3 : Τα 3 Βασικά πρότυπα κατανάλωσης ισχύος

Ένα μη-παρεμβατικό μοντέλο έχει πέντε διακριτές καταστάσεις οι οποίες επιγραμματικά είναι :

1. Μέτρηση ισχύος ή ρεύματος.
2. Ανίχνευση ακμών
3. Διαχωρισμός συστάδων
4. Δημιουργία Μοντέλων Συσκευής
5. Παρακολούθηση συμπεριφοράς

Όσον αφορά την πρώτη εξ' αυτών, η επιστήμη των ηλεκτρολογικών μετρήσεων αποτελεί ένα πεδίο με τεράστιο ερευνητικό ενδιαφέρον σήμερα. Στην παράγραφο αυτή θα εστιάσουμε στα είδη των μετρητών που παρέχονται για τον σκοπό μετρήσεων κτηριακών φορτίων. Υπάρχει μεγάλη ποικιλία μετρητών ισχύος σχεδιασμένων για τη μέτρηση του συνολικού φορτίου του κτιρίου που μπορούν να ταξινομηθούν ως εξής [2]. :

1. **Μετρητές ενέργειας χαμηλής συχνότητας:** Οι εμπορικές λύσεις που είναι διαθέσιμες στην αγορά σήμερα προσφέρουν μια σειρά συχνοτήτων δειγματοληψίας για τους μετρητές. Ο ρυθμός δειγματοληψίας προσδιορίζει τον τύπο πληροφοριών που μπορούν να εξαχθούν από τα ηλεκτρικά σήματα. Προκειμένου να καταγραφούν οι αρμονικές της ανώτερης τάξης του ηλεκτρικού σήματος, οι οποίες είναι ακέραια πολλαπλάσια βασικής συχνότητας (δηλ. 50 Hz), ο ρυθμός δειγματοληψίας του μετρητή ενέργειας πρέπει να πληροί τα κριτήρια δειγματοληψίας Nyquist-Shannon. Για παράδειγμα, ένας μετρητής ενέργειας όπως ο Itron έχει ρυθμό δειγματοληψίας 600 Hz, πράγμα που του επιτρέπει να συλλαμβάνει μέχρι την 5η αρμονική των ηλεκτρικών σημάτων (δηλαδή 300 Hz). Από την άλλη πλευρά, οι παραδοσιακές μετρήσεις ισχύος, όπως η πραγματική ισχύς, η άεργη ισχύς, η τάση Root Mean Square (RMS) και οι τρέχουσες τιμές μπορούν να υπολογιστούν με χαμηλό ρυθμό δειγματοληψίας (δηλαδή, 120 Hz). Οι καταγεγραμμένες τιμές είτε αναφέρονται στον υποκείμενο διακομιστή μέσω κάρτας διασύνδεσης δικτύου (NIC) είτε υποβάλλονται σε επεξεργασία εντός του

μετρητή. Οι κάρτες διασύνδεσης δικτύου υψηλών απαιτήσεων μπορούν να διαβάσουν, να εγγράφουν και να αναφέρουν δεδομένα έως 1 kHz, απαιτούνται ωστόσο αλλαγές στο υλικό του μετρητή για την υποστήριξη ρυθμού δειγματοληψίας μεγαλύτερου από 5 kHz.

- 2. Μετρητές Ενέργειας Υψηλής Συχνότητας:** Προκειμένου να καταγραφούν τα μεταβατικά γεγονότα ή ο ηλεκτρικός θόρυβος που παράγεται από τα ηλεκτρικά σήματα, οι κυματομορφές πρέπει να λαμβάνουν δειγματοληψία με πολύ υψηλότερη συχνότητα σε μια περιοχή από 10 έως 100 MHz. Αυτοί οι τύποι μετρητών ενέργειας υψηλής συχνότητας είναι συχνά κατασκευασμένοι κατά παραγγελία, ακριβοί λόγω του εξελιγμένου υλικού που απαιτούνε και είναι προσαρμοσμένοι στο είδος των χαρακτηριστικών που χρειάζεται να εξαχθούν από το σήμα.

Οι ερευνητές ισχυρίζονται ότι οι περισσότεροι από τους διαθέσιμους στο εμπόριο μετρητές παρουσιάζουν απόκλιση 10% έως 20% στις μετρήσεις δεδομένων. Επιπλέον, οι λύσεις μέτρησης χαμηλού κόστους προσφέρουν περιορισμένη λειτουργικότητα, καθώς είναι εξοπλισμένες με μετατροπέα αναλογικού σε ψηφιακό σήμα (A/D) χαμηλής ανάλυσης και μνήμη “flash” μικρού μεγέθους με δομές ολοκληρωμένων κυκλωμάτων (chip) που χρησιμοποιείται από τη μονάδα επεξεργασίας για την αποθήκευση αποτελεσμάτων μετά από διάφορες λειτουργίες. Ως εκ τούτου, προκειμένου να επιτευχθεί δειγματοληψία των ηλεκτρικών σημάτων σε μέσο ή υψηλότερο ποσοστό, οι ερευνητές άρχισαν να αναπτύσσουν τα δικά τους πρωτότυπα για πειραματικές αξιολογήσεις. Η απόκτηση δεδομένων για το μη-παρεμβατικό μοντέλο αναγνώρισης φορτίων μπορεί επίσης να ταξινομηθεί σε απόκτηση δεδομένων σε επίπεδο συνόλου και σε απόκτηση δεδομένων από μεμονωμένα κυκλώματα. Το τυπικό σύστημα μη - παρεμβατικής αναγνώρισης χρησιμοποιεί τα δεδομένα που έχουν συγκεντρωθεί από ένα μόνο μετρητή. Ωστόσο, ένας περιορισμός μιας τέτοιας προσέγγισης είναι ο εντοπισμός των συσκευών χαμηλής ισχύος καθώς και των συσκευών μεταβλητών καταστάσεων, υπό την παρουσία φορτίων υψηλής ισχύος, από τα δεδομένα του συνόλου της εγκατάστασης. Έχει προταθεί εναλλακτική προσέγγιση [5] για τη χρήση των μετρήσεων ισχύος

σε επίπεδο κυκλώματος, καθώς συχνά οι συσκευές υψηλής ισχύος λαμβάνουν αποκλειστικό κύκλωμα εντός κατοικιών. Το έργο της αποσύνθεσης ισχύος γίνεται πολύ ευκολότερο καθώς υπάρχουν λιγότερες συσκευές σε κάθε κύκλωμα σε αντίθεση με ολόκληρο το σπίτι, αλλά εις βάρος της αυξημένης πολυπλοκότητας και κόστους εγκατάστασης. Επίσης μια άλλη προσέγγιση στο θέμα είναι η χρήση μετρήσεων ρεύματος και αρμονικών αντί για τις αντίστοιχες τιμές ισχύος καθώς όπως αναφέρετε στο [6] οι ιδιόρρυθμες συμπεριφορές των αρμονικών τιμών ρεύματος παρέχουν επιπρόσθετη πληροφορία για το δείγμα.

Το επόμενο βήμα μετά την απόκτηση δεδομένων είναι η επεξεργασία των ακατέργαστων δεδομένων (δηλ. στιγμιαίες τιμές τάσης και κυματομορφές ρεύματος) προκειμένου να υπολογιστούν οι μετρήσεις ισχύος (ενεργή, άεργη ισχύς, αρμονικές κτλ.). Στη συνέχεια μετά την επεξεργασία των μη επεξεργασμένων δεδομένων λαμβάνει χώρα η ανίχνευση συμβάντων όπως η μετάβαση καταστάσεων της συσκευής (π.χ. ενεργή "ON" σε ανενεργή "OFF"). Μια μονάδα ανίχνευσης συμβάντος ανιχνεύει τη μετάβαση ενεργών / ανενεργών των συσκευών με ανάλυση των αλλαγών στα επίπεδα ισχύος ή ρεύματος. Βιβλιογραφικά έχουν προταθεί διάφορες μέθοδοι ανίχνευσης συμβάντων, ενώ για να χαρακτηριστούν τα ανιχνευμένα συμβάντα, αναπτύσσονται μέθοδοι ταξινόμησης χαρακτηριστικών σταθερής και μεταβατικής κατάστασης. Οι μέθοδοι σταθερής κατάστασης αναγνωρίζουν συσκευές που βασίζονται σε παραλλαγές στις υπογραφές σταθερής κατάστασής τους, για παράδειγμα, η αλλαγή της ενεργού μέτρησης σταθερής κατάστασης από υψηλή έως χαμηλή τιμή μπορεί να προσδιορίσει εάν η συσκευή είναι ενεργοποιημένη ή απενεργοποιημένη. Οι μεταβατικές μέθοδοι, από την άλλη πλευρά, χρησιμοποιούν μεταβατικές υπογραφές που καθορίζουν με μοναδικό τρόπο τις μεταβολές της κατάστασης της συσκευής, εξαγοντας χαρακτηριστικά όπως το σχήμα, το μέγεθος, τη διάρκεια και τις αρμονικές των παροδικών κυματομορφών. Ωστόσο, οι μεταβατικές υπογραφές μπορούν να εξαχθούν μόνον εάν ο ρυθμός δειγματοληψίας είναι υψηλότερος από 1000 δείγματα ανά δευτερόλεπτο. Πρόσφατα γίνεται συζήτηση σχετικά με τη χρήση μεθόδων σταθερής κατάστασης ή μεταβλητών χαρακτηριστικών για τον προσδιορισμό του φορτίου, καθώς και οι δύο αυτές προσεγγίσεις έχουν τα πλεονεκτήματα και τα μειονεκτήματά τους. Λαμβάνοντας υπόψη το κόστος της λύσης, οι μέθοδοι σταθερής

κατάστασης φαίνεται να είναι μια πιο εφικτή προσέγγιση επειδή απαιτούνε υλικό χαμηλού κόστους. Από την άλλη, οι αλγόριθμοι ανίχνευσης φορτίου μπορούν να ενσωματώσουν παροδικά χαρακτηριστικά για να βελτιώσουν τον διαχωρισμό συσκευών με αλληλεπικαλυπτόμενα χαρακτηριστικά σταθερής κατάστασης, αλλά με μεγάλο κόστος υλικού. Εκτός από τις προσεγγίσεις που βασίζονται σε γεγονότα, έχουν επίσης καταβληθεί ερευνητικές προσπάθειες για να αποφευχθεί εντελώς το βήμα ανίχνευσης συμβάντων, είτε χρησιμοποιώντας ακατέργαστες μετρήσεις ρεύματος και τάσης [6] είτε με ανάλυση των πληροφοριών στο φάσμα συχνοτήτων, προκειμένου να ανιχνευθεί η παρουσία ορισμένων συσκευών κατά τη λειτουργία τους.

Τα εξαγόμενα χαρακτηριστικά της συσκευής αναλύονται περαιτέρω από τους αλγόριθμους αναγνώρισης φορτίου, ώστε να προσδιορίζονται οι συγκεκριμένες καταστάσεις των συσκευών από τη συγκεντρωτική μέτρηση. Οι περισσότερες από τις ερευνητικές εργασίες στη μέθοδο μη-παρεμβατικής αναγνώρισης επικεντρώνονται σε ελεγχόμενες τεχνικές μηχανικής μάθησης που απαιτούν στοιχεία με ετικέτα για την κατάρτιση της τάξης. Οι περισσότερες, μέχρι σήμερα, από τις εποπτευόμενες μεθόδους μηχανικής μάθησης που είναι προσαρμοσμένες για τη αναγνώριση φορτίου είναι είτε προσεγγίσεις βασισμένες στη βελτιστοποίηση του σφάλματος είτε στην αναγνώριση μοτίβων.

Η προσέγγιση βελτιστοποίησης σφάλματος προσπαθεί να ταιριάξει τις μετρήσεις ισχύος που παρατηρούνται $P(t)$ σε έναν πιθανό συνδυασμό σημάτων ισχύος συσκευής (που υπάρχουν ήδη στη βάση δεδομένων) για να μειωθεί το σφάλμα αντιστοίχισης. Ωστόσο, ένα σημαντικό μειονέκτημα είναι ότι η παρουσία αγνώστων φορτίων στο $P(t)$ περιπλέκει το πρόβλημα βελτιστοποίησης, καθώς η μέθοδος επιχειρεί να παράσχει μια λύση βασισμένη στο συνδυασμό γνωστών συσκευών [7]. Ως εκ τούτου, η προσέγγιση αναγνώρισης μοτίβων ήταν μια προτιμώμενη μέθοδος από τους ερευνητές για το έργο της ταυτοποίησης φορτίου. Παρόμοια, με την αντιστοιχία μοτίβων, τα χαρακτηριστικά που εξάγονται αντιστοιχίζονται με ένα σύνολο υπογραφών φορτίου που είναι ήδη διαθέσιμες στη βάση δεδομένων, προκειμένου να προσδιοριστεί ένα συμβάν που σχετίζεται με τη λειτουργία μιας συσκευής.

Η απαίτηση των δεδομένων εκπαίδευσης για τους αλγόριθμους είναι ένα από τα σημαντικότερα εμπόδια στην ευρεία υιοθέτηση των λύσεων μη-παρεμβατικών μεθόδων, όπως περιγράφεται παρακάτω. Πρόσφατα, οι ερευνητές έδειξαν αυξημένο ενδιαφέρον σε μεθόδους χωρίς επίβλεψη για την αποσύνθεση του φορτίου έτσι ώστε να μπορεί να εξλειφθεί η ανάγκη για δεδομένα εκπαίδευσης. Σε αντίθεση με τις περισσότερες από τις προσεγγίσεις αποσύνθεσης φορτίου που βασίζονται στην ανίχνευση γεγονότων για ταξινόμηση, οι μέθοδοι χωρίς επίβλεψη δεν βασίζονται σε συμβάντα. Οι μέθοδοι αυτές, κάνουν χρήση των τεχνικών εκμάθησης χωρίς επίβλεψη και προσπαθούν να αποσυνθέσουν απευθείας τις συνολικές μετρήσεις φορτίου, δίχως να εκτελεστεί οποιοδήποτε είδος ανίχνευσης συμβάντων.

Η εκπαίδευση του συστήματος, ή μια φάση προ-μάθησης είναι συχνά προϋπόθεση για τα μη-παρεμβατικά συστήματα. Οι εποπτευόμενοι (supervised) αλγόριθμοι αποσύνθεσης χρειάζονται επαρκή στοιχεία με ετικέτα για την εκμάθηση των παραμέτρων του μοντέλου, προκειμένου να εκτελέσουν την εργασία της αναγνώρισης συσκευών. Οι μέθοδοι εκπαίδευσης μπορούν περαιτέρω να ταξινομηθούν σε on-line ή off-line. Σε περιπτώσεις on-line εκπαίδευσης, οι ερευνητές χρησιμοποιούν τη χρονική τομή για τον εντοπισμό, σε πραγματικό χρόνο, και την εκμάθηση των χαρακτηριστικών της συσκευής. Ωστόσο, κατά την ανίχνευση των συμβάντων φορτίου η χειροκίνητη επισήμανση των συσκευών είναι χρονοβόρα και περίπλοκη, καθώς παρατηρούνται μόνο οι συνολικές τιμές φορτίου αντί των μετρήσεων μεμονωμένων συσκευών. Για να διευκολυνθεί η on-line διαδικασία μάθησης, προτείνονται αρκετοί μηχανισμοί υποβοηθούμενοι από αισθητήρες [8- 10].

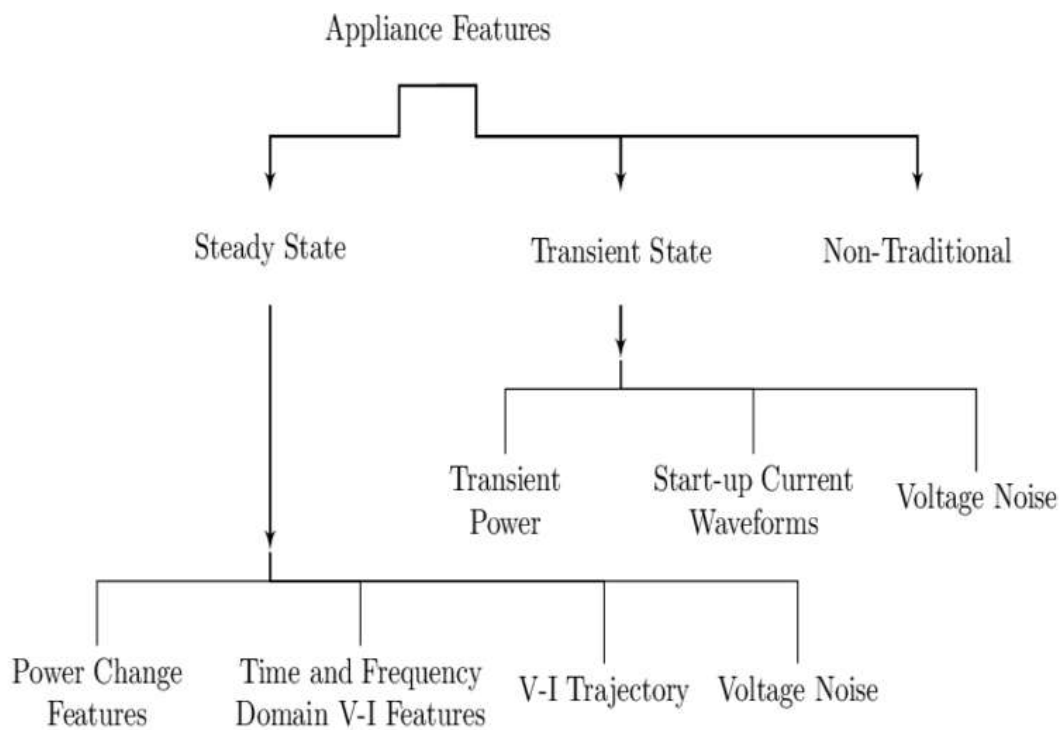
Αντιστρόφως, οι off-line μέθοδοι εκπαίδευσης αποκτούν τις συνολικές μετρήσεις φορτίου από το περιβάλλον-στόχο για μια χρονική περίοδο, όπως για συγκεκριμένες ημέρες, ή οι συσκευές χαρακτηρίζονται με βάση μια προϋπάρχουσα υπογραφή στη βάση δεδομένων. Εναλλακτικά, μια παραδοσιακή μέθοδος εκπαίδευσης, κατά την οποία ο χρήστης αλλάζει χειροκίνητα τις καταστάσεις της συσκευής μία προς μία για να ολοκληρώσει τη βάση

δεδομένων των καταστάσεων της συσκευής μπορεί επίσης να χρησιμοποιηθεί, ωστόσο είναι πολύ χρονοβόρα.

Για να διευκολυνθεί η διαδικασία ετικετοποίησης δεδομένων “labeling”, χρησιμοποιείται επίσης η προσέγγιση υπομέτρησης (sub-metering) που απαιτεί την εγκατάσταση ενός μετρητή ενέργειας ανά συσκευή για την καταγραφή των προτύπων κατανάλωσης της συγκεκριμένης συσκευής. Ωστόσο, δεν προκαλεί μόνο επιπλέον κόστος, αλλά απαιτεί επίσης μια σύνθετη εγκατάσταση αισθητήρων σε κάθε συσκευή ενδιαφέροντος, η οποία δεν είναι εφικτή για την ανάπτυξη ευρείας κλίμακας συστημάτων μη-παρεμβατικής αναγνώρισης φορτίων. Μια σύγχρονη λύση σε αυτό το πρόβλημα, αποτελεί η υλοποίηση μιας πλατφόρμας προσομοίωσης ενός έξυπνου συστήματος μετρήσεων με ενεργειακή επίγνωση, με στόχο την επίσπευση της διαδικασίας σχεδιασμού. Αν και η βασική εστίαση του προσομοιωτή είναι να επιτρέψει στους σχεδιαστές να βελτιστοποιήσουν την αρχιτεκτονική και την επικοινωνία του συστήματος, παρέχει επίσης μια ευκαιρία να προσομοιώσει τη συμπεριφορά των διαφόρων συσκευών μέσα σε ένα σπίτι. Οι ερευνητές που εργάζονται στον τομέα της μη-παρεμβατικής αναγνώρισης φορτίων μπορούν να επωφεληθούν σε μεγάλο βαθμό από αυτή την επιπλέον λειτουργικότητα, καθώς θα μπορούσε να χρησιμοποιηθεί για την ανάλυση των μοτίβων κατανάλωσης ενέργειας των διαφόρων συσκευών και των αλληλεπιδράσεων τους εκ των προτέρων χωρίς την ανάγκη για μία πραγματική εγκατάσταση.

Η όλη διαδικασία δημιουργίας μιας βάσης δεδομένων υπογραφών συσκευής και ενός μοντέλου ετικετοποίησης “labeling” των δεδομένων είναι μια κουραστική διαδικασία και απαιτεί ανθρώπινη παρέμβαση και επίβλεψη, ενώ επί του παρόντος δεν υπάρχουν τυποποιημένες αυτοματοποιημένες λύσεις που να μπορούν να διευκολύνουν τη διαδικασία της μάθησης. Αυτός ήταν ένας από τους περιοριστικούς παράγοντες που εμπόδιζαν την εκτεταμένη επιτυχία των λύσεων μη-παρεμβατικής αναγνώρισης φορτίων. Οι ερευνητές [11] οραματίζονται τη δημιουργία βάσεων δεδομένων γενικών υπογραφών συσκευής, οι οποίες θα είναι ανοιχτής πρόσβασης σε ένα “σύννεφο” για μελλοντικές εφαρμογές, προκειμένου να προχωρήσουν σε ένα ενιαίο πλαίσιο.

Τα χαρακτηριστικά των συσκευών, όπως αναφέρθηκε νωρίτερα, μπορούν να ταξινομηθούν σε γενικές γραμμές σε χαρακτηριστικά σταθερής κατάστασης και μεταβατικής κατάστασης, ωστόσο υπάρχουν και μη παραδοσιακές υπογραφές που χρησιμοποιούνται συχνά σε συνδυασμό με παραδοσιακά χαρακτηριστικά συσκευών για τη βελτίωση της απόδοσης των αλγορίθμων διαχωρισμού φορτίου. Η ανάλυση σταθερής κατάστασης εξετάζει τις σταθερές καταστάσεις λειτουργίας των συσκευών, ενώ η προσωρινή ανάλυση λαμβάνει υπόψη τη μεταβατική κατάσταση κατά την οποία η συμπεριφορά κατανάλωσης ενέργειας των συσκευών είναι ασταθής. Περαιτέρω υποδιαίρεση των δύο αυτών κατηγοριών παρουσιάζεται στο Σχήμα 4.



Σχήμα 4 : Ταξινόμηση των χαρακτηριστικών συσκευών για την ανάλυση ενέργειας

2.3 Υπογραφές φορτίων (LS)

Η αναγνώριση φορτίου – Load Identification (LI) βασίζεται στην ιδέα ότι η συμπεριφορά κατανάλωσης / λειτουργίας κάθε συσκευής θα μπορούσε να περιγραφεί με μια μοναδική και ξεχωριστή δομή δεδομένων, η οποία ενσωματώνει πληροφορίες σχετικά με όλα τα χαρακτηριστικά που σχετίζονται με τις λειτουργικές / ηλεκτρικές ιδιότητες της συσκευής. Αυτή η διατύπωση είναι κοινώς γνωστή ως υπογραφή φορτίου – Load Signature (LS). Η πρόκληση, όμως, στις μεθοδολογίες μη-παρεμβατικής παρακολούθησης (NILM) έγκειται στην αποδοτικότητα της αναγνώρισης φορτίου (LI), καθώς το καταγεγραμμένο και μετρημένο σήμα είναι στην πραγματικότητα ένα συγκεντρωτικό σήμα που ενσωματώνει τις πληροφορίες από όλες ταυτόχρονα τις συσκευές. Ο αποτελεσματικός διαχωρισμός αυτού του σήματος, εξαρτάται σε μεγάλο βαθμό από τη διαμόρφωση μοναδικών και διακριτών υπογραφών φορτίου. Σε περίπτωση ταυτόχρονης λειτουργίας δύο ή περισσότερων συσκευών, η αναγνώριση φορτίου περιλαμβάνει την εξεύρεση κατάλληλου συνδυασμού των επιμέρους υπογραφών φορτίου των ενεργών συσκευών. Επομένως, όσο μεγαλύτερη είναι η μοναδικότητα αυτών των υπογραφών φορτίου, τόσο ευκολότερη είναι η διαδικασία αναγνώρισης [12].

Η αναγνώριση φορτίου (LI) αναφέρεται η ακριβής γνώση της λειτουργικής κατάστασης ή/και του τρόπου λειτουργίας κάθε συσκευής. Η ιδέα πίσω από τις υπογραφές φορτίου (LS) είναι ότι η λειτουργική κατάσταση κάθε συσκευής μπορεί να περιγραφεί με απλό τρόπο χρησιμοποιώντας αντιπροσωπευτικό τύπο δεδομένων. Τα δεδομένα αυτά εμπεριέχουν πληροφορίες σχετικά με τα ηλεκτρικά χαρακτηριστικά της συσκευής και σχετικά με τυχόν ειδικά λειτουργικά χαρακτηριστικά που μπορούν εύκολα να διακρίνουν την αναφερόμενη συσκευή από άλλες εντός της εγκατάστασης. Ως εκ τούτου, ο πυρήνας του συστήματος

αναγνώρισης φορτίου (LI) βασίζεται στην ανάπτυξη κατάλληλων δακτυλικών αποτυπωμάτων με την ονομασία υπογραφές φορτίου (LS). Στην πράξη, οι υπογραφές φορτίου είναι μια κατάλληλα διαμορφωμένη δομή δεδομένων που περιγράφει συνοπτικά τη συμπεριφορά του φορτίου κατά τη λειτουργία του. Οι πρώτες απόπειρες προς την κατεύθυνση αυτή αφορούσαν την εγκατάσταση τουλάχιστον ενός αισθητήρα σε κάθε συσκευή ή ενδιάμεσου συστήματος παρακολούθησης, γνωστού ως μεθοδολογίες ILM. Η προσέγγιση αυτή, αν και πολύ αποδοτική, είναι ακατάλληλη για εφαρμογές μεγάλης κλίμακας λόγω του υψηλού κόστους και του φόρτου υλοποίησης [1]. Ο Hart [3] πρότεινε μια απλούστερη προσέγγιση στις αρχές της δεκαετίας του '90, δηλαδή τη μη-παρεμβατική παρακολούθηση φορτίου (NILM). Η προσέγγιση αυτή βασίστηκε στην εγκατάσταση μόνο μιας διάταξης μέτρησης στον κύριο πίνακα τροφοδοσίας της εγκατάστασης. Αν και το κόστος και η πολυπλοκότητα της εγκατάστασης στην περίπτωση αυτή είναι σημαντικά χαμηλότερα σε σύγκριση με μια προσέγγιση παρεμβατικής αναγνώρισης, η βασική πρόκληση είναι να αποσυντεθούν τα συγκεντρωτικά μετρηθέντα δεδομένα στα σωστά στοιχεία (π.χ. σχήμα ανάλυσης). Αυτή η διαδικασία απαιτεί τη χρήση ειδικών αλγορίθμων, που ονομάζονται αλγόριθμοι NILM. Γενικά, οι αλγόριθμοι NILM ανήκουν είτε σε εποπτευόμενες είτε σε μη εποπτευόμενες κατηγορίες. Το πρώτο χρειάζεται τις γνώσεις των υπογραφών φορτίου των συσκευών για την εφαρμογή του συστήματος διαχωρισμού, ενώ το δεύτερο δεν χρειάζεται δεδομένα μεμονωμένων συσκευών και οι πληροφορίες του μοντέλου καταγράφονται μόνο με το συγκεντρωτικό φορτίο [13]. Λόγω της απαιτούμενης γνώσης για τις υπογραφές φορτίων των συσκευών, οι εποπτευόμενοι (supervised) αλγόριθμοι μη-παρεμβατικού χαρακτήρα έχουν γενικά καλύτερες επιδόσεις από τους μη εποπτευόμενους. Η αποτελεσματικότητα αυτών των αλγορίθμων NILM εξαρτάται από δύο παραμέτρους:

- Τον υπολογιστικό χρόνο που απαιτείται για την εφαρμογή του αλγόριθμου.
- Τις επιδόσεις όσον αφορά την επιτυχημένη ταυτοποίηση της συσκευής.

Αμφότερα συνδέονται άμεσα με τα χαρακτηριστικά και τα ηλεκτρικά χαρακτηριστικά που ενσωματώνονται στις υπογραφές φορτίου. Οι δημοσιευμένες μεθοδολογίες χρησιμοποιούν μια ποικιλία χαρακτηριστικών, όπως η ενεργός και άεργος ισχύς, αρμονικές ρεύματος και

τάσης, παροδική συμπεριφορά ιδίως κατά την ενεργοποίηση ή/και την απενεργοποίηση, τρέχοντα χαρακτηριστικά κυματομορφής, μετρήσεις ρεύματος και τάσης, και τεχνικές βασισμένες σε τροχιά V-I, είτε μεμονωμένα είτε συνδυαστικά, προκειμένου να βελτιωθεί η αποτελεσματικότητα των υπογραφών φορτίου [14].

Για τα συστήματα αποσύνθεσης - αναγνώρισης, οι υπάρχουσες προσεγγίσεις βασίζονται κυρίως στον υπολογισμό βελτιστοποίησης σφάλματος, ειδικά για τους εποπτευόμενους αλγόριθμους NILM. Άλλοι αλγόριθμοι NILM χρησιμοποιούν τεχνικές νευρωνικού δικτύου και τεχνητής νοημοσύνης, όπως προσεγγίσεις μηχανών βαθιάς μάθησης, ενώ τελευταία, οι αλγόριθμοι NILM που βασίζονται σε κρυφά μοντέλα Markov προσέλκυσαν μεγάλη προσοχή λόγω της υψηλής απόδοσης ταυτοποίησης και της μειωμένης υπολογιστικής επιβάρυνσης. Στις μεθοδολογίες αυτές, οι επιδόσεις των συσκευών χαμηλής κατανάλωσης ισχύος είναι υψηλές και το σύστημα αποσύνθεσης είναι εφαρμόσιμο ακόμη και σε ένα περιβάλλον με θόρυβο, όμως το απαραίτητο μέρος της εκπαίδευσης αυξάνει την πολυπλοκότητα του αλγορίθμου.

2.4 Αρμόνικες συχνότητες

Τα συστήματα παραγωγής της ηλεκτρικής ενέργειας είναι τριφασικά συστήματα. Η παραγόμενη AC τάση στις σύγχρονες γεννήτριες είναι τριφασική ημιτονοειδής εναλλασσόμενη τάση και με την ίδια μορφή, αλλά με διαφορετικά επίπεδα τάσεων, μεταφέρεται μέχρι τα σημεία των καταναλωτών της ηλεκτρικής ενέργειας. Τα ηλεκτρικά φορτία κατατάσσονται σε δύο κατηγορίες:

- Γραμμικά ηλεκτρικά φορτία
- Μη γραμμικά ηλεκτρικά φορτία

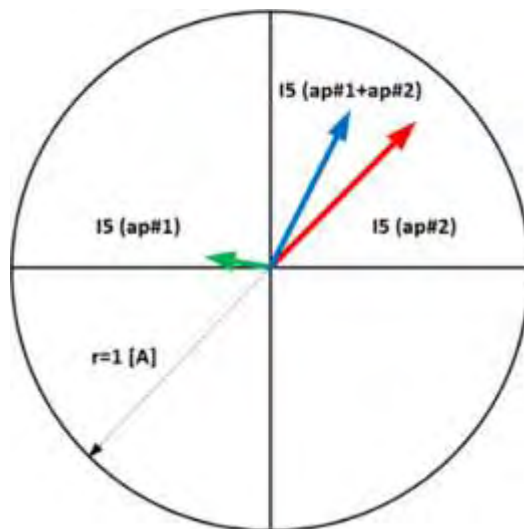
Τα γραμμικά ηλεκτρικά φορτία, όπως είναι οι ωμικές αντιστάσεις, τα πηνία, οι πυκνωτές, οι ηλεκτροκινητήρες AC κτλ., απορροφούν ρεύματα ημιτονοειδώς εναλλασσόμενα, όπως η

μορφή της ηλεκτρικής τάσης. Η πλειοψηφία όμως των ηλεκτρικών φορτίων, στις μέρες μας, λειτουργούν απορροφώντας από το δίκτυο ρεύματα τα οποία δεν είναι ημιτονοειδή, αλλά αποκλίνουν σε μικρότερο ή μεγαλύτερο βαθμό από αυτή την κυματομορφή. Τέτοια φορτία είναι οι ρυθμιστές στροφών AC κινητήρων ή VFDs (Variable Frequency Drives), τα παλμοτροφοδοτικά ή SMPS (Switched Mode Power Supplies), οι ανορθωτές (Rectifiers), οι ηλεκτροσυγκολλήσεις inverter, τα τροφοδοτικά αδιάλειπτης παροχής ή UPS (Uninterruptible Power Supplies), οι λαμπτήρες φθορισμού, οι ρυθμιστές φωτεινότητας λαμπτήρων (Dimmers), τα κλιματιστικά inverter, οι ηλεκτρονικοί υπολογιστές, οι εκτυπωτές και πλήθος ακόμα ηλεκτρικών/ηλεκτρονικών συσκευών. Γενικά όλοι οι ηλεκτρονικοί μετατροπείς ισχύος AC/DC, DC/AC, DC/DC, AC/AC είναι μη γραμμικά ηλεκτρικά φορτία και τα ρεύματα λειτουργίας τους περιέχουν αρμονικές συνιστώσες.

Η εμφάνιση αρμονικών συνιστωσών ρεύματος κατά τη λειτουργία των ηλεκτρικών φορτίων μπορεί να προκαλέσει σοβαρές δυσλειτουργίες, τόσο στα ηλεκτρικά φορτία, όσο και στα ηλεκτρικά δίκτυα μεταφοράς και διανομής της ηλεκτρικής ενέργειας. Γίνετε συνεπώς σήμερα μια προσπάθεια προσδιορισμού των αρμονικών συνιστωσών, με κατάλληλες μετρήσεις, και στη συνέχεια επιχειρείται ο περιορισμός τους με κατάλληλες διατάξεις. Αυτό, καθώς σε πλήθος εφαρμογών είναι επιθυμητό να μετριάσουμε την ποσότητα των αρμονικών των κυματομορφών και των επιπτώσεων που αυτές έχουν στο δυτικό πράγμα το οποίο είναι εφικτό με χρήση φίλτρων ή με την αύξηση της πολυπλοκότητας των ηλεκτρονικών μετατροπέων ισχύος.

Η ιδέα να χρησιμοποιούνται αρμονικές συχνότητες ρευμάτων για να σχηματίσουν μοναδικές και εύκολα διακριτές υπογραφές φορτίων δεν είναι νέα καθώς με την χρήση των αρμονικών στις υπογραφές φορτίου μπορούμε εύκολα να διαχωρίσουμε τα γραμμικά από τα μη γραμμικά φορτία. Ωστόσο οι προτεινόμενες προσεγγίσεις δεν λαμβάνουν υπόψη τις γωνίες φάσης των αρμονικών συχνοτήτων. Μια υπογραφή φορτίου θα πρέπει ιδανικά να είναι απλή και ταυτόχρονα μοναδική, κάτι που δεν είναι πάντα εφικτό. Το δεύτερο βήμα είναι η διαδικασία της αναγνώρισης φορτίου, όπου ένα συγκεντρωτικό σήμα ως αποτέλεσμα της ταυτόχρονης

λειτουργίας πολλαπλών φορτίων αποσυντίθεται στις μεμονωμένες συνεισφορές κάθε φορτίου. Σύμφωνα με το [6] οι πρόσθετες πληροφορίες που παρέχονται από τις γωνίες φάσης αυξάνουν την αποδοτικότητα των αναπτυγμένων υπογραφών φορτίου σε σύγκριση με την περίπτωση που χρησιμοποιούνται μόνο τα αρμονικά πλάτη. Στην πραγματικότητα για την 3η και 5η αρμονική είναι δυνατόν να αποδοθούν αντίθετα αρμονικά διανύσματα για διαφορετικές συσκευές. Σε αυτή την περίπτωση, η ταυτόχρονη λειτουργία αυτών των δύο συσκευών θα εξαλείψει ορισμένες αρμονικές από το συγκεντρωτικό σήμα, οι οποίες δεν θα αποκαλυφθούν εάν χρησιμοποιούνται μόνο αρμονικά πλάτη. Για παράδειγμα, στο Σχήμα 5, η 5η αρμονική ρεύματος δύο τυχαίων συσκευών παρουσιάζεται σε διανυσματικές μορφές [6].



Σχήμα 5 : Διανυσματικές τιμές 5ης αρμονική ρεύματος για δυο διαφορετικές συσκευές

Με βάση το Σχήμα 5 γίνεται προφανές ότι η συγκεντρωτική 5η αρμονική ρεύματος είναι το αποτέλεσμα της συσσωμάτωσης των αντίστοιχων μεμονωμένων αρμονικών ρευμάτων και όχι της αλγεβρικής άθροισης του πλάτους για κάθε μία από αυτές. Επομένως, η συμβολή των γωνιών της φάσης των αρμονικών ρευμάτων στην ανάπτυξη ισχυρών και διακριτών LS μπορεί να αποδειχθεί αποφασιστική. Σύμφωνα με το [6] τόσο τα πλάτη των αρμονικών τάξεων όσο και τα ορίσματα αυτών υπολογίζονται με την εφαρμογή Fast Fourier Transformation (FFT) στις μετρήσεις στιγμιαίων τιμών ρεύματος. Έτσι, κάθε αρμονικό ρεύμα, που μετράτε με ρυθμό

εγγραφής ίσο με 1 Hz, θεωρείται ως τρέχων διάνυσμα. Για να αναλυθούν οι επιπτώσεις της εφαρμογής των αρμονικών φάσεων στις υπογραφές φορτίου, έχουν δοκιμαστεί τρεις διαφορετικές συνθέσεις υπογραφών [6]. Για κάθε σχηματισμό, η υπογραφή φορτίου της συσκευής i βασίζεται στις πληροφορίες που παρέχονται από τα τρία αρμονικά ρεύματα, όπως εξηγείται νωρίτερα. Έτσι, η γενική μορφή κάθε υπογραφής φορτίου (LS) είναι η εξής [6]:

$$LS_i^X = [(I_{i1}^X), (I_{i3}^X), (I_{i5}^X)] \quad (2)$$

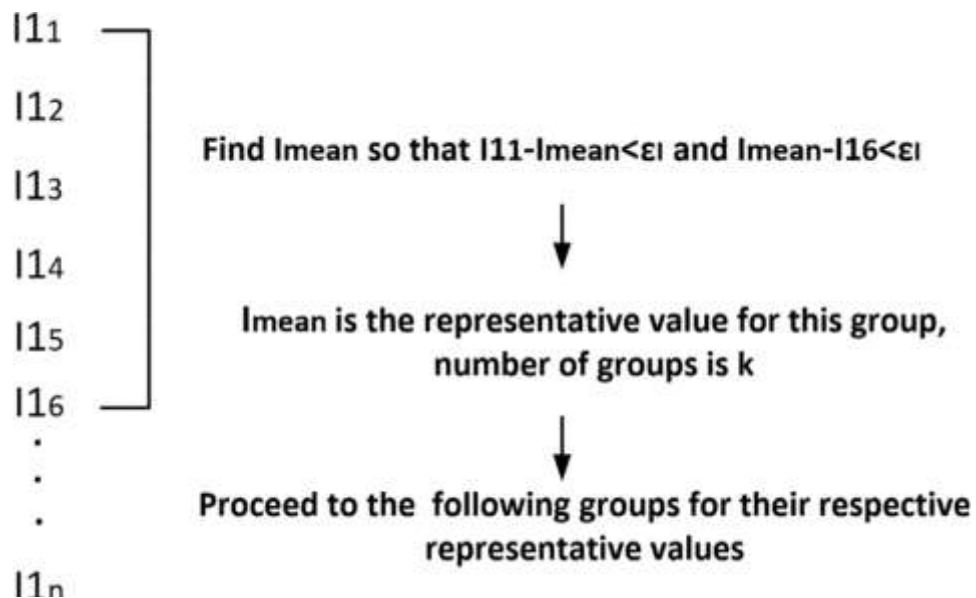
όπου I_1 , I_3 και I_5 είναι τα αρμονικά ρεύματα των πρώτων τριών αρμονικών. Η γενική μορφή στο (2) είναι περαιτέρω εξειδικευμένη με βάση το αν οι αρμονικές των ρευμάτων θεωρούνται ως διανύσματα ή απλά αποτυπώνονται αλγεβρικά ως πλάτη. Μια άλλη κρίσιμη πτυχή σχετικά με τη δημιουργία των υπογραφών φορτίου αφορά τον αριθμό των τιμών για κάθε αρμονικό ρεύμα στο σώμα της υπογραφής. Αυτό έχει μεγάλη σημασία, αφού επηρεάζει την υπολογιστική επιβάρυνση του αλγορίθμου. Ως εκ τούτου, η ανάλυση αυτή προτείνει την εξέταση των όσο το δυνατόν λιγότερων, αλλά ταυτόχρονα πιο αντιπροσωπευτικών τιμών για κάθε αρμονική συχνότητα, υπό την προϋπόθεση ότι αυτές οι τιμές θα μπορούσαν να περιγράψουν αποτελεσματικά τη λειτουργία της συσκευής σε σταθερή κατάσταση. Οι τρεις εξεταζόμενες συνθέσεις υπογραφών είναι οι εξής [6]:

- Ο πρώτος σχηματισμός υπογραφών φορτίου, δηλ. LS^A , βασίζεται στη χρησιμοποίηση μόνο του εύρους των αρμονικών ρευμάτων. Για να καταλήξει με τις αντιπροσωπευτικές τιμές που θα χρησιμοποιηθούν για κάθε αρμονική σειρά, εφαρμόζει τα ακόλουθα βήματα:
 - Τα τρέχοντα πλάτη κάθε αρμονικής σειράς σχηματίζουν μια λίστα με φθίνουσα σειρά.
 - Μετακινούμενοι από πάνω προς τα κάτω σε αυτή τη λίστα, δημιουργούμε ομάδες δεδομένων k, i και m για τις αντίστοιχες αρμονικές συχνότητες, έτσι ώστε η μέση τιμή της ομάδας να μην αποκλίνει περισσότερο από ένα προκαθορισμένο όριο σφάλματος από τις ανώτερες και χαμηλότερες τιμές της ομάδας. Έτσι, για κάθε

αρμονική σειρά, τα k , i και m αντιπροσωπευτικά τρέχοντα πλάτη θα χρησιμοποιηθούν για τη δημιουργία υπογραφών οι οποίες θα εκφράζονται όπως παρουσιάζεται στο (3)

$$LS_i^A = [(I_{i1j}^A, \dots, I_{i1k}^A), (I_{i3j}^A, \dots, I_{i3k}^A), (I_{i5j}^A, \dots, I_{i5k}^A)] \quad (3)$$

Η διαδικασία αυτή παρουσιάζεται επίσης εν συντομία στο Σχήμα 6, όπου οι τρέχουσες τιμές για την 1η αρμονική παρατίθενται κατά φθίνουσα σειρά και σχηματίζονται οι ομάδες k με τις αντιπροσωπευτικές τους τιμές.



Σχήμα 6 : Διαδικασία σχηματισμού υπογραφών φορτίου

- Ο δεύτερος σχηματισμός υπογραφών φορτίου, δηλ. LS^B , χρησιμοποιεί τις πληροφορίες που παρέχονται από τη γωνία φάσης κάθε αρμονικού ρεύματος. Προκειμένου να διατηρηθεί χαμηλή η υπολογιστική επιβάρυνση και η μαθηματική πολυπλοκότητα, αυτή η διατύπωση βασίζεται στις προβολές των διανυσμάτων στην x-άξονα αντί των ίδιων των διανυσμάτων. Πρέπει να αναφερθεί ότι, στην προκειμένη περίπτωση, οι διακυμάνσεις της γωνίας φάσης δεν έχουν ληφθεί υπόψη. Έτσι, για κάθε αρμονικό ρεύμα η γωνία φάσης

θεωρήθηκε σταθερή και ίση με τη μέση τιμή. Το τελευταίο ήταν δικαιολογημένο, καθώς στις περισσότερες περιπτώσεις αυτές οι διακυμάνσεις της γωνίας φάσης ήταν ομαλές και μόνο λίγες τιμές συμπεριφέρονταν ως ακραίες. Οι αντιπροσωπευτικές τρέχουσες αρμονικές και πλάτη υπολογίστηκαν όπως στο πρώτο σχηματισμό υπογραφών.

$$LS_i^B = [(I_{i1j}^B * \cos\varphi_1, \dots, I_{i1k}^B * \cos\varphi_1), (I_{i3j}^B * \cos\varphi_3, \dots, I_{i3k}^B * \cos\varphi_3), (I_{i5j}^B * \cos\varphi_5, \dots, I_{i5k}^B * \cos\varphi_5)] \quad (4)$$

Οπού $\varphi_1, \varphi_3, \varphi_5$ είναι οι γωνίες φάσης των αντίστοιχων αρμονικών και $j = 1, \dots, k$ ή i ή m ανάλογα με την αντίστοιχη αρμονική σειρά.

- Ο τρίτος σχηματισμός υπογραφών φορτίου, δηλ. LS^C , χρησιμοποιεί συνολικά τα αρμονικά ρεύματα με τη μορφή διανυσμάτων για κάθε αρμονική τάξη, εκμεταλλευόμενο έτσι πλήρως τις πληροφορίες που παρέχονται από τις μετρήσεις τόσο του εύρους όσο και της γωνίας φάσης κάθε αρμονικού ρεύματος. Αν και η αποτελεσματικότητα αυτού του σχηματισμού υπογραφών φορτίου αναμένεται να είναι υψηλότερη, το ίδιο ισχύει και για την υπολογιστική επιβάρυνση [6].

$$LS_i^C = [(\vec{I}_{i1j}^C, \dots, \vec{I}_{i1p}^C), (\vec{I}_{i3j}^C, \dots, \vec{I}_{i3r}^C), (\vec{I}_{i5j}^C, \dots, \vec{I}_{i5s}^C)] \quad (5)$$

2.5 Επιστήμη των υπολογιστών και μη εποπτευόμενη μάθηση.

Τα τελευταία χρόνια όλο και μεγαλύτερο ερευνητικό ενδιαφέρον στρέφεται στον τομέα της επιστήμης των υπολογιστών καθώς η τεχνολογική πρόοδος έχει οδηγήσει σε σημαντική αύξηση της επεξεργαστικής ισχύος και κατά συνέπεια στην απόδοση των συστημάτων. Δυο από τα σημαντικότερα παιδιά του κλάδου είναι η μηχανική μάθηση (Machine Learning) και η βαθιά μάθηση (Deep Learning).

Σύμφωνα με τον Άρθουρ Σάμιουελ που έδωσε έναν από τους πρώτους ορισμούς της μηχανικής μάθησης το 1959, σαν μηχανική μάθηση όρισε: “Το πεδίο μελέτης που δίνει στους

υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί". Η μηχανική μάθηση χρησιμοποιείται σε αναζητήσεις στο διαδίκτυο, φίλτρα ανεπιθύμητης αλληλογραφίας, συστήματα ασφάλειας, τοποθέτηση διαφημίσεων on-line, πιστωτικές κάρτες, ανίχνευση απατών, εμπόριο μετοχών, σχεδιασμό φαρμάκων και πολλές άλλες εφαρμογές [15]. Τα μοντέλα μηχανικής μάθησης (ML) έχουν χρησιμοποιηθεί ευρέως και στη μοντελοποίηση, το σχεδιασμό και την πρόβλεψη στα ενεργειακά συστήματα. Κατά τη διάρκεια των τελευταίων δύο δεκαετιών, σημειώθηκε δραματική αύξηση στην πρόοδο και εφαρμογή διαφόρων τύπων μοντέλων μηχανικής μάθησης (ML) για ενεργειακά συστήματα. Στον τομέα τις αναγνώρισης φορτίων ειδικότερα αποτέλεσε μια από τις βασικές αιτίες επανεμφάνισης, μετά από δεκαετίες των τεχνικών μη-παρεμβατικής αναγνώρισης φορτίων (NILM) καθώς η αύξηση σε επεξεργαστική συχνή έχει οδηγήσει τη μηχανική μάθηση (ML) σε πολύ πιο αποδοτικά μοντέλα που μπορούνε πολύ εύκολα να ανταγωνιστούν ένα μοντέλο έξυπνου σπιτιού [17].

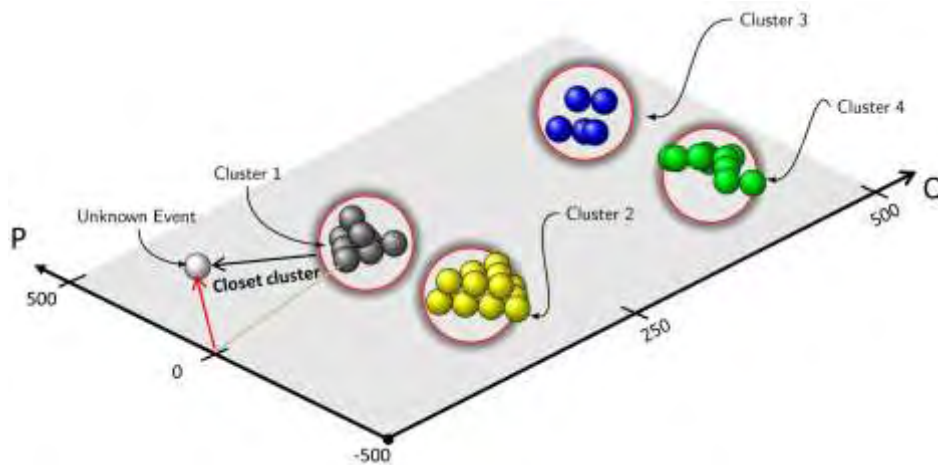
Από την άλλη, αν και βιβλιογραφικά το Deep Learning (DL) θεωρείται ως μια υποκατηγορία της μηχανικής μάθησης (ML) τα τελευταία χρονιά λαμβάνει όλο και περισσότερο προσοχή με αποτέλεσμα σήμερα να θεωρείται ένας από τους πιο αναπτυσσόμενους επιστημονικούς κλάδους. Η αρχιτεκτονικές βαθιάς μάθησης (DL) όπως τα βαθιά νευρωνικά δίκτυα, τα επαναλαμβανόμενα νευρωνικά δίκτυα και τα συνελκτικά νευρωνικά δίκτυα έχουν εφαρμοστεί σε τομείς όπως η όραση των υπολογιστών, η αναγνώριση ομιλίας, η επεξεργασία φυσικής γλώσσας, η αναγνώριση ήχου, το φιλτράρισμα κοινωνικών δικτύων, η αυτοματοποιημένη μετάφραση, η βιοπληροφορική, η ανάλυση ιατρικών εικόνων, τα προγράμματα επιθεώρησης υλικών και επιτραπέζιων παιχνιδιών, όπου παρήγαγαν αποτελέσματα συγκρίσιμα και, σε ορισμένες περιπτώσεις, ακόμα και ανώτερα από αυτά του αντίστοιχου ανθρώπινου δυναμικού [16].

Όσον αφορά την αναγνώριση φορτίων οι παραπάνω τεχνικές φαίνεται ότι δεν αφήνουν αδιάφορο τον τομέα αυτό καθώς είναι πολλά υποσχόμενες και μπορεί να οδηγήσουν σε μη εποπτευόμενες εφαρμογές με αρκετά ανταγωνιστικό προφίλ, σε σχέση με τις αντίστοιχες εποπτευόμενες εφαρμογές, παρέχοντας υψηλότερη απόδοση και χαμηλότερο κόστος.

2.6 Συσταδοποίηση.

2.6.1 Εισαγωγή

Η εξόρυξη δεδομένων, αποτελεί επίσης ένα παρακλάδι της μηχανικής μάθησης, είναι η διαδικασία εξαγωγής κρυφών και ενδιαφερόντων μοτίβων ή χαρακτηριστικών από πολύ μεγάλα σύνολα δεδομένων με σκοπό τη χρήση αυτών στη λήψη αποφάσεων και στην πρόβλεψη μελλοντικής συμπεριφοράς. Μία από αυτές τις μεθόδους εξαγωγής πληροφοριών είναι η συσταδοποίηση όπου ένα σύνολο αντικειμένων χωρίζεται σε πολλές συστάδες με σκοπό τη μεγιστοποίηση της ομοιότητας στο εσωτερικό των συστάδων και την ελαχιστοποίηση της ομοιότητας μεταξύ των συστάδων (Σχήμα 7).



Σχήμα 7 : Διαχωρισμός συνόλου δεδομένων σε συστάδες

Τα μειονεκτήματα των περισσότερων παραδοσιακών αλγορίθμων συσταδοποίησης είναι η υψηλή υπολογιστική πολυπλοκότητα και η μείωση της απόδοσης τους για πολύ μεγάλα σύνολα δεδομένων, οπότε η ανάπτυξη βελτιωμένων αλγορίθμων συσταδοποίησης έχει λάβει μεγάλη προσοχή τα τελευταία χρόνια. Υπάρχουν διάφορες μέθοδοι συσταδοποίησης που μπορούν να χρησιμοποιηθούν για το χειρισμό πολύ μεγάλων συνόλων δεδομένων. Αυτές οι

μέθοδοι μπορούν να κατηγοριοποιηθούν σε αλγόριθμους διαχωρισμού, ιεραρχίας, πλέγματος και πυκνότητας [18,19].

- Οι αλγόριθμοι διαχωρισμού προσπαθούν να προσδιορίσουν τα διαμερίσματα k που βελτιστοποιούν μια συγκεκριμένη αντικειμενική συνάρτηση, όπου κάθε αντικείμενο έχει εκχωρηθεί στο πλησιέστερο σύμπλεγμα. Οι συστάδες αντιπροσωπεύονται από τον μέσο όρο, όπως ο αλγόριθμος $k - mean$ ή από το πιο κεντρικά τοποθετημένο αντικείμενο όπως ο $k - medoid$.
- Οι ιεραρχικές μέθοδοι δημιουργούν μια ιεραρχική αποσύνθεση του εξεταζόμενου συνόλου αντικειμένων / δεδομένων που σχηματίζουν ένα δενδρόγραμμα σε ένα συσσωρευτικό ή διαιρετικό τρόπο. Η μέτρηση απόστασης για τη συγχώνευση / διάσπαση συστάδων μπορεί να επιλεγεί με διάφορους τρόπους προκειμένου να κατασκευαστεί το δενδρόγραμμα. Το ευρέως χρησιμοποιούμενο μέτρο είναι η απόσταση μεταξύ των μέσων των δύο συστάδων, ωστόσο η μέθοδος αυτή αποτυγχάνει στην ανίχνευση τυχαίων σχημάτων συστάδων. Τέλος η απαίτηση για ενημέρωση της απόστασης σε κάθε επανάληψη για τη μέτρηση της ομοιότητας μεταξύ συμπλεγμάτων είναι ένα άλλο μειονέκτημα που καθιστά τους περισσότερους από τους ιεραρχικούς αλγόριθμους ομαδοποίησης μη αποτελεσματικούς για πολύ μεγάλα σύνολα δεδομένων.
- Αλγόριθμοι ομαδοποίησης που βασίζονται σε πλέγμα όπως ο SOM και ο CLIQUE χρησιμοποιούν μια δομή δεδομένων πλέγματος όπου χωρίζει τον χώρο εισόδου σε ορθογώνια τμήματα, απορρίπτοντάς τα τμήματα χαμηλής πυκνότητας και σχηματίζονται συστάδες με τη συγχώνευση γειτονικών τμημάτων υψηλής πυκνότητας.
- Οι αλγόριθμοι με βάση την πυκνότητα όπως οι DBSCAN και DENCLUE ξεπερνούν τη δυσκολία ανίχνευσης αυθαίρετων σχηματισμών συστάδων, εξάγοντας πυκνές

περιοχές και προσδιορίζοντάς τους ως ομάδες που χωρίζονται από περιοχές χαμηλής πυκνότητας, υποθέτοντας ότι τα σημεία εντός συστάδων είναι πυκνά και τα σημεία σε διαφορετικά συσσωματώματα δεν είναι. Οι αλγόριθμοι με βάση την πυκνότητα δεν χρειάζονται εκ' των προτέρων τον αριθμό των συστάδων, αφού ανιχνεύουν αυτόματα τις συστάδες με τον φυσικό τους αριθμό. Ένα από τα μειονεκτήματα αυτής της τεχνικής είναι η υψηλή υπολογιστική πολυπλοκότητα αυτών των μεθόδων επειδή εξετάζονται όλες οι γειτονικές περιοχές κάθε σημείου. Ιδίως κατά τον έλεγχο της κατάστασης πυρήνα για κάθε αντικείμενο, όταν ο αλγόριθμος τρέχει σε πολύ μεγάλα σύνολα δεδομένων, το βήμα μπορεί να γίνει πολύ δαπανηρό.

2.6.2 Ο αλγόριθμος DBSCAN

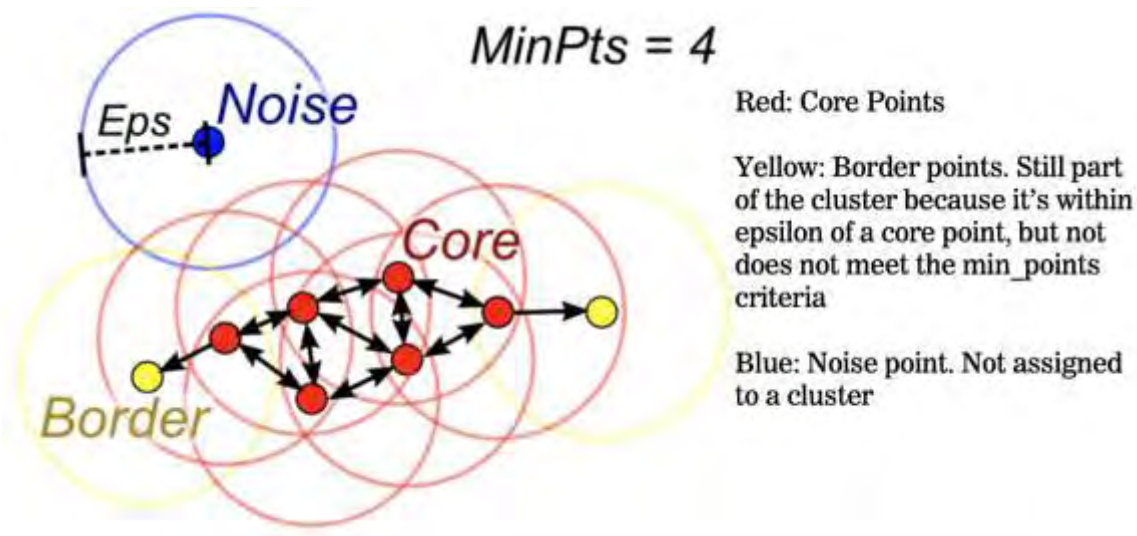
Ο DBSCAN είναι ο “πρωτόπλαστος” αλγόριθμος μεθόδων συσταδοποίησης με βάση την πυκνότητα, ο οποίος χρησιμοποιεί τον αριθμό των αντικειμένων-δεδομένων για να μετρήσει την πυκνότητα τους σε μια συγκεκριμένη εξαιρετικά σφαιρική περιοχή. Αν και ο DBSCAN δεν μπορεί μόνο να ανακαλύψει αυθαίρετες συστάδες, αλλά και να προσδιορίσει σημεία θορύβου ή ακραίες τιμές, οι δύο παράμετροι του *EPS* και ο ελάχιστος αριθμός σημείων (*MinPts*) που απαιτείται για την δημιουργία μιας συστάδας, είναι απαραίτητο να αναληφθούν χειροκίνητα εκ των προτέρων. Για τον προσδιορισμό της τιμής του *EPS* προτείνονται διάφορες μέθοδοι από τη [20]. Ο αλγόριθμος αυτός είναι ιδιαίτερα χρήσιμος για την αντιμετώπιση μεγάλων συνόλων δεδομένων, με θόρυβο και είναι σε θέση να προσδιορίσει συστάδες με διαφορετικά μεγέθη και σχήματα.

Η βασική ιδέα του αλγορίθμου DBSCAN είναι ότι, για κάθε σημείο ενός συμπλέγματος, η γειτονιά μιας δεδομένης ακτίνας (*EPS*) πρέπει να περιέχει τουλάχιστον έναν ελάχιστο αριθμό σημείων (*MinPts*), δηλαδή η πυκνότητα στη γειτονιά πρέπει να υπερβεί κάποιο προκαθορισμένο όριο. Αυτός ο αλγόριθμος χρειάζεται τρεις παραμέτρους εισόδου:

- k , το μέγεθος του συνόλου δεδομένων.
- EPS , η ακτίνα που απομεινείται την περιοχή γειτονίας ενός σημείου.
- $MinPts$, ο ελάχιστος αριθμός σημείων που πρέπει να υπάρχουν στην περιοχή.

Η διαδικασία συσταδοποίησης βασίζεται στην κατάταξη των σημείων στο σύνολο δεδομένων ως βασικών σημείων, συνοριακών σημείων και σημείων θορύβου, καθώς και στη χρήση των σχέσεων πυκνότητας μεταξύ των σημείων αυτών με στόχο το σχηματισμό συστάδων. Η διαδικασία συσταδοποίησης ακολουθεί τα παρακάτω βήματα:

- Διαιρεί το σύνολο δεδομένων σε n διαστάσεις.
- Για κάθε σημείο του συνόλου δεδομένων ο DBSCAN σχηματίζει ένα σχήμα n διαστάσεων γύρω από αυτό το σημείο δεδομένων και, στη συνέχεια, υπολογίζει τον αριθμό των σημείων που εμπίπτουν σε αυτό το σχήμα.
- Ο DBSCAN καταλογίζει αυτό το σχήμα ως συστάδα εάν ο αριθμός των σημείων εντός του σχήματος είναι μεγαλύτερος από τον προκαθορισμένο ελάχιστο αριθμό.
- Στη συνέχεια, εξετάζοντας κάθε σημείο μέσα στη συστάδα και μετρώντας τον αριθμό των άλλων σημείων δεδομένων που βρίσκονται κοντά διευρύνει την συστάδα. Μια συνοπτική αναπαράσταση τις διαδικασίας φαίνεται στο Σχήμα 8.



Σχήμα 8 : Γραφική αναπαράσταση του αλγορίθμου DBSCAN

Με το τέλος της διαδικασίας μπορούμε να διαχωρίσουμε τα σημεία - δεδομένα σε τρεις κατηγορίες. Τα σημεία τα οποία λαμβάνονται ως θόρυβος και ως εκ τούτου δεν περιλαμβάνονται σε κάποια συστάδα, τα σημεία που θεωρούνται συνοριακές τιμές, σημεία που δεν καλύπτουν τις απαιτήσεις του $MinPts$ αλλά ανήκουν σε συστάδες καθώς ικανοποιούν τις απαιτήσεις του EPS και τέλος τα σημεία πυρήνα που καλύπτουν τόσο τις προϋποθέσεις του EPS όσο και του $MinPts$. Είναι εμφανές ότι κρίσιμο ρόλο στην απόδοση του αλγορίθμου διαδραματίζουν οι τιμές των EPS και $MinPts$. Για ένα δεδομένο EPS , ο αριθμός των αντικειμένων πυρήνα μειώνεται σημαντικά όταν η τιμή του $MinPts$ είναι υπερβολική. Σε αυτήν την περίπτωση, ορισμένες συστάδες που περιέχουν λίγα αντικείμενα ενδέχεται να απορριφθούν. Αντιθέτως, ορισμένα αντικείμενα θα κριθούν ως συνοριακά αντικείμενα κατά λάθος όταν η τιμή του $MinPts$ είναι πολύ μικρή ή μερικά σημεία θορύβου μπορεί να ταξινομηθούν στις κανονικές συστάδες εσφαλμένα.

Προτείνεται λοιπόν από τη [21] μια μέθοδος προσαρμοζόμενου προσδιορισμού των παραμέτρων. Η κύρια ιδέα είναι οι τιμές των παραμέτρων EPS και $MinPts$ να διαπιστώνονται με βάση τις στατιστικές ιδιότητες του συνόλου δεδομένων. Για λόγους ευκολίας, πρέπει να υπολογιστεί εκ των προτέρων ένας πίνακας διανομής $DIST$.

$$DIST_{n*n} = \{dist(i,j) | 1 \leq i \leq n, 1 \leq j \leq n\} \quad (6)$$

Στην (6), το n αντιπροσωπεύει τον αριθμό των αντικειμένων στο σύνολο δεδομένων D . Ο $DIST$ είναι ένας συμμετρικός πίνακας με n γραμμές και n στήλες, στον οποίο κάθε στοιχείο υποδηλώνει την απόσταση μεταξύ των αντικειμένων i και j στο D . Υπολογίζουμε την τιμή κάθε στοιχείου στο $DIST_{n*n}$ διαδοχικά και, στη συνέχεια, τα ταξινομήσαμε σε αύξουσα σειρά γραμμή προς γραμμή. Χρησιμοποιούμε τη $DIST_{n*i}$ για να εκφράσουμε την τιμή της i -ωστης στήλης στο $DIST_{n*n}$. Καθώς όλα τα δεδομένα στο $DIST_{n*i}$ υπακούουν στη κατανομή Poisson, χρησιμοποιούμε μαθηματικά και την εκτίμηση μέγιστης πιθανότητας για να εκτιμήσουμε την αξία της παραμέτρου λ . Δηλαδή, το λ μπορεί να υπολογιστεί μέσω του γεωμετρικού μέσου των τιμών του $DIST_{n*n}$.

$$\lambda = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (7)$$

Αφού προσδιοριστεί το EPS , ακολουθεί ο υπολογισμός των $MinPts$. Πρώτον, ο αριθμός των αντικειμένων - δεδομένων στη γειτονιά EPS κάθε σημείου του συνόλου δεδομένων υπολογίζεται ένα προς ένα. Στη συνέχεια η μαθηματική πιθανότητα, όλων αυτών των αντικειμένων - δεδομένων υπολογίζεται, η οποία είναι η τιμή των $MinPts$.

$$MinPts = \frac{1}{n} \sum_{i=1}^n p_i \quad (8)$$

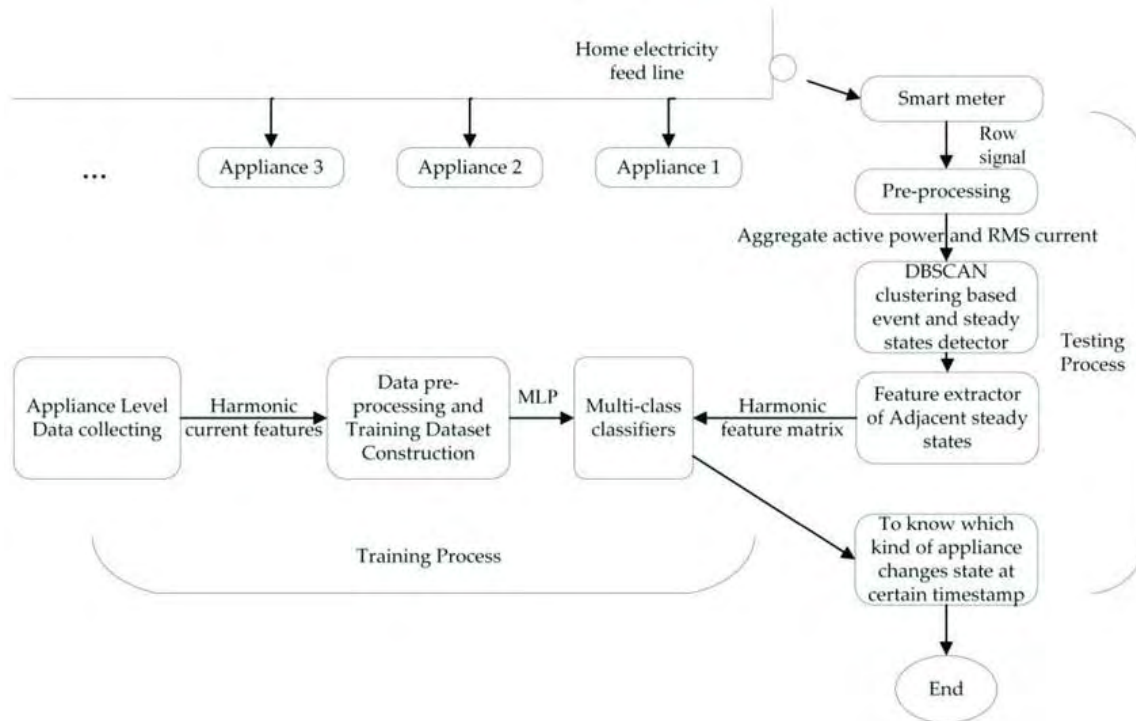
Τέλος τα βασικά μειονεκτήματα του αλγορίθμου είναι :

- Η υψηλή πολυπλοκότητα, σε ορισμένες περιπτώσεις η πολυπλοκότητα του φθάνει στο $O(n^2)$.
- Δεν είναι δυνατή η σωστή λειτουργία μεταξύ συνόλων δεδομένων διαφορετικής πυκνότητας.

- Απαιτείται πολύς χώρος στη μνήμη για τη φόρτωση ολόκληρου του συνόλου δεδομένων.

2.7 Η προσέγγιση της παρούσας εργασίας

Στην παρούσα εργασία υλοποιήθηκε ένα μοντέλο παραγωγής υπογραφών φορτίου με μη εποπτευόμενες τεχνικές μηχανικής μάθησης με σκοπό τη χρήση αυτών σε ένα σύστημα μη-παρεμβατικής αναγνώρισης φορτίων. Το μοντέλο που υλοποιήθηκε κάνει χρήση μετρήσεων του ρεύματος, από έναν μοναδικό μετρητή εγκατεστημένο στον κεντρικό πίνακα της οικίας, των αρμονικών παραμορφώσεων και των γωνιών υπό τις οποίες αυτές εμφανίζονται και παράγει αυτόματα τις υπογραφές των φορτίων. Εν συνεχεία χρησιμοποιούμε τις υπογραφές που έχουμε παράγει για την αναγνώριση της εκάστοτε συσκευής ή του συνδυασμού αυτών, Σχήμα 9. Πρόκειται επί της ουσίας για ένα συνδυασμό εννοιών και προβλημάτων που αναφέρθηκαν παραπάνω [20].



Σχήμα 9 : Σχηματικό διάγραμμα της προσέγγισης που ακολουθήθηκε

ΚΕΦΑΛΑΙΟ 3

ΠΡΑΚΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

3.1 Εισαγωγή

Στο κεφάλαιο αυτό γίνεται μικρή παρουσίαση της δομής και των τεχνολογιών που χρησιμοποιήθηκαν για την ανάπτυξη του συστήματος. Πιο συγκεκριμένα, το σύστημα αποτελείται από τρία ξεχωριστά μέρη στο προγραμματιστικό σκέλος. Ο πυρήνας του συστήματος είναι μια αλγοριθμική αλληλουχία η οποία παράγει τις υπογραφές των φορτίων. Με αυτό επικοινωνούν μια βάση δεδομένων η οποία συνεχώς ανανεώνεται με μετρήσεις πραγματικού χρόνου καθώς και ένα δεύτερο προγραμματιστικό κομμάτι το οποίο λαμβάνοντας τις υπογραφές που έχει παράγει ο αλγόριθμος αναγνωρίζει τις ενεργές συσκευές.

3.2 Μετρήσεις

3.2.1 Ρύθμιση μέτρησης

Στην εργασία [6], για τον σκοπό της διεξαγωγής των μετρήσεων με όσο το δυνατόν περισσότερη αυτονομία, έχει προταθεί η δημιουργία μιας προσαρμοσμένης ρύθμισης

μέτρησης. Αυτό περιλαμβάνει έναν αισθητήρα εναλλασσόμενου ρεύματος (LEM PR630, εύρος μετρήσεων $\pm 60\text{A}$ με ακρίβεια $\pm 1\%$, εύρος συχνοτήτων DC έως 20 kHz) και έναν οπτό - απομονωμένο μετατροπέα τάσης (εύρος μέτρησης $\pm 600\text{V}$). Οι συντελεστές μεταφοράς του μετατροπέα τάσης και του αισθητήρα ρεύματος είναι 5V/600V και 10 mV/A, αντίστοιχα. Μια κάρτα απόκτησης δεδομένων (κάρτα DAQ, NI PXI-6123) καταγράφει τις εκροές του μετατροπέα τάσης και του αισθητήρα ρεύματος. Η προτεινομένη δειγματοληψία των δεδομένων ορίστηκε στα 50 kHz. Επιπλέον, συνιστάται μια περίοδος απόκτησης 1 ή 50 περιόδων της βασικής συχνότητας της γραμμής τροφοδοσίας, έτσι ώστε κάθε ρεύμα και κυματομορφή τάσης που καταγράφεται να περιείχε 50 kHz. Η ταυτόχρονη απόκτηση των τιμών ρεύματος και τάσης επιβεβαιώθηκε καθώς η καταγραφή των δύο ποσοτήτων ενεργοποιήθηκε όταν η τάση γραμμής ήταν σε μηδενικό επίπεδο με θετική κλίση, δηλαδή, στην αρχή μιας νέας περιόδου.

Η αποκληθείσα αναλογική κυματομορφή μεταποιείται και αναλύεται σε πραγματικό χρόνο μέσω της ταχείας μετατροπής Φουριέ (FFT) μέσω μιας εφαρμογής που αναπτύχθηκε με τη χρήση της γραφικής γλώσσας προγραμματισμού στο LabVIEW. Αυτή η ανάλυση [6] FFT επιτρέπει τον προσδιορισμό των φάσεων και των συχνοτήτων των θεμελιωδών και αρμονικών συστατικών του ρεύματος (τρέχον φάσμα ισχύος). Έχοντας ορίσει την προαναφερθείσα διαμόρφωση απόκτησης (50 kHz και 1s περίοδος απόκτησης), το ελάχιστο χρονικό διάστημα συχνότητας που επιτεύχθηκε στο φάσμα ισχύος (ανάλυση συχνότητας του φάσματος) ήταν 1 Hz.

Ένα χαμηλοπερατό φίλτρο είναι απαραίτητο για τον αλγόριθμο FFT, ώστε να συμμορφώνεται με το θεώρημα του Nyquist. Για τις μετρήσεις της μελέτης έχουμε εφαρμόσει στις αποκληθείσες κυματομορφές ένα χαμηλοπερατό φίλτρο με συχνότητα αποκοπής στα 25 kHz. Η ανάλυση FFT που εφαρμόζεται στο φιλτραρισμένο σήμα ρεύματος στο φάσμα της συχνότητας των 25 kHz αποκάλυψε ένα φάσμα ισχύος που αποτελείται από μια κανονική σειρά κορυφών σε θεμελιώδεις (50 Hz) και υψηλότερες μονές αρμονικές (150 Hz, 250 Hz,...) χωρίς ενδιάμεσες κορυφές και ένα σχεδόν αμελητέο ποσοστό θορύβου. Το μέτρο της

συμβολής του θορύβου στη διακύμανση του εύρους των παρατηρούμενων αρμονικών δίνεται από τη μετρούμενη τυπική απόκλιση των κορυφών.

Η μέθοδος μαθηματικής ανάλυσης FFT, που χρησιμοποιείται σε αυτό το κομμάτι, λαμβάνει το τρέχον σήμα και αναζητά ένα καθορισμένο εύρος συχνοτήτων, επιστρέφοντας ένα ενιαίο τόνο (αρμονικό περιεχόμενο) συχνότητας, πλάτος και φάση. Αυτή η περιοχή αναζήτησης συχνότητας ελέγχεται από το χρήστη ορίζοντας την κεντρική συχνότητα και το πλάτος. Ως κεντρικές συχνότητες για την αναζήτηση της βασικής, τρίτης και πέμπτης αρμονικής είχαν οριστεί, οι τιμές των 50, 150 και 250 Hz αντίστοιχα. Ο αλγόριθμος FFT μπορεί να αναζητήσει ένα μόνο τόνο στο εύρος πλήρους συχνότητας ($0, f_s/2$), όπου, f_s αντιπροσωπεύει το ποσοστό δειγματοληψίας, στην περίπτωσή μας, 50 kHz. Από την άλλη πλευρά, το πλάτος συχνότητας δίνεται ως ποσοστό του ρυθμού δειγματοληψίας, και χρησιμοποιείται για τον περιορισμό του εύρους αναζήτησης, βελτιώνοντας έτσι την αναλογία σήματος προς θόρυβο. Φυσικά το πλάτος της συχνότητας δεν πρέπει να είναι υπερβολικά μικρό, υπό την έννοια ότι ένα πλάτος μικρότερο από το εγγενές FWHM (foul with at half maximum) της αρμονικής συν κάποιες πιθανές διακυμάνσεις στην κεντρική συχνότητα της αρμονικής. Στις μετρήσεις μας, η παράμετρος πλάτους έχει οριστεί στο 0,004, με αποτέλεσμα το πλάτος της περιοχής αναζήτησης συχνότητας να είναι $\pm 0,004\% \times 50 \text{ kHz} = \pm 2 \text{ Hz}$. Δεδομένου ότι το FWHM των κανονικών αρμονικών είναι της τάξης των $\pm 1 \text{ Hz}$, είναι προφανές ότι οι διακυμάνσεις της κεντρικής συχνότητας από $\pm 0,1 \text{ Hz}$ (ή ακόμα υψηλότερες) δεν επηρεάζουν την εκτίμηση συχνότητας και πλάτους των αρμονικών.

3.2.2 Μετρούμενες συσκευές και συνδυασμοί

Οι μετρήσεις εκτελέστηκαν σε μια τυπική οικιακή εγκατάσταση 3 φάσεων με ονομαστική ισχύ 25 kVA. Η απόκτηση δεδομένων τάσης και ρεύματος εκτελέστηκε απευθείας στον κύριο πίνακα τροφοδοσίας της εγκατάστασης. Κάθε συσκευή μετρήθηκε για περίπου 4 – 5 λεπτά, καθώς για τις περισσότερες από αυτές ο χρόνος αυτός ήταν επαρκής για να απεικονίσει την

λειτουργική συμπεριφορά σταθερής κατάστασης. Στον Πίνακα 1 εμφανίζονται τόσο οι μετρούμενες συσκευές όσο και οι συνδυασμοί συσκευών. Οι εξεταζόμενες συσκευές έχουν κατηγοριοποιηθεί σε τρεις βασικές ομάδες: συσκευές με συμπεριφορά αντίστασης (δηλ. συσκευές 3,4,5), συσκευές με κινητήρες (δηλ. συσκευές 1,2) και συσκευές με ηλεκτρονικά μέσα (δηλ. συσκευές 6,7) [6].

Για να διευκολυνθεί ο διασταυρούμενος έλεγχος όσον αφορά την επιτυχή ταυτοποίηση του φορτίου κατά τη διάρκεια του συστήματος διαχωρισμού, οι συνδυασμοί συσκευών εκτελέστηκαν με ελεγχόμενο τρόπο. Συγκεκριμένα, για κάθε συνδυασμό οι συσκευές ενεργοποιήθηκαν με διαδοχική σειρά με προκαθορισμένα χρονικά διαστήματα.

Πίνακας 1 : Μετρούμενες συσκευές και συνδυασμοί αυτών

	<u>ΣΥΣΚΕΥΕΣ</u>		<u>Ενδεικτικοί συνδυασμοί</u>
1	Πιστολάκι	A	Τηλεόραση + πιστολάκι + υπολογιστής
2	Σκούπα	B	Τηλεόραση + πιστολάκι + σίδερο
3	Τοστιέρα	Γ	Σκούπα + υπολογιστής + σίδερο
4	Μάτι κουζίνας	Δ	Τοστιέρα + σκούπα + πιστολάκι
5	Σίδερο		
6	Υπολογιστής		
7	Τηλεόραση		

3.2.3 Μορφή ακατέργαστων δεδομένων

Ενδεικτικά στον Πίνακα 2 διακρίνεται το αποτέλεσμα των μετρήσεων και η μορφή των αρχικών ακατέργαστων δεδομένων. Όπως είναι εμφανές, η πληροφορία που παρέχεται αφορά μόνο το ολικό ρεύμα της εγκατάστασης και παρέχει πληροφορίες για αρμονικές και γωνίες αυτών. Τέλος κάθε συνδυασμός συσκευών αναπαρίσταται σε διαφορετικό πίνακα και αποτελείται από 300 εγγραφές (5 λεπτά καταγραφής με δειγματοληψία /sec).

Πίνακας 2 : Συνολικά αποτελέσματα και δομή μετρήσεων

Ολικό (2/1/2019)								
50Hz	I50 [A]	Φ50 [deg]	150Hz	I150 [A]	Φ150 [deg]	250Hz	I250 [A]	Φ250 [deg]
50.030	3.504	175.400	150.000	0.515	10.840	250.100	0.424	-155.900
50.010	3.613	178.600	150.000	0.606	10.480	250.100	0.443	-150.400
49.990	3.514	179.900	150.100	0.566	0.923	250.000	0.469	-147.500
49.970	3.444	-178.200	150.000	0.549	9.659	249.900	0.426	-134.300
50.010	3.473	175.700	150.000	0.539	12.700	250.000	0.395	-148.800
50.020	3.478	175.100	150.100	0.508	1.452	250.100	0.408	-165.100
50.030	3.494	172.400	149.900	0.498	27.400	250.100	0.415	-158.800
49.990	3.404	179.300	150.100	0.535	6.822	250.100	0.412	-161.400
50.000	3.470	-179.400	150.100	0.560	1.460	250.100	0.400	-159.100

Επιπλέον, με σκοπό την επαλήθευση των αποτελεσμάτων, εκτός του ολικού φορτίου κάθε συνδυασμού κατεγράφησαν παράλληλα και οι επιμέρους τιμές των συσκευών που συμμετείχαν στην διάταξη. Η αντίστοιχη πληροφορία απεικονίζεται στον Πίνακα 3.

Πίνακας 3 : Επιμέρους αποτελέσματα και δομή μετρήσεων για φορτίου υπολογιστή

Laptop								
50Hz	I50 [A]	Φ50 [deg]	150Hz	I150 [A]	Φ150 [deg]	250Hz	I250 [A]	Φ250 [deg]
50.020	0.315	12.120	150.100	0.277	-157.700	250.100	0.233	38.820
50.020	0.321	10.170	150.100	0.282	-159.600	250.100	0.238	36.740
50.010	0.311	13.160	150.000	0.272	-155.700	250.000	0.230	42.350
50.010	0.311	12.130	150.000	0.273	-155.900	250.000	0.231	41.420
50.010	0.311	12.660	150.000	0.272	-157.000	250.100	0.231	39.850
50.000	0.309	13.460	150.000	0.270	-155.300	250.000	0.229	41.940
50.010	0.309	12.210	150.000	0.272	-157.700	250.000	0.229	39.240
50.010	0.310	11.930	150.000	0.272	-158.300	250.100	0.230	36.860
50.010	0.310	11.790	150.000	0.272	-157.700	250.100	0.229	38.440
50.010	0.310	11.450	150.000	0.271	-158.500	250.100	0.229	37.140
50.010	0.312	-168.400	150.000	0.273	22.740	250.000	0.230	-140.400

3.3 Μεθοδολογίες

3.3.1 Ο αλγόριθμος του Hart

Ο αλγόριθμος που εισήγαγε ο Hart στις αρχές του 80' θεωρείται θεμελιώδης και είναι η βάση πολλών από τις υπάρχουσες τεχνικές NILM. Η γενική ιδέα του αλγορίθμου του Hart είναι να μετρήσει τη συνολική κατανάλωση ισχύος ενός νοικοκυριού, να προσδιορίσει συσκευές και να παρακολουθήσει τη συμπεριφορά τους. Ο αλγόριθμος εκτελεί τις ακόλουθες εργασίες [3]:

1. Μέτρηση ισχύος και τάσης.
2. Υπολογισμός κανονικοποιημένης ισχύος
3. Ανίχνευση ακμών
4. Διαχωρισμός συστάδων
5. Δημιουργία Μοντέλων Συσκευής
6. Παρακολούθηση συμπεριφοράς
7. Δημιουργία πίνακα στατιστικών
8. Ονοματολογία συσκευής

Ο αλγόριθμος που αναπτύχθηκε ακολουθεί τα βήματα του αλγορίθμου του Hart με κάποιες τροποποιήσεις σε αυτά, με σκοπό την βελτιστοποίηση του και μπορεί να περιγραφεί στις φάσεις που αναφέρονται παρακάτω.

Το συνολικό πλαίσιο της προτεινόμενης προσέγγισης της NILM απεικονίζεται στο Σχήμα 9. Η προσέγγιση διαιρείται σε διαδικασία εκπαίδευσης και διαδικασία δοκιμών. Κατά τη διαδικασία εκπαίδευσης, τα δεδομένα σε επίπεδο συσκευών συλλέγονται από συσκευές και ένα σύνολο δεδομένων εκπαίδευσης κατασκευάζεται μετά από επεξεργασία. Στη διαδικασία δοκιμής, τα συγκεντρωτικά δεδομένα σε επίπεδο σπιτιού αποκτώνται από τον έξυπνο μετρητή και στη συνέχεια διεξάγεται προ επεξεργασία για να εξομαλύνει την διαδικασία αναγνώρισης. Ο αλγόριθμος DBSCAN χρησιμοποιείται στη συνέχεια για την ανίχνευση των σταθερών καταστάσεων. Μετά από αυτό, η εύρεση του μέσου όρου και η αφαίρεση μεταξύ σταθερών καταστάσεων διεξάγονται για την εξαγωγή του νέου πίνακα καταστάσεων. Ο εντοπισμός και η ταξινόμηση συμβάντων μπορούν να χρησιμοποιήσουν διάφορες μετρήσεις. Σε αυτή την προσέγγιση όμως χρησιμοποιούνται οι τιμές ρεύματος και αρμονικών συχνοτήτων. Στην

επόμενη φάση το προτεινόμενο μοντέλο εκτελεί ξανά την διαδικασία συσταδοποίησης και καταλήγει στις τελικές υπογραφές φορτίου.

3.3.2 Αναγνώριση συμβάντων (ανίχνευση ακμών)

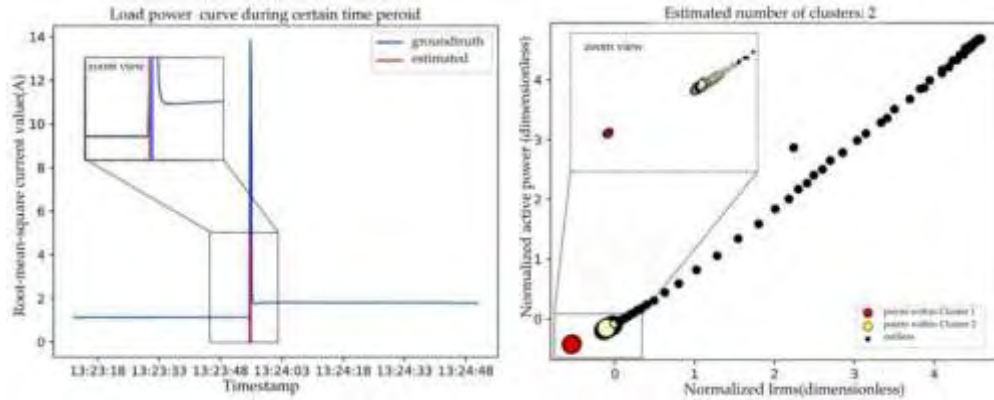
Κύριος στόχος αυτής της εργασίας είναι η χρήση των αρμονικών συχνοτήτων και της μεθόδου συσταδοποίησης DBSCAN για την ανίχνευση συμβάντων και παρακείμενων σταθερών καταστάσεων. Λαμβάνοντας υπ' όψη την μεθοδολογία του Hart που αποτελεί ακρογωνιαίο λίθο στη λειτουργία κάθε μοντέλου NILM και θεωρώντας πως τα δυο πρώτα βήματα περιγράφονται επαρκώς από τα κεφάλαιο 3.2, το κεφάλαια που ακολουθούν καλύπτουν τα επόμενα βήματα του αλγορίθμου με το παρόν να αντιστοιχεί στο βήματα τρία, τέσσερα και πέντε.

Σε αυτό το μοντέλο, δύο παρακείμενες σταθερές καταστάσεις θεωρούνται ως δύο συστάδες και τα παροδικά τους διαστήματα θεωρούνται θόρυβος ή ακραίες τιμές, όπως φαίνεται στο Σχήμα 10. Τα πλεονεκτήματα της χρήσης του αλγορίθμου DBSCAN για ανίχνευσή συμβάντων περιλαμβάνουν [20]:

- Απόρριψη αποτελέσματος θορυβωδών σημείων που προκαλούνται από συσκευή μέτρησης ή άλλους λόγους.
- Δυνατότητα διακριτοποίησης συμβάντων με μικρό επίπεδο αλλαγής ενέργειας, επειδή ο αλγόριθμος αυτός δεν βασίζεται στο όριο αλλαγής ισχύος.
- Δυνατότητα διακριτοποίησης συμβάντων με μεγάλο μεταβατικό διάστημα, όπως υπολογιστές (υπό προϋποθέσεις).

Αυτός ο αλγόριθμος δεν χρειάζεται να προσδιορίσει τον αριθμό των συστάδων εκ των προτέρων και έχει διαισθητικές παραμέτρους στην εφαρμογή του. Η βασική ιδέα όπως αναλύθηκε στο 2.7.2 είναι ότι δεδομένης μιας καθορισμένης ακτίνας *EPS* για ένα σημείο σε

ένα σύμπλεγμα, ο αριθμός των σημείων στη γειτονιά του θα πρέπει να υπερβεί το δεδομένο κατώφλι [20].



Σχήμα 10 : Αναγνώριση συμβάντων με την χρήση του αλγορίθμου DBSCAN

Βασικό ζήτημα για την απόδοση του αλγορίθμου αποτελεί η επιλογή της τιμής του EPS . Η επιλογή της τιμής αυτής στη δικιά μας πρόταση έγινε με βάση την αναλογία των μέσων τιμών και τυπικής απόκλισης των τιμών του υπό εξέταση δείγματος κάθε φορά. Με το EPS να ορίζεται σαν :

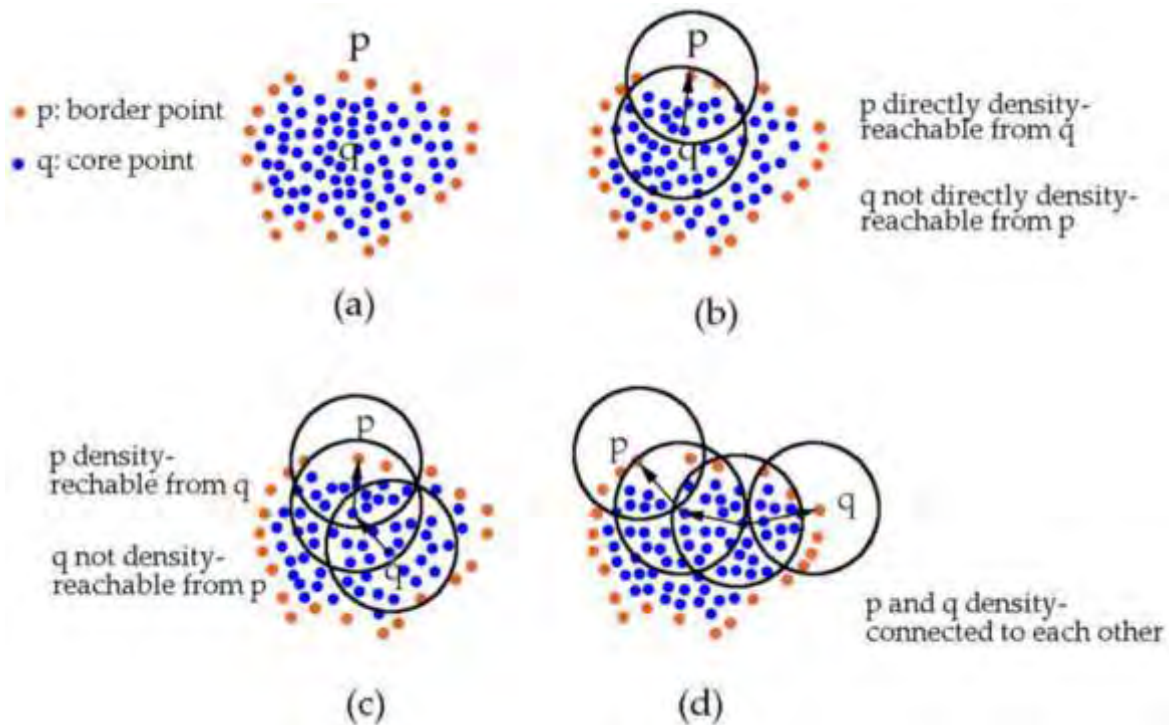
$$EPS = Std/2 \quad (8)$$

Όσον αφορά την τιμή $MinPts$ η επιλογή της τιμής έγινε με βάση το μέγεθος του δείγματος N και ορίστηκε σαν :

$$MinPts = (N/100) + 1 \quad (9)$$

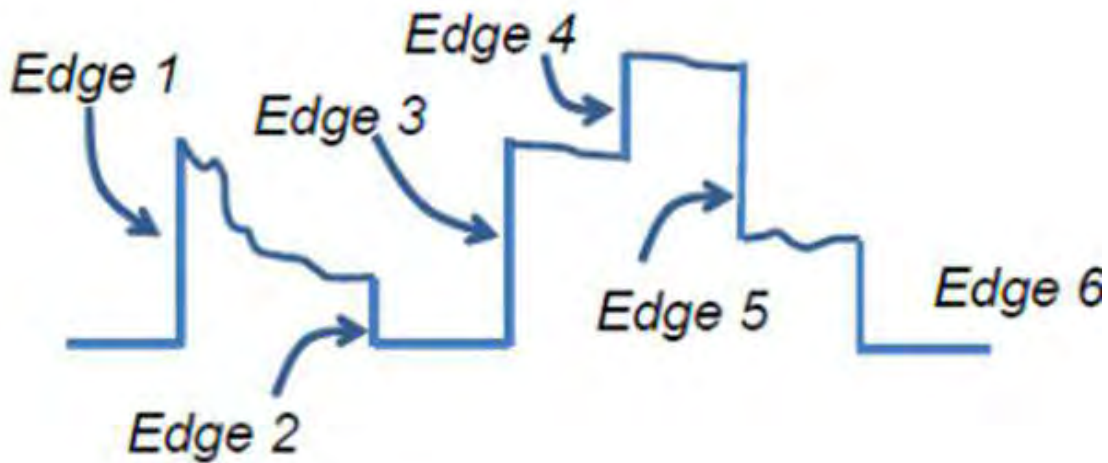
Όπως δείχνει το Σχήμα 11, υπάρχουν δύο τύποι σημείων σε ένα σύμπλεγμα: σημεία μέσα στο σύμπλεγμα (βασικά σημεία) και σημεία στα σύνορα του συμπλέγματος (σημεία περιγράμματος). Για να συμπεριληφθούν τα σημεία περιγράμματος στο σύμπλεγμα, ορίζονται οι έννοιες άμεσα προσβάσιμο, προσβάσιμο και συνδεδεμένο. Το σημείο p είναι άμεσα προσβάσιμο από το σημείο q εάν το $p \in N_{EPS}(q)$ και $|N_{EPS}(q)| \geq MinPts$, όπου $N_{EPS}(q)$ υποδηλώνει το σύνολο των σημείων στη γειτονιά EPS του q . Το σημείο q λογίζεται ως σημείο πυρήνα. Ένα σημείο p είναι προσβάσιμο από ένα σημείο q αν υπάρχει μια αλυσίδα των πόντων $P_1, \dots, P_n, P_1 = q, P_n = p$ έτσι ώστε P_{i+1} να είναι άμεσα προσβάσιμη από P_i . Ένα

σημείο p είναι συνδεδεμένο σε ένα σημείο q αν υπάρχει ένα σημείο k τέτοιο ώστε p και q να είναι αμφότερα προσβάσιμα από το k .



Σχήμα 11 : Διαδικασία συσταδοποίησης του αλγορίθμου DBSCAN

Ο ορισμός του άμεσα προσβάσιμου κρατά μαζί τα σημεία των συνόρων και τα βασικά σημεία. Ο ορισμός του προσβάσιμου συνδέεται με τα σημεία περιγράμματος. Έτσι, όλα τα σημεία σε ένα σύμπλεγμα περιλαμβάνονται με τον καθορισμό των EPS και του $MinPts$. Στη δικιά μας περίπτωση απαιτούνται δύο βήματα για την ανακάλυψη ενός συμπλέγματος. Πρώτον, επιλέγουμε ένα αυθαίρετο σημείο που ικανοποιεί τον ορισμό του κύριου σημείου από τη βάση δεδομένων. Δεύτερον, ανακαλύπτουμε όλα τα σημεία που είναι προσβάσιμα από αυτό το σημείο (Σχήμα 11). Με την εφαρμογή της παραπάνω διαδικασίας καταλήγουμε σε ένα σύνολο δεδομένων με διακριτές πλέον τις ακμές (Σχήμα 12) και έχοντας απομακρύνει ακραίες τιμές. Στο τέλος της διαδικασίας αποθηκεύουμε τις μέσες τιμές κάθε συστάδας. Έχουμε λοιπόν εξάγει μια πρώιμη μορφή υπογραφών φορτίου, οι οποίες περιγράφουν τον τρόπο λειτουργίας των συνδυασμών που επιλέγουμε να εξετάσουμε σε κάθε σύνολο δεδομένων [20].



Σχήμα 12 : Σχηματική απεικόνιση των ακμών κατά τη διάρκεια μιας μέτρησης

3.3.3 Μετά – επεξεργασία δεδομένων

Δεδομένου ότι κάθε μία από τις μέσες τιμές που ανακαλύφθηκαν στο 3.2.3 αποτελεί ένα συνδυασμό από ενεργές συσκευές προχωράμε στην αφαίρεση κάθε δυνατού συνδυασμού μέσων τιμών. Δεδομένο ότι ο τύπος των υπογραφών φορτίου που χρησιμοποιούμε είναι της μορφής :

$$LS_i^B = [(I_{i1j}^B | \cos\varphi_1), (I_{i3j}^B | \cos\varphi_3), (I_{i5j}^B | \cos\varphi_5)] \quad (4)$$

Η παραπάνω διαδικασία μπορεί να περιγραφεί μαθηματικά ως έξις :

$$LS_i^B = | [(I_{(t)}^{B1} - I_{(t+1)}^{B1} | \cos(\varphi_{1(t+1)}) - \cos(\varphi_{1(t+1)}), (I_{(t)}^{B3} - I_{(t+1)}^{B3} | \cos(\varphi_{3(t+1)}) - \cos(\varphi_{3(t+1)}), (I_{(t)}^{B5} - I_{(t+1)}^{B5} | \cos(\varphi_{5(t+1)}) - \cos(\varphi_{5(t+1)})] \quad (5)$$

Το νέο σύνολο δεδομένων αποτελείται πλέον από ένα μεγάλο όγκο τιμών μέσα στις οποίες θα υπάρχει κάθε δυνατός συνδυασμός συσκευών, κατά συνέπεια και οι μεμονωμένες υπογραφές κάθε συσκευής [20].

Για την υλοποίηση της παραπάνω διαδικασίας προτείνεται η χρήση όλης της δυνατής πληροφορίας που προσφέρουν οι μετρήσεις έτσι:

- Αρχικά μετατρέπουμε τις υπογραφές των τιμών ρεύματος και αρμονικών σε μιγαδικούς αριθμούς ακολουθώντας τις κλασικές μεθόδους μετατροπής μιγαδικών $Re(z) = |I| * \cos(\varphi)$ και $Im(z) = |I| * \sin(\varphi)$. Κάθε υπογραφή αποτελεί ένα διάνυσμα \vec{x} έξι διαστάσεων.
- Δημιουργούμε τον πίνακα $DIFF_{n*6}$ όπου n ο αριθμός των υπογραφών που ήδη έχουμε και στην συνέχεια σαν στήλες και γραμμές ορίζουμε τα αρχικά διανύσματα \vec{x} με τις συνιστώσες του. Ακολούθως πολλαπλασιάζουμε τον $DIFF_{n*6}$ ένα μοναδιαίο πίνακα A_{n*n*6} . Αφαιρώντας από τον πίνακα που προκύπτει (Πίνακας 4) τον αντίστροφο του (Πίνακας 5) έχουμε πλέον τον νέο πίνακα $DIFF2_{n*n*6}$ ο οποίος αποτελείται από μηδενικά στη διαγώνιο και την διαφορά μεταξύ οποιουδήποτε δυνατού συνδυασμού υπογραφών κάτω από αυτή (Πίνακας 6).

Πίνακας 4 : Ο πίνακας $DIFF_{n*6} * A_{n*n*6}$

\vec{X}_i	\vec{X}_i	\vec{X}_i	\vec{X}_i	\vec{X}_i	\vec{X}_i	\vec{X}_i
\vec{X}_{i+1}	\vec{X}_{i+1}	\vec{X}_{i+1}	\vec{X}_{i+1}	\vec{X}_{i+1}	\vec{X}_{i+1}	\vec{X}_{i+1}
\vec{X}_{i+2}	\vec{X}_{i+2}	\vec{X}_{i+2}	\vec{X}_{i+2}	\vec{X}_{i+2}	\vec{X}_{i+2}	\vec{X}_{i+2}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
\vec{X}_{i+n}	\vec{X}_{i+n}	\vec{X}_{i+n}	\vec{X}_{i+n}	\vec{X}_{i+n}	\vec{X}_{i+n}	\vec{X}_{i+n}

Πίνακας 5 : Ο αντίστροφος του $DIFF_{n*6} * A_{n*n*6}$

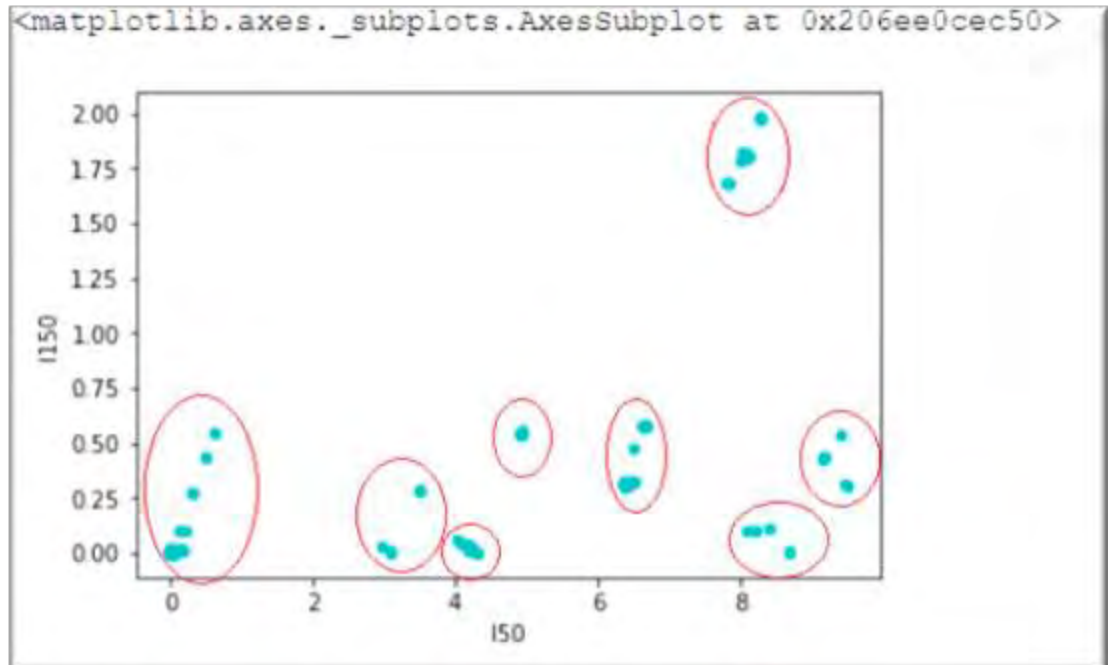
\vec{X}_i	\vec{X}_{i+1}	\vec{X}_{i+2}	\vec{X}_{i+n}	\vec{X}_i
\vec{X}_i	\vec{X}_{i+1}	\vec{X}_{i+2}	\vec{X}_{i+n}	\vec{X}_i

\vec{X}_i	\vec{X}_{i+1}	\vec{X}_{i+2}	\vec{X}_{i+n}	\vec{X}_i
\vec{X}_i	\vec{X}_{i+1}	\vec{X}_{i+2}	\vec{X}_{i+n}	\vec{X}_i
\vec{X}_i	\vec{X}_{i+1}	\vec{X}_{i+2}	\vec{X}_{i+n}	\vec{X}_i
\vec{X}_i	\vec{X}_{i+1}	\vec{X}_{i+2}	\vec{X}_{i+n}	\vec{X}_i

Πίνακας 6 : Ο πίνακας DIFF2

	\vec{X}_i	\vec{X}_{i+1}	\vec{X}_{i+2}	\vec{X}_{i+n}
\vec{X}_i	0	$\vec{X}_{i+1} - \vec{X}_i$	$\vec{X}_{i+2} - \vec{X}_i$	$\vec{X}_{i+n} - \vec{X}_i$
\vec{X}_{i+1}	$\vec{X}_i - \vec{X}_{i+1}$	0	$\vec{X}_{i+2} - \vec{X}_{i+1}$	$\vec{X}_{i+n} - \vec{X}_{i+1}$
\vec{X}_{i+2}	$\vec{X}_i - \vec{X}_{i+2}$	$\vec{X}_i - \vec{X}_{i+1}$	0	$\vec{X}_{i+n} - \vec{X}_{i+2}$
⋮	⋮	⋮	⋮	0	...	⋮
\vec{X}_{i+n}	$\vec{X}_i - \vec{X}_{i+n}$	$\vec{X}_{i+1} - \vec{X}_{i+n}$	$\vec{X}_{i+2} - \vec{X}_{i+n}$	0

- Στην συνέχεια επαναφέρουμε τις τιμές που βρίσκονται κάτω της διαγωνίου σε ένα νέο σύνολο δεδομένων και τις επαναφέρουμε στην αρχική τους μορφή (μέτρο - γωνία). Με το πέρας της διαδικασίας αυτής έχουμε ένα καινούργιο Dataset το οποίο περιέχει όλους τους συνδυασμούς συσκευών που εμφανίστηκαν. Κατά την αντιστροφή μετατροπής των μιγαδικών είναι πιθανό να εμφανιστούν αρνητικές τιμές ρεύματος οι οποίες όμως δεν απορρίπτονται λόγω συμμετρικότητας του πίνακα DIFF2. Στο επόμενο βήμα απομακρύνουμε από τα δεδομένα μας τις πολύ χαμηλές (<0.2A) αλλά και τις πολύ υψηλές (>12A) ρεύματος. Στο σημείο αυτό έχουν αρχίσει να εμφανίζονται διακριτές πλέον υπογραφές φορτίου όπως φαίνεται και στο Σχήμα 13.



Σχήμα 13 : Δισδιάστατη απεικόνιση των δεδομένων μετά την πρώτη συσταδοποίηση

- Τέλος εκτελούμε ξανά την διαδικασία συσταδοποίησης στο νέο σύνολο δεδομένων με τον αλγόριθμο DBSCAN ο οποίος με βάση τις ιδιότητες που αναφέρθηκαν στο 3.2.2 και λαμβάνοντας υπ' όψιν μας το γεγονός πως στο νέο σύνολο δεδομένων οι υπογραφές φορτίου που αποτελούν συνδυασμό συσκευών και όχι αυτόνομες συσκευές, οι οποίες είναι πολύ λιγότερες και διάσπαρτες στον χώρο, ομαδοποιεί τα δεδομένα και κρατάει μόνο τις υπογραφές που προέρχονται από μεμονωμένες συσκευές. Έτσι μετά από την δημιουργία των νέων συστάδων και υπολογισμού των μέσων τιμών αυτών καταλήγουμε στην δημιουργία των υπογραφών φορτίου.

Έχουμε πλέον εξάγει τις υπογραφές φορτίου για τις συσκευές μας και πλέον μπορούμε να αναγνωρίσουμε με χρήση μεθόδων όπως τα Hidden Markov Models (HMM) ποιος συνδυασμός συσκευών είναι ενεργός ανά πασα στιγμή. Έτσι είμαστε σε θέση πλέον να εξάγουμε δεδομένα στατιστικών για την χρήση των συσκευών όπως τη χρονική διάρκεια λειτουργίας της κάθε συσκευής τις χρονικές στιγμές που αυτή είναι ενεργή και να παρέχουμε στο χρήστη ή τον

πάροχο αναλυτικά το καταναλωτικό του προφίλ, χωρίς να έχουμε παρέμβει σε κανένα σημείο του χώρου πέρα από τον κεντρικό ηλεκτρικό πίνακα.

3.4 Παραδοχές

Στο σημείο αυτό είναι απαραίτητο να αναφερθούν κάποιες παραδοχές σχετικά με τον τρόπο λειτουργίας του αλγορίθμου. Αρχικά ο αλγόριθμος αναπτύχθηκε με τη χρήση της γλώσσας προγραμματισμού Python, με την χρήση του editor 'Spyder', και διάφορων βιβλιοθηκών της Python όπως η sklearn, από την οποία υιοθετήθηκε η υλοποίηση του αλγορίθμου DBSCAN (Παράρτημα Α). Οι μετρήσεις οι οποίες χρησιμοποιήθηκαν είναι πραγματικές μετρήσεις από δυο οικίες σε Δάσκιο Βέροιας και Βέροια και η μετρητική διάταξη υλοποιήθηκε με σκοπό την υλοποίηση της [6]. Το σύστημα υλοποιήθηκε με σκοπό την αυτοματοποιημένη, χωρίς ανθρώπινη παρέμβαση, εξαγωγή των υπογραφών φορτίου των παραπάνω μετρήσεων. Επιπλέον ο αλγόριθμος που περιγράφεται βασίζεται στην ανάλυση σταθερής κατάστασης, πράγμα που καθιστά την εξαγωγή υπογραφών για συσκευές τύπου IV (συσκευές οι οποίες έχουν σταθερή κατανάλωση καθ' όλη την διάρκεια της μέτρησης, 2.3) αδύνατη. Σε πολλές από τις μετρήσεις μας συσκευές όπως ο υπολογιστής και η τηλεόραση, δεν παρουσιάζουν καμία μεταβολή στην κατανάλωση τους (εξαιτίας του ότι με την αρχικοποίηση των μετρήσεων είναι ήδη σε λειτουργία) γεγονός που τις καθιστά μη ανιχνεύσιμες. Τέλος όλες οι υπόλοιπες συσκευές εμπίπτουν είτε στην κατηγορία I είτε στην κατηγορία II (2.3) έχουν δηλαδή είτε πολλές διακριτές καταστάσεις είτε μόνο On και Off.

ΚΕΦΑΛΑΙΟ 4

ΑΠΟΤΕΛΕΣΜΑΤΑ

4.1 Πίνακες συσχέτισης

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα που εξήχθησαν μέσα από την εφαρμογή του αλγορίθμου. Αρχικά καθώς σημαντικό κομμάτι της εργασίας αυτής επικεντρώνεται στην χρήση των αρμονικών και των γωνιών τους με σκοπό την εξαγωγή έξυπνων υπογραφών φορτίου, ένα βασικό ερώτημα που γεννάται αποτελεί, το ποσοστό συσχέτισης που υπάρχει μεταξύ πλάτους και γωνίας των αρμονικών και κατά ποσό αυτό επηρεάζει τη διαδικασία συσταδοποίησης. Ως δείκτης συσχέτισης χρησιμοποιείται συνήθως ο συντελεστής γραμμικής συσχέτισης r του Pearson. Τα διαγώνια στοιχεία του πίνακα είναι ίσα με τη μονάδα διότι κάθε μεταβλητή συσχετίζεται απόλυτα με τον εαυτό της. Τα μη διαγώνια στοιχεία αντιστοιχούν στις τιμές του δείκτη συσχέτισης ανάμεσα στα ζευγάρια των μεταβλητών γραμμών και στηλών. Η ύπαρξη μιας ομάδας μεταβλητών που συσχετίζονται μεταξύ τους υψηλά είναι μια ένδειξη υπέρ της υπόθεσης ότι αυτές οι μεταβλητές μετρούν την ίδια σύνθετη ή λανθάνουσα μεταβλητή. Η μεταβλητή αυτή ονομάζεται παράγοντας (factor). Με τη μείωση του πλήθους μιας ομάδας συσχετιζόμενων μεταβλητών, ένας μικρός αριθμός παραγόντων εξηγεί τη μέγιστη δυνατή κοινή διακύμανση μεταξύ των μεταβλητών. Έτσι έχουμε την δυνατότητα επιλογής μεταξύ των βέλτιστων μεταβλητών με σκοπό την υλοποίηση της συσταδοποίησης. Ένα παράδειγμα ενός πίνακα συσχετίσεων είναι ο Πίνακας 7 ο οποίος παραθέτει τις συσχετίσεις μεταξύ των μεταβλητών (I50, Φ50, I150, Φ150, I250, Φ250) των μετρήσεων που υλοποιήθηκαν στο [6].

Αξίζει να σημειωθεί ότι ο πίνακας συσχετίσεων μπορεί να περιέχει τόσο αρνητικές όσο και θετικές τιμές από μηδέν έως ένα με τον αριθμό μηδέν να υποδηλώνει μηδαμινή συσχέτιση, το “θετικό” ένα απολυτή συσχέτιση και το “αρνητικό” ένα αντίστοιχα απολυτή διαφοροποίηση, η οποία μπορεί να μεταφραστεί και ως αντίστροφη συσχέτιση. Για τον λόγο αυτό και με σκοπό την ευκολότερη ανάγνωση των πινάκων στους πίνακες συσχέτισης που παρατίθενται επιλέχθηκε η λήψη της απολυτής τιμής σας δείκτη συσχέτισης.

Πίνακας 7 : Πίνακας συσχετίσεων για τον συνδυασμό “ΣΙΔΕΡΟ – ΜΑΤΙ – Η/Υ”

A/A	I50 [A]	Φ50 [deg]	I150 [A]	Φ150 [deg]	I250 [A]	Φ250 [deg]
I50	1	0.83	0.01	0.03	0.3	0.44
Φ50	0.83	1	0.02	0.04	0.52	0.56

I150	0.01	0.02	1	0.02	0.1	0.06
Φ150	0.03	0.04	0.02	1	0.01	0.12
I250	0.3	0.52	0.1	0.01	1	0.49
Φ250	0.44	0.56	0.06	0.12	0.49	1

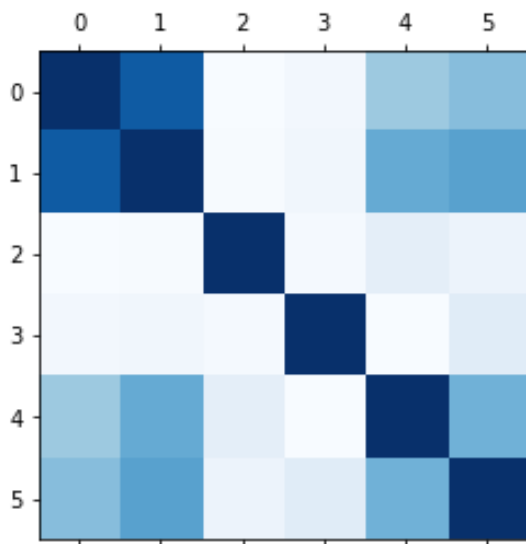
Όπως είναι εμφανές υπάρχει απολυτή συσχέτιση μεταξύ των ιδίων μεταβλητών ενώ επίσης έντονη συσχέτιση, δηλαδή πάνω από 80%, παρατηρείται μεταξύ των τιμών του μέτρου μεταξύ 1^{ης} αρμονικής και της γωνίας της. Από τον παραπάνω πίνακα μπορούν να εξαχθούν συμπεράσματα για το είδος των φορτίων που επιδρούν στην μέτρηση και κατά κύριο λόγο για τον τρόπο που αυτά επηρεάζονται από αντίστοιχα αρμονικά περιεχόμενα, εάν δηλαδή το φορτίο είναι γραμμικό ή όχι. Στην παραπάνω μέτρηση ο συνδυασμός “ ΣΙΔΕΡΟ – ΜΑΤΙ – Η/Υ ” είναι γνωστό ότι επηρεάζεται από την αρμονική παραμόρφωση που προσδίδει το “laptop”. Έτσι, όπως είναι αναμενόμενο οι συσχετίσεις μεταξύ των μέτρων της 1^{ης}, 3^{ης} και 5^{ης} αρμονικής είναι μικρή. Στον Πίνακα 8 παρουσιάζεται ο πίνακας συσχετίσεων για των συνδυασμό “ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – ΜΑΤΙ”, τα φορτία που οποίου δεν εμφανίζουν επαγωγική συμπεριφορά. Όπως είναι εμφανές οι συσχετίσεις τόσο των μέτρων όσο και των γωνιών είναι εμφανώς μεγαλύτερες από αυτές του Πίνακα 7.

Πίνακας 8 : Πίνακας συσχετίσεων για τον συνδυασμό “ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – ΜΑΤΙ”

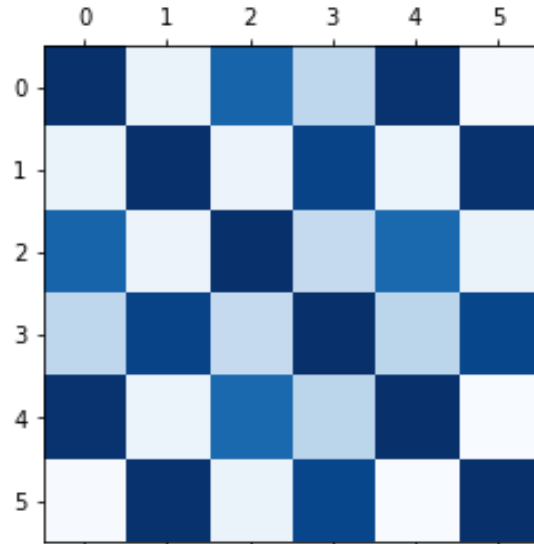
A/A	I50 [A]	Φ50 [deg]	I150 [A]	Φ150 [deg]	I250 [A]	Φ250 [deg]
I50	1	0.19	0.82	0.37	0.98	0.14
Φ50	0.19	1	0.18	0.93	0.18	0.99
I150	0.82	0.18	1	0.35	0.8	0.2
Φ150	0.37	0.93	0.35	1	0.37	0.92
I250	0.98	0.18	0.8	0.37	1	0.13
Φ250	0.14	0.99	0.2	0.92	0.13	1

Τέλος ένας έξυπνος και παραστατικός τρόπος αναπαράστασης ενός πίνακα συσχετίσεων είναι η γραφική αναπαράσταση αυτή με χρωματικούς πίνακες (colormaps), όπου τα πιο σκουρόχρωμα στοιχεία έχουν μεγάλη συσχέτιση μεταξύ τους ενώ η συσχέτιση φθίνει μαζί με

την φωτεινότητα του χρώματος καταλήγοντας στο μηδέν το οποίο αντιπροσωπεύεται από το λευκό. Το αντίστοιχα colormaps για τους Πίνακες 7 - 8 παρατίθενται στα Σχήματα 14 – 15. Το σύνολο των colormaps παρατεθείτε στο Παράρτημα Β.



Σχήμα 14 : Colormap για τον συνδυασμό
“SIDERO+MATI+LAPTOP”



Σχήμα 15 : Colormap για τον συνδυασμό
“PISTOLAKI + TOSTRIERA + MATI”

4.2 Εξαγωγή υπογραφών με χρήση παρεμβατικές μεθόδους

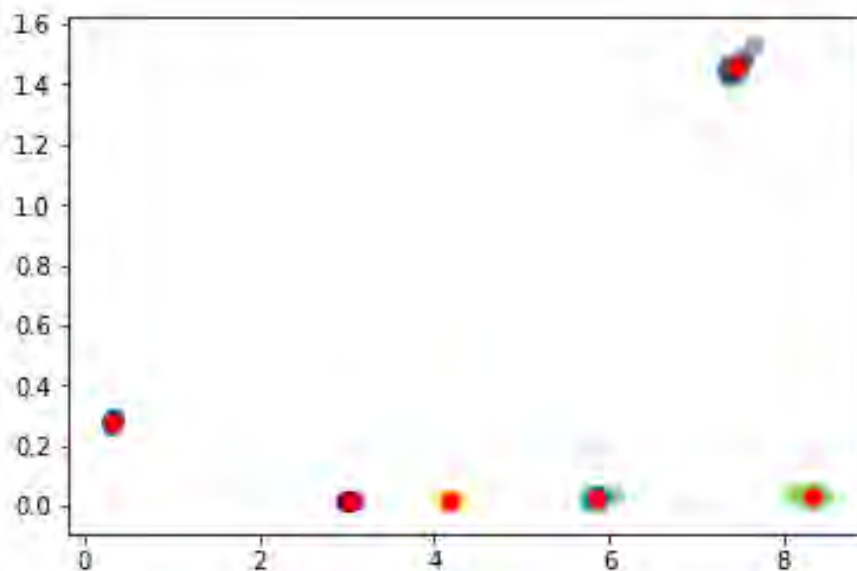
Για τον σκοπό εξαγωγής αποτελεσμάτων όπως αναφέρεται στο 3.2.3, στην [6], εκτός από την καταγραφή των συνολικών τιμών του ρεύματος και των τιμών των αρμονικών ακολούθησε καταγραφή των μεμονωμένων τιμών αυτών για κάθε συσκευή με σκοπό την επαλήθευση των αποτελεσμάτων. Έτσι παράλληλα με την διαδικασία μη-παρεμβατικής παρακολούθησης του φορτίου υλοποιήθηκε μια σειρά συναρτήσεων οι οποίες έχουν ως σκοπό την μαθηματική εξαγωγή των υπογραφών φορτίου, από τις μεμονωμένες μετρήσεις, με σκοπό την σύγκριση των αποτελεσμάτων. Η ροή δεδομένων για τον παρεμβατικό αλγόριθμο είναι η εξής :

- Ανάγνωση των μεμονωμένων μετρήσεων.
- Απόρριψη τιμών για τις οποίες η συσκευή είναι ανενεργή (Δηλ. $I_{50} < 0.2A$).
- Υπολογισμός μέσου ορού των τιμών που επέμειναν.

- Αποθήκευση δεδομένων για κάθε μεμονωμένη μέτρηση.
- Υπολογισμός μέσου ορού ανά συσκευή.

Μετά την εκτέλεση του αλγορίθμου και δεδομένου ότι κανένα από τα φορτία που αναλύουμε δεν είναι τύπου III, είμαστε σε θέση να γνωρίζουμε τον τρόπο συμπεριφοράς του συνόλου των συσκευών. Οι υπογραφές μπορούν εύκολα πλέον να αναπαρασταθούν γραφικά σε άξονες Χ - Υ όπου ο άξονας Χ αποτελείται από τις τιμές πλάτους της 1^{ης} αρμονικής και ο Υ από τα πλάτη της 3^{ης} αρμονικής. Η αναπαράσταση αυτή γίνεται στο Σχήμα 16 όπου διαφορετικές συσκευές παρουσιάζονται με διαφορετικά χρώματα και με κόκκινο χρώμα παρουσιάζονται οι μέσοι όροι κάθε συσκευής. Επιπλέον, παρατίθενται στον Πίνακα 9 οι μαθηματικές τιμές των υπογραφών ανά συσκευή. Οι υπογραφές ακολουθούν τη μορφή :

$$LS_i^B = [(I_{i1j}^B | \cos\varphi1), (I_{i3j}^B | \cos\varphi3), (I_{i5j}^B | \cos\varphi5)] \quad (5)$$



Σχήμα 16 : Γραφική αναπαράσταση των μαθηματικά υπολογισμένων υπογραφών φορτίου και των μέσων τιμών τους.

Όπως αναφέρεται στο 3.4 οι συσκευές “υπολογιστής” και “τηλεόραση” είναι τύπου IV και δεν μεταβάλουν την κατάσταση τους καθ’ όλη την διάρκεια των μετρήσεων έτσι δεν είναι εύκολη

η διάκριση μεταξύ τους. Για το λόγο αυτό επιλέχθηκε η αναπαράσταση τους σαν μια συσκευή ακόμα και στο κομμάτι των μαθηματικών – παρεμβατικών υπολογισμών.

Πίνακας 9 : αριθμητικές τιμές μαθηματικά υπολογισμένων υπογραφών φορτίου

A/A	I50 [A]	Φ50 [deg]	I150 [A]	Φ150 [deg]	I250 [A]	Φ250 [deg]
Πιστολάκι	5.973	45.341	0.031	112.137	0.085	104.060
Σκούπα	7.422	23.236	1.470	128.146	0.100	136.156
Τοστιέρα	4.135	40.926	0.017	7.552	0.044	95.915
Μάτι κουζίνας	3.016	8.934	0.006	39.092	0.039	146.050
Σίδερο	8.327	18.940	0.012	4.247	0.108	-44.883
Υπολογιστής*	0.327	-14.252	0.273	-130.045	0.229	12.916
Τηλεόραση*	0.327	-14.252	0.273	-130.045	0.229	12.916

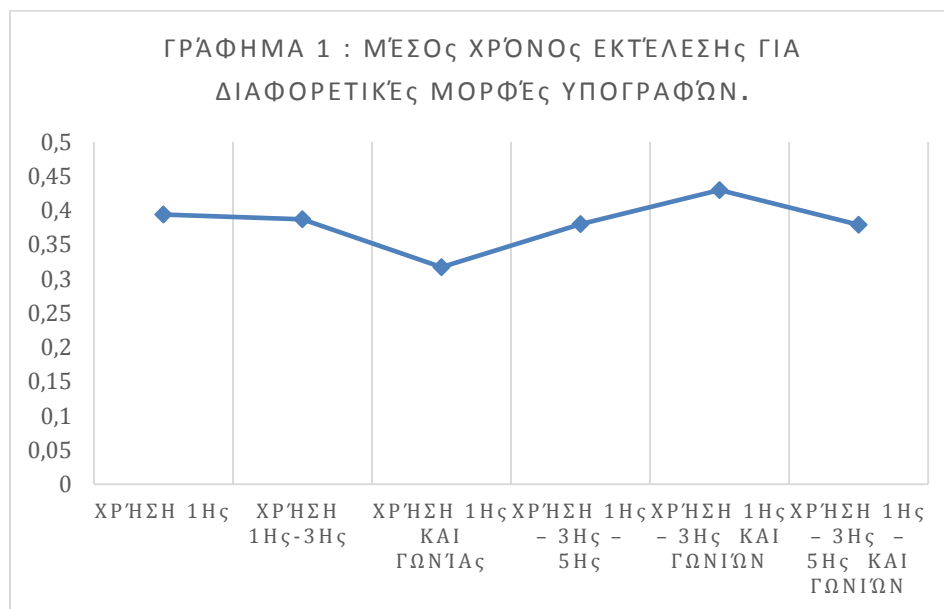
4.3 Χρόνος εκτέλεσης

Επίσης σημαντικός παράγοντας για τον καθορισμό της απόδοσης οποιουδήποτε συστήματος μη-παρεμβατικής αναγνώρισης είναι ο χρόνος που απαιτείται από το εκάστοτε πρόγραμμα για την εξαγωγή των υπογραφών φορτίου. Ο χρόνος αυτός μπορεί να ποικίλει ανάλογα με τον τύπο των υπογραφών που επιλέγεται για χρήση όπως επίσης και από την μεθοδολογία που ακολουθείται για την εξαγωγή των υπογραφών (υπολογιστικές διεργασίες). Κρίθηκε λοιπόν σκόπιμο στο κεφάλαιο αυτό να παρουσιαστούν οι χρόνοι που χρειάστηκε ο προτεινόμενος αλγόριθμος. Στον Πίνακα 10 παρουσιάζονται οι χρόνοι τρεξίματος αναλυτικά για επτά σύνολα δεδομένων.

Πίνακας 10 : Χρόνος τρεξίματος σε sec για τα επτά πρώτα σύνολα δεδομένων.

Σύνολο δεδομένων	Πλήρης υπογραφή	Χρήση 1 ^{ης} – 3 ^{ης} και γωνιών	Χρήση 1 ^{ης} και γωνίας	Χρήση 1 ^{ης} – 3 ^{ης} – 5 ^{ης}	Χρήση 1 ^{ης} - 3 ^{ης}	Χρήση 1 ^{ης}
1	0.341	0.4905	0.374	0.375	0.420	0.486
2	0.534	0.420	0.417	0.342	0.452	0.315
3	0.433	0.411	0.279	0.284	0.435	0.278
4	0.334	0.480	0.312	0.314	0.325	0.429
5	0.279	0.455	0.341	0.433	0.416	0.399
6	0.298	0.382	0.291	0.459	0.355	0.396
7	0.436	0.369	0.362	0.453	0.306	0.452
ΜΟ	0.379	0.4296	0.317	0.38	0.387	0.394

Όπως είναι εμφανές ο αλγόριθμος λειτουργεί πιο αποδοτικά αναφορικά με τους χώρους δύο η τριών διαστάσεων και καθώς οι διαστάσεις αυξάνονται ο απαιτούμενος χρόνος για την εξαγωγή των υπογραφών επίσης αυξάνεται χωρίς ωστόσο να αγγίζει απαγορευτικές τιμές, όπως φαίνεται στο Σχήμα 17.

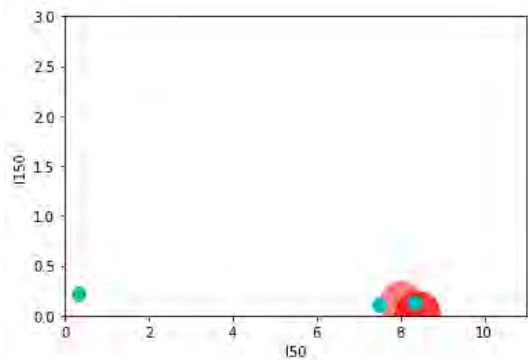


Σχήμα 17 : Μέσος χρόνος εκτέλεσης για διαφορετικές μορφές υπογραφών

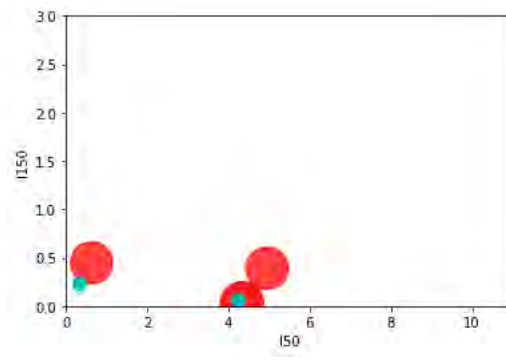
4.4 Εξαχθείσες υπογραφές φορτίου

Στην παράγραφο αυτή παρουσιάζονται οι υπογραφές που εξάγονται με τη χρήση του προτεινομένου αλγορίθμου καθώς σαν μέτρο σύγκρισης για την απόδοση των υπογραφών λαμβάνονται οι μαθηματικά υπολογισμένες υπογραφές φορτίου. Για την ευκολότερη αναπαράσταση των υπογραφών και την ασφαλέστερη εξαγωγή συμπερασμάτων κατά την εφαρμογή του αλγορίθμου γίνεται γραφική αναπαράσταση σε άξονες X – Y (όπου X αντιστοιχεί στο μέτρο της 1^{ης} αρμονικής και Y στο μέτρο της 3^{ης}) τόσο των υπογραφών που εξήχθησαν από την παρεμβατική όσο και από την μη παρεμβατική μέθοδο. Λαμβάνοντας υπ' όψιν την ελάχιστη διαφορά που χρειάζεται στο χρόνο εκτέλεσης για την πλήρη υπογραφή σε σχέση με τις άλλες επιλογές, τα Σχήματα 18-21 που παρουσιάζονται παρακάτω είναι γραφικές απεικονίσεις των ανεπτυγμένων υπογραφών (πλήρων) δηλαδή :

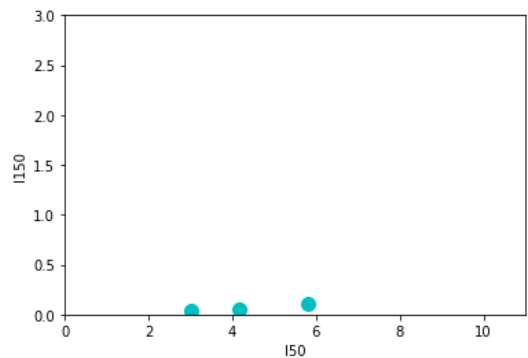
$$LS_i^B = [(I_{i1j}^B | \cos\varphi_1), (I_{i3j}^B | \cos\varphi_3), (I_{i5j}^B | \cos\varphi_5)] \quad (5)$$



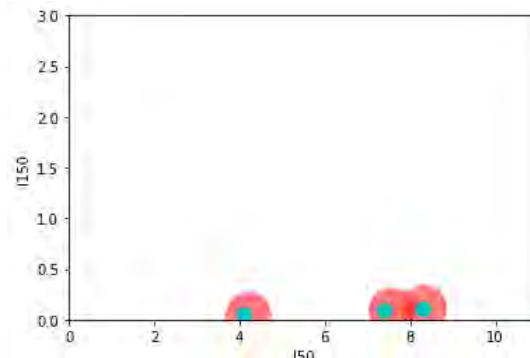
Σχήμα 18 : Τηλεόραση – σκούπα – σίδηρο.



Σχήμα 19 : Τηλεόραση – τοστιέρα – Η/Υ.



Σχήμα 20 : Πιστολάκι – τοστιέρα – μάτι.



Σχήμα 21 : Τοστιέρα – σκούπα – σίδηρο .

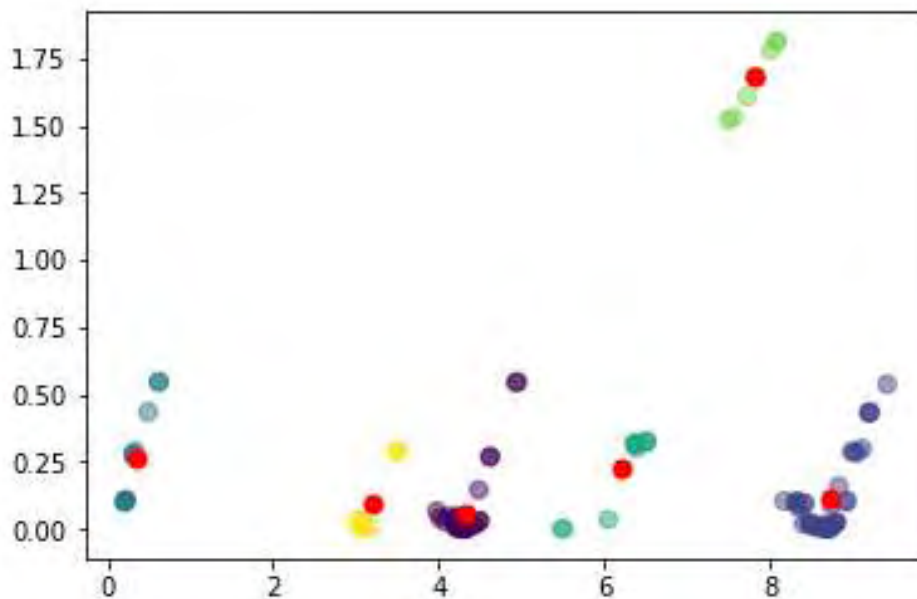
Σχήματα 18-21 : Γραφική αναπαράσταση των αποτελεσμάτων του αλγορίθμου.

Με κόκκινο εμφανίζονται οι μη παρεμβατικά υπολογισμένες υπογραφές φορτίου ενώ με μπλε οι παρεμβατικά υπολογισμένες. Όπως φαίνεται και στα Σχήματα 18 - 21 η μη παρεμβατική μέθοδος οδηγεί σε αρκετά αποκλίνοντα αποτελέσματα σε σχέση με την παρεμβατική που μπορούν να διακριθούν στις ακόλουθες κατηγορίες :

- Εξαγωγή μηδενικών υπογραφών (Σχήμα 20).
- Εξαγωγή λιγότερων υπογραφών από τις αναμενόμενες (Σχήμα 18).
- Εξαγωγή ακριβή αριθμού υπογραφών (Σχήμα 21).
- Εξαγωγή περισσότερων υπογραφών (Σχήμα 19).

Στο σημείο αυτό έρχεται να βοηθήσει τη διαδικασία η δεύτερη ιδιότητα του αλγορίθμου DBSCAN. Όπως αναφέρθηκε στο 2.6.2 βασικό χαρακτηριστικό του αλγορίθμου αποτελεί η δυνατότητα απομάκρυνσης ακραίων τιμών. Έτσι εφόσον η πλειοψηφία των εξαγόμενων υπογραφών συμβαδίζει με τις πραγματικές τιμές αυτών η αναγνώριση θα αποδίδει θετικά αποτελέσματα και δεν θα επηρεάζεται από λανθασμένες υπογραφές που μπορεί να παράγει το αρχικό τρέξιμο. Αναλυτικά τα γραφήματα αναπαράστασης των αποτελεσμάτων του αλγορίθμου για όλα τα σύνολα δεδομένων παρατίθενται στο Παράρτημα Γ.

Τέλος οι υπογραφές που εξήχθησαν ομαδοποιήθηκαν σε ένα νέο σύνολο δεδομένων του οποίου η αναπαράσταση ακολουθεί στο Σχήμα 22. Διαφορετικές συσκευές απεικονίζονται με διαφορετικά χρώματα και με κόκκινο απεικονίζονται οι μέσες τιμές κάθε συσκευής, δηλαδή οι τελικές υπογραφές, οι οποίες είναι ίδιες σε αριθμό με αυτές που υπολογίστηκαν με τη χρήση της μη παρεμβατικής μεθόδου.



Σχήμα 22 : Γραφική αναπαράσταση μη παρεμβατικά υπολογισμένων υπογραφών φορτίου

4.5 Αποτελεσματικότητα διαφόρων δομών υπογραφής

Επιπλέον για την περαιτέρω διερεύνηση της βέλτιστης δομής των υπογραφών έγινε μια προσπάθεια συσχέτισης τους με τις μη παρεμβατικά υπολογισμένες υπογραφές φορτίου. Ως μέτρο συσχέτισης επιλέχθηκε η ποσοστιαία απόκλιση μεταξύ των δύο. Έτσι έγινε μια προσπάθεια διερεύνησης του ποσοστού των υπογραφών που είχαν απόκλιση από τις πραγματικές μικρότερη του $\pm 10\%$. Τα αποτελέσματα παρουσιάζονται στον Πίνακα 11. Στη στήλη “Τύπος υπογραφής” παρουσιάζονται οι διαφορετικές συνιστώσες, των συνόλων δεδομένων, οι οποίες παρέχονται στον αλγόριθμο για τη διαδικασία συσταδοποίησης (διαφορετικά μοντέλα υπογραφών), στη στήλη “ Συνολικές υπογραφές για το σύνολο δεδομένων ” καταγράφεται ο συνολικός αριθμός των υπογραφών που παράγει ο αλγόριθμος, στη στήλη “ Υπογραφές εντός του 10% των πραγματικών ” παρουσιάζονται οι υπογραφές οι οποίες κυμαίνονται σε ακτίνα $\pm 10\%$ από τις πραγματικές και τέλος στην στήλη “Αναλογία” παρουσιάζεται η αναλογία αυτών, το ποσοστό δηλαδή των υπογραφών που υπολογιστήκαν μη παρεμβατικά και έχουν απόκλιση από τις πραγματικές λιγότερο από 10%.

Πίνακας 11 : Ποσοστά ευστοχίας της μεθόδου.

Τύπος υπογραφής	Συνολικές υπογραφές για το σύνολο δεδομένων	Υπογραφές εντός του 10% των πραγματικών	Αναλογία
Μετρό 1 ^{ης}	84	52	0.61
Μετρό 1 ^{ης} – 3 ^{ης}	252	198	0.78
Μετρό και γωνία 1 ^{ης}	252	198	0.78
Μετρό 1 ^{ης} – 3 ^{ης} – 5 ^{ης}	84	51	0.60
Μετρό 1 ^{ης} – 3 ^{ης} – 5 ^{ης} και γωνία 1 ^{ης}	257	198	0.77
Μετρό και γωνία 1 ^{ης} – 3 ^{ης}	173	140	0.80
Μετρό και γωνία 1 ^{ης} – 3 ^{ης} και μετρό 5 ^{ης}	173	140	0.80
Μετρό και γωνία 1 ^{ης} – 3 ^{ης} – 5 ^{ης}	119	105	0.882

Τέλος θεωρήθηκε σημαντικό να υπολογιστεί ο συντελεστής συσχέτισης μεταξύ των κέντρων των κλάσεων που υπολογιστήκαν από τις δυο μεθόδους. Για τον σκοπό αυτό παρατίθενται οι Πίνακες 12-13 όπου παρουσιάζονται τόσο οι αριθμητικές τιμές των κέντρων (δηλαδή της κάθε συσκευής) των δυο μεθόδων όσο και η μεταξύ τους διαίρεση η οποία δίνει το ποσοστό συσχέτισης [22].

Πίνακας 12 : Αριθμητικές τιμές μη παρεμβατικά υπολογισμένων υπογραφών φορτίου.

A/A	I50 [A]	I150 [A]	I250 [A]
Πιστολάκι	6.202	0.226	0.106
Σκούπα	7.820	1.677	0.142
Τοστιέρα	4.322	0.112	0.105
Μάτι κουζίνας	3.203	0.112	0.105
Σίδερο	8.709	0.260	0.142
Υπολογιστής*	0.339	0.058	0.062
Τηλεόραση*	0.339	0.058	0.062

Πίνακας 13 : Ποσοστά συσχέτισης υπολογισμένων υπογραφών με πραγματικές τιμές.

A/A	I50 [A]	I150 [A]	I250 [A]
Πιστολάκι	0.963	0.137	0.802
Σκούπα	0.950	0.877	0.704
Τοστιέρα	0.956	0.152	0.419
Μάτι κουζίνας	0.940	0.054	0.371
Σίδερο	0.956	0.046	0.761
Υπολογιστής*	0.964	4.707	3.694
Τηλεόραση*	0.964	4.707	3.694

ΚΕΦΑΛΑΙΟ 5

ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΣΘΗΚΕΣ

Σε αυτή τη διπλωματική εργασία, προτείνεται και υλοποιείται ένα μη εποπτευόμενο πλαίσιο NILM που βασίζεται στην αναγνώριση συμβάντων, χρησιμοποιώντας τόσο σύνολα δεδομένων και αλλά και εργαστηριακό πείραμα. Βασικό χαρακτηριστικό τόσο των μετρήσεων όσο και των πειραμάτων είναι η χρήση των αρμονικών τιμών ρεύματος για την εξαγωγή αποτελεσμάτων. Επίσης βασικό πυλώνα των πειραμάτων αποτελεί και η επιστήμη της μηχανικής μάθησης και πιο συγκεκριμένα ο αλγόριθμος DBSCAN. Τέλος κύρια παραδοχή της παραπάνω διαδικασίας είναι το γεγονός πως η χρήση των αρμονικών παρέχει σημαντική, επιπλέον, πληροφορία για την παραγωγή βέλτιστων υπογραφών φορτίου.

Αναγνωρίζονται βέβαια περιορισμοί της προτεινόμενης προσέγγισης. Η προσέγγιση αυτή εξετάζει μόνο συσκευές με ενεργή και ανενεργή κατάσταση (Τύπος I) και συσκευές πολλαπλών καταστάσεων (Τύπος II). Εκτός αυτού, δεδομένου ότι αυτή η μέθοδος εξάγει χαρακτηριστικά από παρακείμενες σταθερές καταστάσεις, δεν μπορεί να διακρίνει μεταβολές συσκευών χωρίς σταθερές καταστάσεις.

Στο μέλλον, οι προαναφερόμενοι περιορισμοί θα πρέπει να εξεταστούν σοβαρά. Για παράδειγμα, ορισμένες συσκευές είναι πολύπλοκες και μεταβάλλουν την κατάσταση τους κατά τη διάρκεια της λειτουργίας, ενώ το NILM που βασίζεται σε ανίχνευση καταστάσεων χρησιμοποιείται καλύτερα για την αναγνώριση γεγονότων με παρακείμενες σταθερές καταστάσεις. Στην περίπτωση αυτή, οι μέθοδοι χωρίς την ανίχνευση γεγονότων μπορεί να λειτουργήσουν καλύτερα, όπως βαθιά νευρωνικά δίκτυα. Επιπλέον, διαφορετικοί αλγόριθμοι που βασίζονται σε διαφορετικά χαρακτηριστικά μπορεί να είναι αποτελεσματικοί για ορισμένους τύπους συσκευών. Ως εκ τούτου, η ενσωμάτωση συμπληρωματικών μοντέλων NILM είναι απαραίτητη και εφικτή

BIBΛΙΟΓΡΑΦΙΑ

- [1] Ahmed Zoha, Alexander Gluhak, Muhammad Ali Imran and Sutharshan Rajasegarar: “Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey”, (2012)
- [2] Ismail Hakkı ÇAVDAR and Vahid FARYAD: “New Design of a Supervised Energy Disaggregation Model Based on the Deep Neural Network for a Smart Grid”, Energies (2019).
- [3] George W. Hart: “Nonintrusive Appliance Load Monitoring”, Proceedings of the IEEE, (1992).
- [4] Albert Fiol: “Algorithms for Energy Disaggregation”, (2016)
- [5] Marchiori A., Hakkarinen D., Han Q., Earle L.: “Circuit-level load monitoring for household energy management.”, IEEE Pervas. Comput., (2011).
- [6] Aggelos S. Bouhouras, Paschalis A. Gkaidatzis, Evangelos Panagiotou, Nikolaos Poulakis, Georgios C. Christoforidis: “A NILM algorithm with enhanced disaggregation scheme under harmonic current vectors”, (2018).
- [7] Liang J., Ng S.K.K., Kendall G., Cheng J.W.M.: “Load signature study Part I: Basic concept, structure, and methodology”, IEEE Trans. Power Del. (2010).
- [8] Schoofs A., Guerrieri A., Delaney D., O’Hare G., Ruzzelli A.: “Automated Electricity Data Annotation Using Wireless Sensor Networks.”, IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks, Boston, MA, USA, (21–25 June 2010).
- [9] Rowe A., Berges M., Rajkumar R.: “Contactless Sensing of Appliance State Transitions through Variations in Electromagnetic Fields.”, ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building, Zurich, Switzerland, (3–5 November 2010).
- [10] Giri S., Berges M.: “A Study on the Feasibility of Automated Data Labeling and Training Using an EMF Sensor in NILM Platforms.”, International EG-ICE Workshop on Intelligent Computing, Munich, Germany, (4–6 June 2012).
- [11] Kolter J.Z., Johnson M.J.: “REDD: A Public Data Set for Energy Disaggregation Research.” SustKDD Workshop on Data Mining Applications in Sustainability, San Diego, CA, USA, (2011).

- [12] Aggelos S. Bouhouras 1, Paschalis A. Gkaidatzis 2, Konstantinos C. Chatzisavvas, Evangelos Panagiotou 1, Nikolaos Poulakis 1 and Georgios C. Christoforidis: "Load Signature Formulation for Non-Intrusive Load Monitoring Based on Current Measurements", *Energies* (2017).
- [13] C. Laughman, Kwangduk Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, P. Armstrong: "Power signature analysis", *IEEE Power Energy Mag.* 1, (2003).
- [14] Y. Du, L. Du, B. Lu, R. Harley, T. Habetler: "A review of identification and monitoring methods for electric loads in commercial and residential buildings", *IEEE Energy Conversion Congress and Exposition, ECCE 2010*, (2010).
- [15] Pedro Domingos: "Tapping into the "folk knowledge" needed to advance machine learning applications.", *Communications of the ACM*, (2012).
- [16] Schmidhuber J.: "Deep Learning in Neural Networks: An Overview", *Neural Networks* 61, (2015).
- [17] H.H. Chang: "Load identification of non-intrusive load-monitoring system in smart home", *WSEAS Trans. Syst.* 9, (2010).
- [18] Yasser El-sonbaty: "An efficient density-based clustering algorithm for large databases", *Conference Paper*, Source: *IEEE Xplore*, (2004)
- [19] Nizar Grira, Michel Crucianu, Nozha Boujemaa: "Unsupervised and Semi-supervised Clustering: A Brief Survey", (2005)
- [20] Zhuang Zheng, Hainan Chen and Xiaowei Luo: "A Supervised Event-Based Non-Intrusive Load Monitoring for Non-Linear Appliances", *Sustainability Multidisciplinary Open Access Journal*, (2018).
- [21] Ali. Tariq, Asghar. Sohail: "Critical analysis of DBSCAN variations", *International Conference on Information and Emerging Technologies (ICIET)*, (2010).
- [22] Valeria Amenta, Giuseppe Marco Tina: "Load Demand Disaggregation based on Simple Load Signature and User's Feedback", *7th International Conference on Sustainability in Energy and Buildings*, (2015)

ΠΑΡΑΡΤΗΜΑ Α

ΚΩΔΙΚΑΣ

Η ΥΛΟΠΟΙΗΣΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ DBSCAN ΑΠΟ ΤΗΝ ΒΙΒΛΙΟΘΗΚΗ SKITLEARN

```
def dbscan (X, eps=0.5, min_samples=5, metric='minkowski', metric_params=None,
            algorithm='auto', leaf_size=30, p=2, sample_weight=None,
            n_jobs=None):

    if not eps > 0.0:
        raise ValueError ("eps must be positive.")

    X = check_array (X, accept_sparse='csr')
    if sample_weight is not None:
        sample_weight = np. asarray(sample_weight)
        check_consistent_length (X, sample_weight)

    # Calculate neighborhood for all samples. This leaves the original point
    # in, which needs to be considered later (i.e. point i is in the
    # neighborhood of point i. While True, its useless information)
    if metric == 'precomputed' and sparse. issparse(X):
        neighborhoods = np. empty (X. shape [0], dtype=object)
        X.sum_duplicates () # XXX: modifies X's internals in-place

    # set the diagonal to explicit values, as a point is its own neighbor
    with warnings. catch_warnings ():
        warnings. simplefilter ('ignore', sparse. SparseEfficiencyWarning)
        X.setdiag (X. diagonal ()) # XXX: modifies X's internals in-place

    X_mask = X.data <= eps
    masked_indices = X. indices. astype (np. intp, copy=False) [X_mask]
    masked_indptr = np. concatenate ([[0], np. cumsum(X_mask)])
    masked_indptr = masked_indptr [X. indptr [1: -1]]

    # split into rows
    neighborhoods [:] = np. split (masked_indices, masked_indptr)
else:
    neighbors_model = NearestNeighbors (radius=eps, algorithm=algorithm,
                                       leaf_size=leaf_size,
```

```

        metric=metric,
        metric_params=metric_params, p=p,
        n_jobs=n_jobs)
neighbors_model.fit(X)
# This has worst case O(n^2) memory complexity
neighborhoods = neighbors_model.radius_neighbors(X, eps,
        return_distance=False)

if sample_weight is None:
    n_neighbors = np.array([len(neighbors)
        for neighbors in neighborhoods])
else:
    n_neighbors = np.array([np.sum(sample_weight[neighbors])
        for neighbors in neighborhoods])

# Initially, all samples are noise.
labels = np.full(X.shape[0], -1, dtype=np.intp)

# A list of all core samples found.
core_samples = np.asarray(n_neighbors >= min_samples, dtype=np.uint8)
dbscan_inner(core_samples, neighborhoods, labels)
return np.where(core_samples)[0], labels

class DBSCAN (BaseEstimator, ClusterMixin):
    def __init__(self, eps=0.5, min_samples=5, metric='euclidean',
        metric_params=None, algorithm='auto', leaf_size=30, p=None,
        n_jobs=None):
        self.eps = eps
        self.min_samples = min_samples
        self.metric = metric
        self.metric_params = metric_params
        self.algorithm = algorithm
        self.leaf_size = leaf_size
        self.p = p
        self.n_jobs = n_jobs

    def fit(self, X, y=None, sample_weight=None):
        X = check_array(X, accept_sparse='csr')
        clust = dbscan(X, sample_weight=sample_weight,

```

```

        **self.get_params ()
self. core_sample_indices_, self. labels_ = clust
if len (self. core_sample_indices_):
    # fix for scipy sparse indexing issue
    self. components_ = X [self. core_sample_indices_]. copy ()
else:
    # no core samples
    self. components_ = np. empty ((0, X. shape [1]))
return self

```

```

def fit_predict (self, X, y=None, sample_weight=None):

```

```

    self.fit (X, sample_weight=sample_weight)
    return self.labels_

```

ΧΡΗΣΗ ΤΟΥ DBSCAN ΣΤΟΝ ΑΛΓΟΡΙΘΜΟ ΠΟΥ ΑΝΑΠΤΗΧΘΗΚΕ

```

def db(dataset):

```

```

    std1=dataset.std ()
    std1=std1[1]
    mean1=dataset.mean()
    mean1=mean1[1]

```

```

    e= abs(std1/2)
    m = (n/100) + 1

```

```

    f1 = dataset. iloc [:1]
    f2 = dataset. iloc [:2]
    f3 = dataset. iloc [:4]
    f4 = dataset. iloc [:5]
    f5 = dataset. iloc [:7]
    f6 = dataset. iloc [:8]

```

```

    X = np. array (list (zip (f1, f2, f3, f4, f5, f6)))

```

```
dbscan = DBSCAN (eps=e, min_samples = m)
clusters = dbscan.fit_predict(X)
Z = np. amax(clusters)+1
Y = clusters

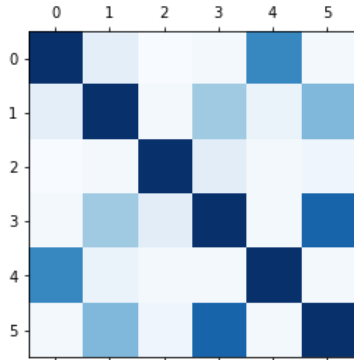
Z=np.arange(0, dataset. shape [0], 1)
X = np. array (list (zip (Z, f1, f2, Y)))
dataset = pd. DataFrame ({'I50':X [::1],'Φ50': f2,'Cluster':X [::3],'I150': f3,'Φ150': f4,'I250': f5,'Φ250': f6})
dataset = dataset [dataset. Cluster >= 0]

dataset = dataset. sort_values(by=['Cluster'])
dataset = dataset. reset_index ()
dataset = dataset. drop(['index'], axis=1)
return(dataset)
```

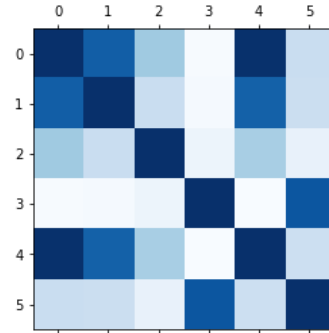
ΠΑΡΑΡΤΗΜΑ Β

ΓΡΑΦΙΚΕΣ ΠΑΡΑΣΤΑΣΕΙΣ ΠΙΝΑΚΩΝ ΣΥΣΧΕΤΗΣΗΣ

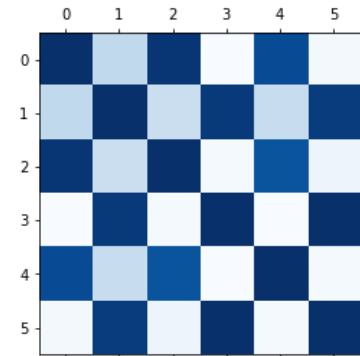
ΤΗΛΕΟΡΑΣΗ – ΜΑΤΙ – Η/Υ



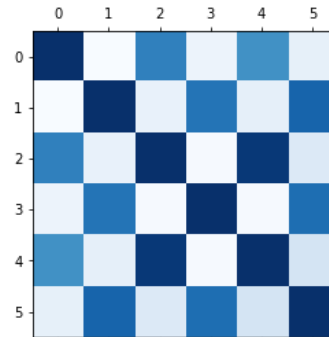
ΤΗΛΕΟΡΑΣΗ – ΜΑΤΙ – ΣΙΔΕΡΟ



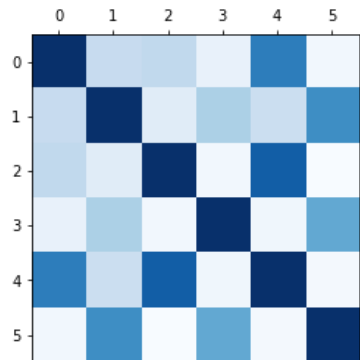
ΤΗΛΕΟΡΑΣΗ – ΜΑΤΙ – ΣΚΟΥΠΑ



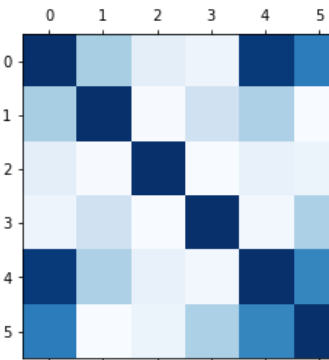
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – Η/Υ



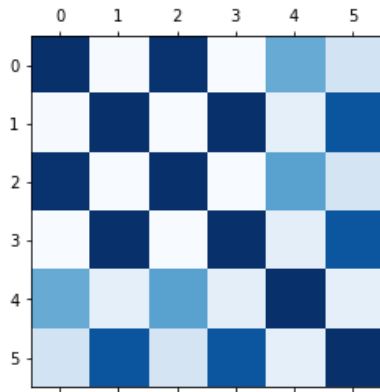
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – ΜΑΤΙ



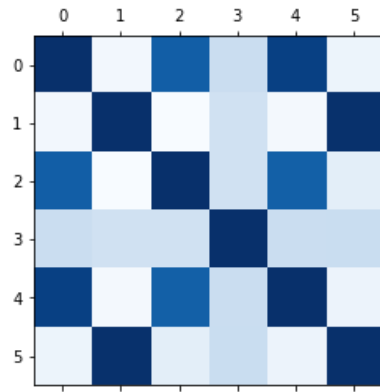
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – ΣΙΔΕΡΟ



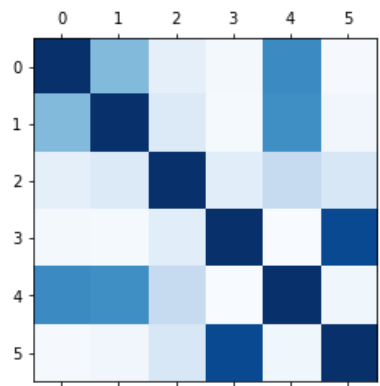
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – ΣΚΟΥΠΑ



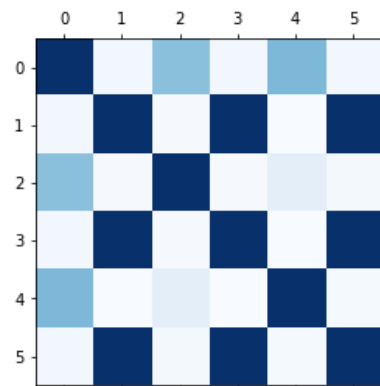
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ



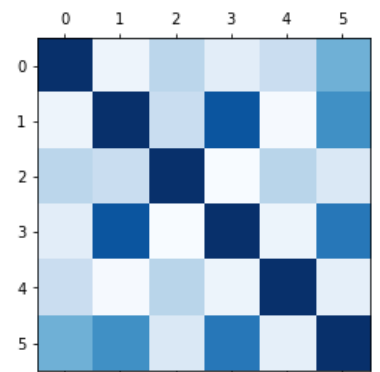
ΤΗΛΕΟΡΑΣΗ – ΣΙΔΕΡΟ – Η/Υ



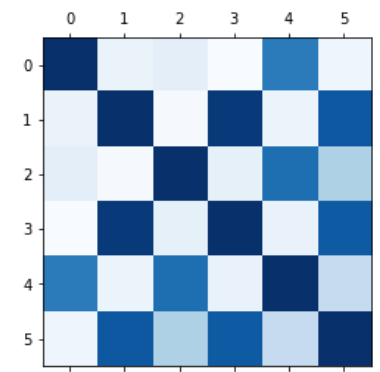
ΤΗΛΕΟΡΑΣΗ – ΣΚΟΥΠΑ – Η/Υ



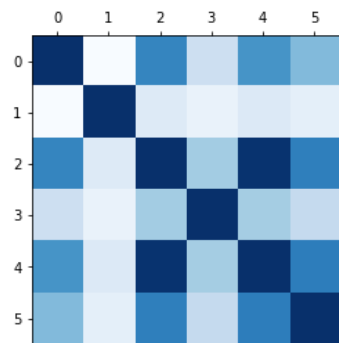
ΤΗΛΕΟΡΑΣΗ – ΣΚΟΥΠΑ – ΣΙΔΕΡΟ



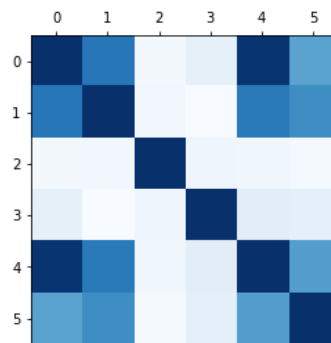
ΤΗΛΕΟΡΑΣΗ – ΤΟΣΤΙΕΡΑ – Η/Υ



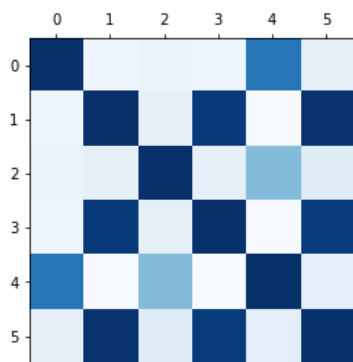
ΤΗΛΕΟΡΑΣΗ – ΤΟΣΤΙΕΡΑ – ΜΑΤΙ



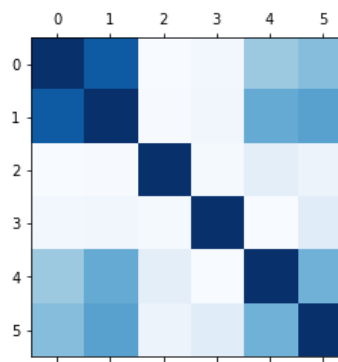
ΤΗΛΕΟΡΑΣΗ – ΤΟΣΤΙΕΡΑ – ΣΙΔΕΡΟ



ΤΗΛΕΟΡΑΣΗ – ΤΟΣΤΙΕΡΑ – ΣΚΟΥΠΑ

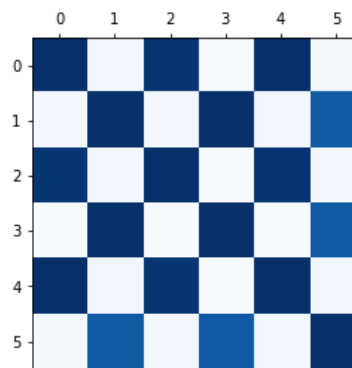


ΜΑΤΙ – ΣΙΔΕΡΟ – Η/Υ



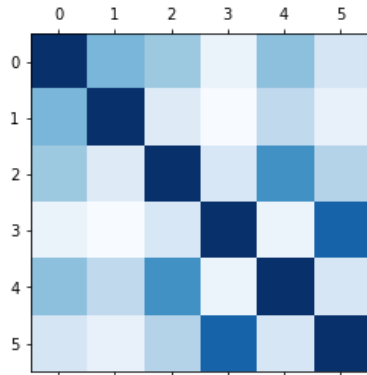
ΜΑΤΙ – ΣΚΟΥΠΑ – Η/Υ

ΜΑΤΙ – ΣΚΟΥΠΑ – ΣΙΔΕΡΟ

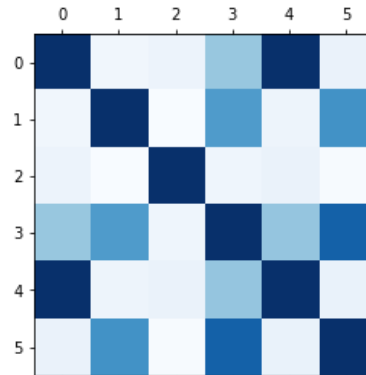


ΔΕΝ ΕΠΕΞΕΡΓΑΣΤΗΚΕ ΛΟΓΟ ΚΑΚΗΣ
ΠΟΙΟΤΗΤΑΣ ΤΩΝ ΕΓΓΡΑΦΩΝ.

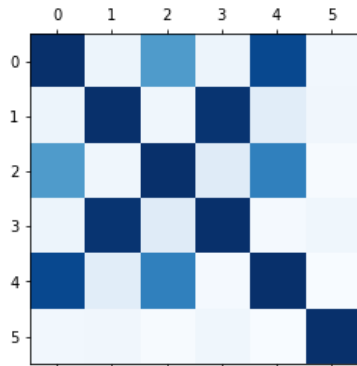
ΠΙΣΤΟΛΑΚΙ – ΜΑΤΙ – Η/Υ



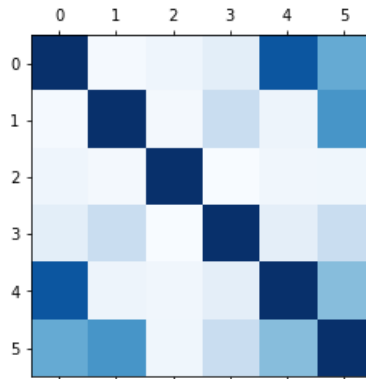
ΠΙΣΤΟΛΑΚΙ – ΜΑΤΙ – ΣΙΔΕΡΟ



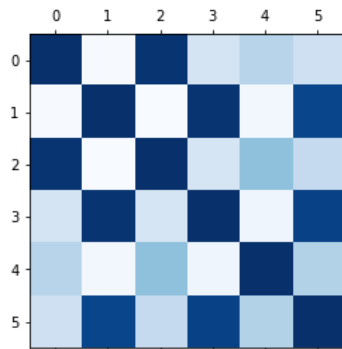
ΠΙΣΤΟΛΑΚΙ – ΜΑΤΙ – ΣΚΟΥΠΑ



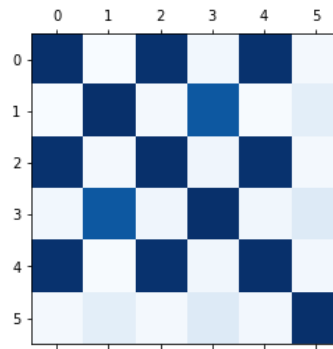
ΠΙΣΤΟΛΑΚΙ – ΣΙΔΕΡΟ – Η/Υ



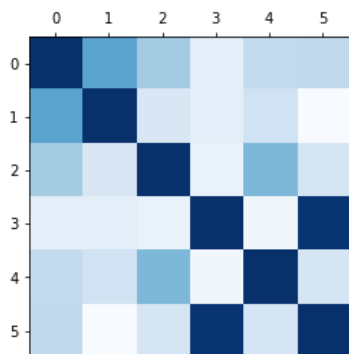
ΠΙΣΤΟΛΑΚΙ – ΣΚΟΥΠΑ – Η/Υ



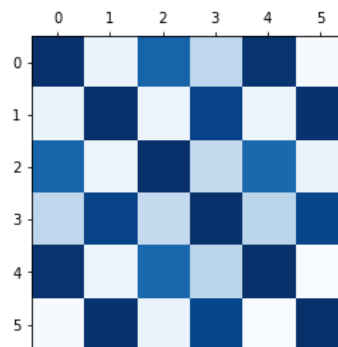
ΠΙΣΤΟΛΑΚΙ – ΣΚΟΥΠΑ – ΣΙΔΕΡΟ



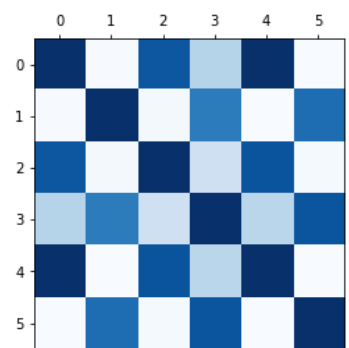
ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – Η/Υ



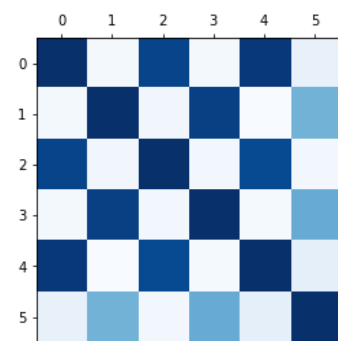
ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – ΜΑΤΙ



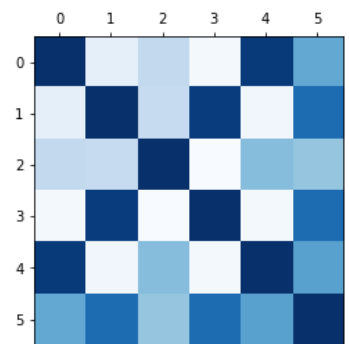
ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – ΣΙΔΕΡΟ



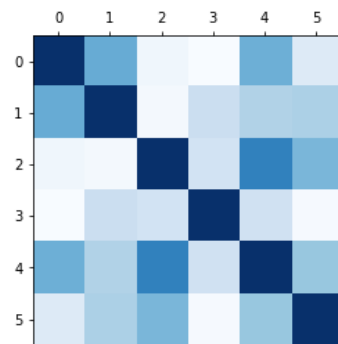
ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – ΣΚΟΥΠΑ



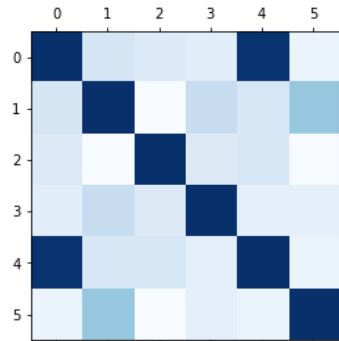
ΣΚΟΥΠΑ – ΣΙΔΕΡΟ – Η/Υ



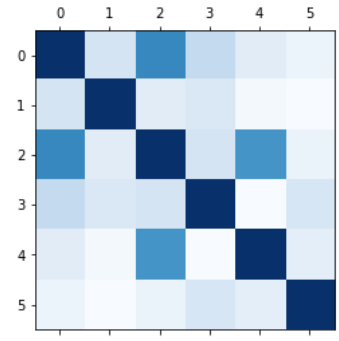
ΤΟΣΤΙΕΡΑ – ΜΑΤΙ – Η/Υ



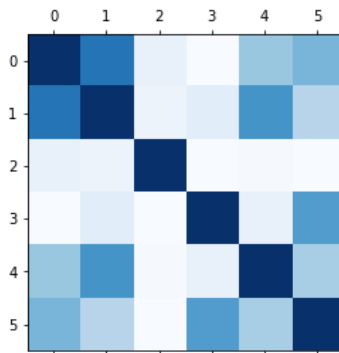
ΤΟΣΤΙΕΡΑ – ΜΑΤΙ – ΣΙΔΕΡΟ



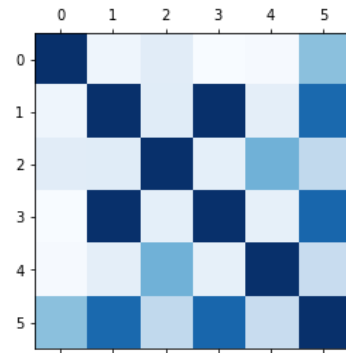
ΤΟΣΤΙΕΡΑ – ΜΑΤΙ – ΣΚΟΥΠΑ



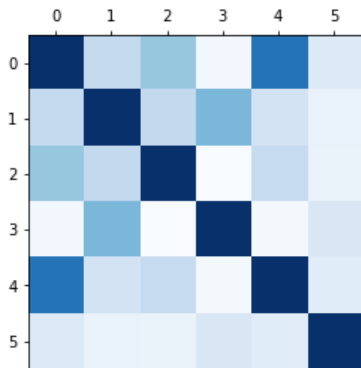
ΤΟΣΤΙΕΡΑ – ΣΙΔΕΡΟ – Η/Υ



ΤΟΣΤΙΕΡΑ – ΣΚΟΥΠΑ – Η/Υ



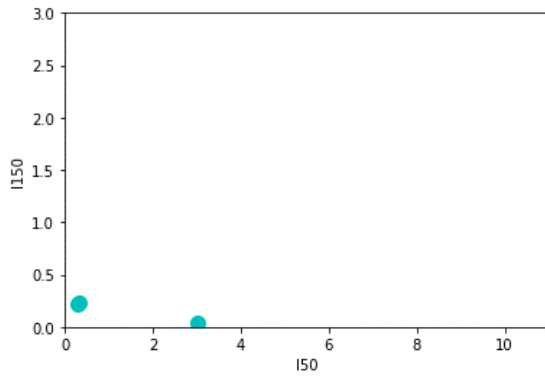
ΤΟΣΤΙΕΡΑ – ΜΑΤΙ – ΣΙΔΕΡΟ



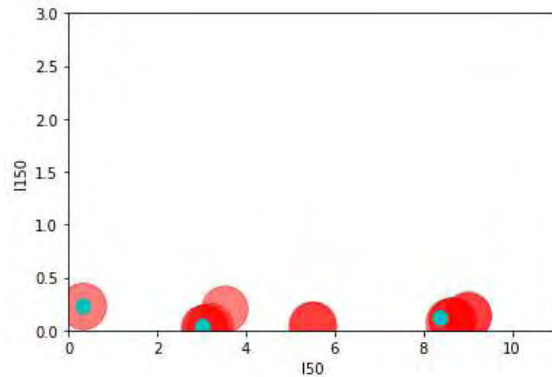
ΠΑΡΑΡΤΗΜΑ Γ

ΓΡΑΦΙΚΕΣ ΠΑΡΑΣΤΑΣΕΙΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΑΛΓΟΡΙΘΜΟΥ

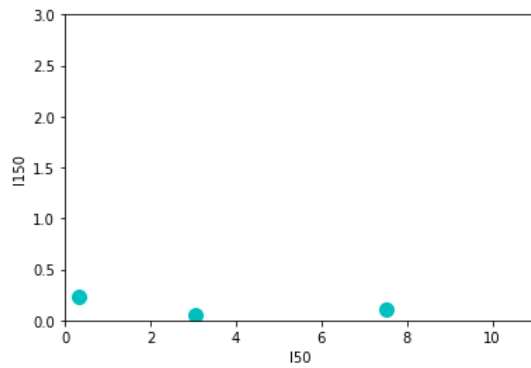
ΤΗΛΕΟΡΑΣΗ – ΜΑΤΙ – Η/Υ



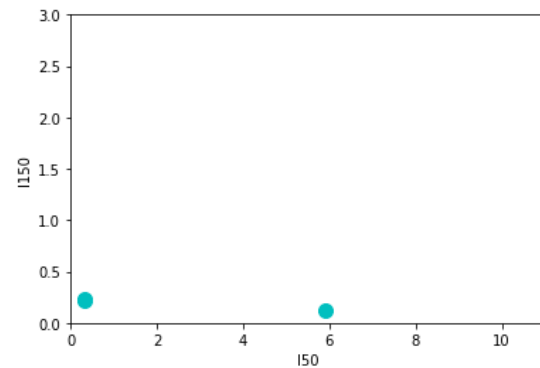
ΤΗΛΕΟΡΑΣΗ – ΜΑΤΙ – ΣΙΔΕΡΟ



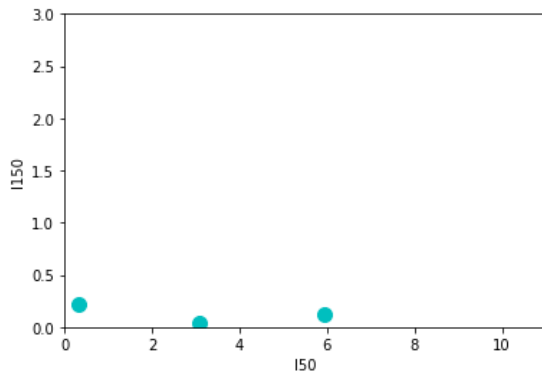
ΤΗΛΕΟΡΑΣΗ – ΜΑΤΙ – ΣΚΟΥΠΑ



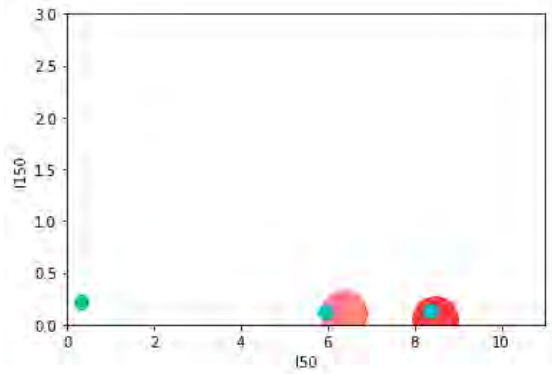
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – Η/Υ



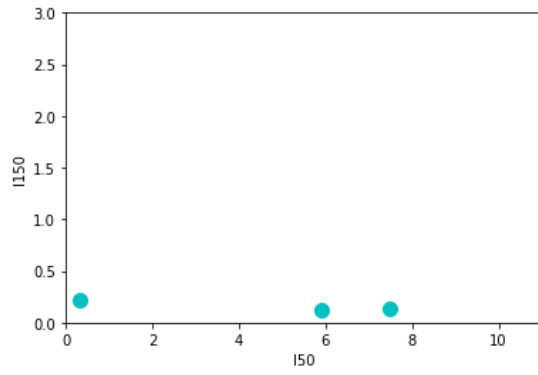
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – ΜΑΤΙ



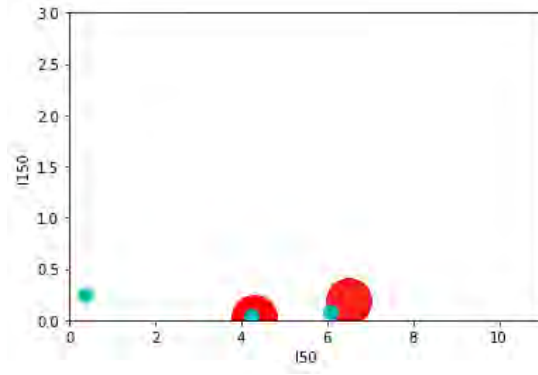
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – ΣΙΔΕΡΟ



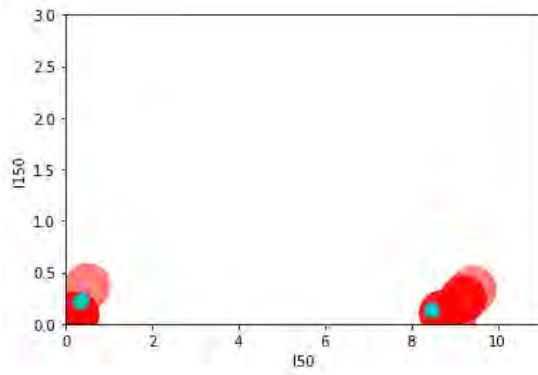
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – ΣΚΟΥΠΑ



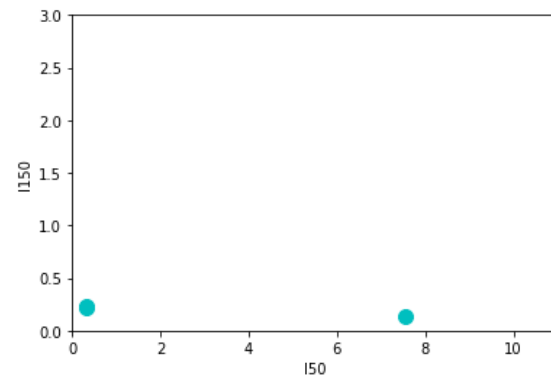
ΤΗΛΕΟΡΑΣΗ – ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ



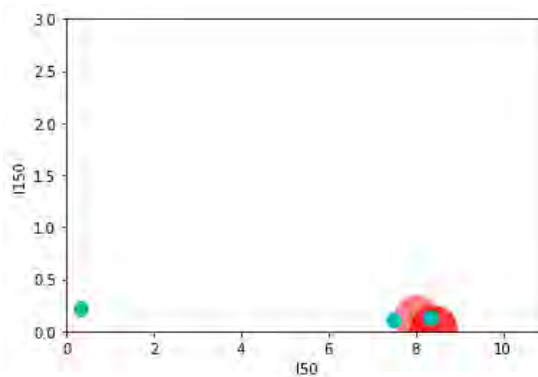
ΤΗΛΕΟΡΑΣΗ – ΣΙΔΕΡΟ – Η/Υ



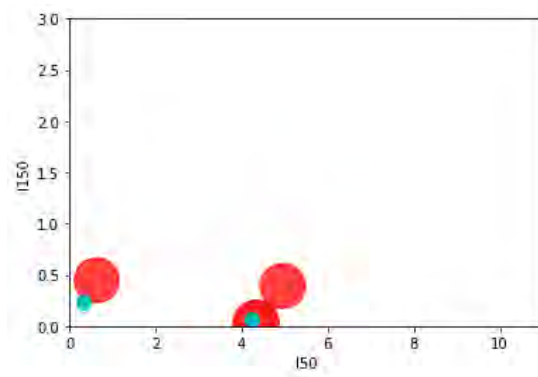
ΤΗΛΕΟΡΑΣΗ – ΣΚΟΥΠΑ – Η/Υ



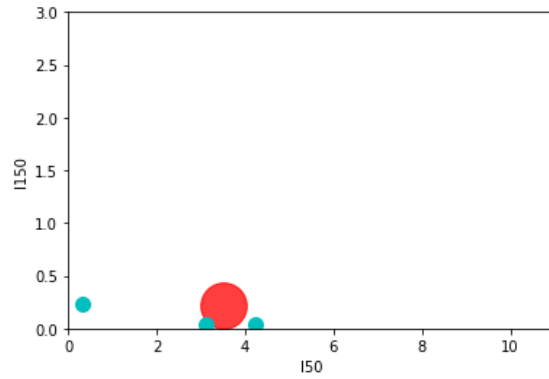
ΤΗΛΕΟΡΑΣΗ – ΣΚΟΥΠΑ – ΣΙΔΕΡΟ



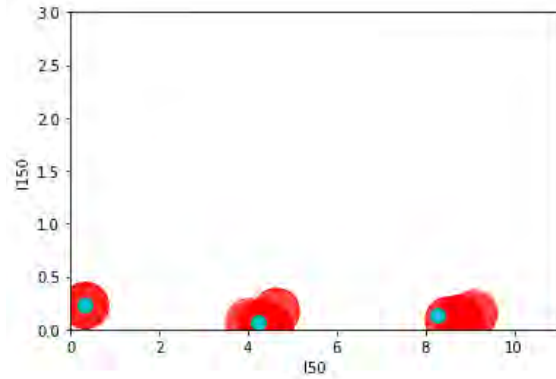
ΤΗΛΕΟΡΑΣΗ – ΤΟΣΤΙΕΡΑ – Η/Υ



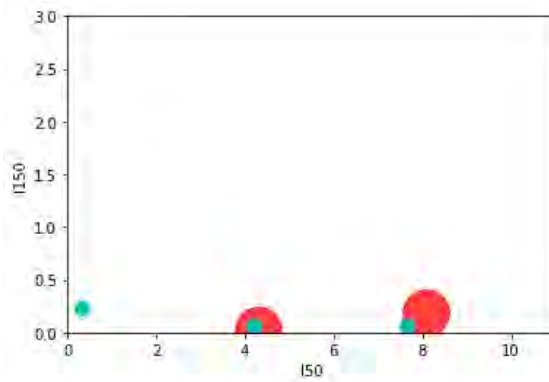
ΤΗΛΕΟΡΑΣΗ – ΤΟΣΤΙΕΡΑ – ΜΑΤΙ



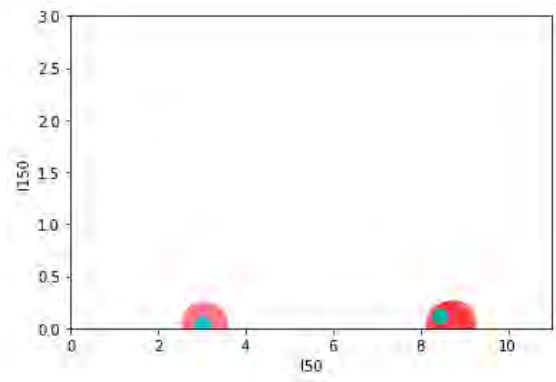
ΤΗΛΕΟΡΑΣΗ – ΤΟΣΤΙΕΡΑ – ΣΙΔΕΡΟ



ΤΗΛΕΟΡΑΣΗ – ΤΟΣΤΙΕΡΑ – ΣΚΟΥΠΑ



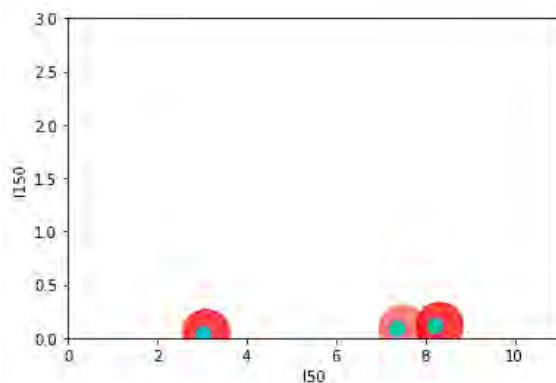
ΜΑΤΙ – ΣΙΔΕΡΟ – Η/Υ



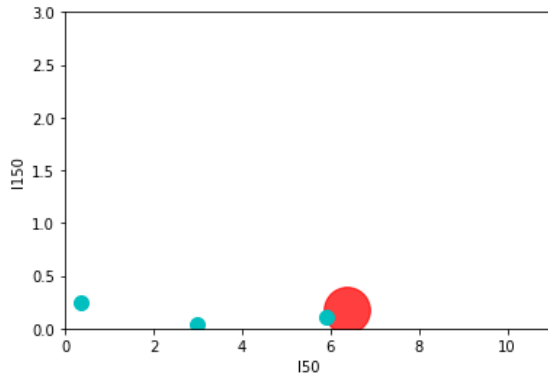
ΜΑΤΙ – ΣΚΟΥΠΑ – Η/Υ

ΔΕΝ ΕΠΕΞΕΡΓΑΣΤΗΚΕ ΛΟΓΟ ΚΑΚΗΣ
ΠΟΙΟΤΗΤΑΣ ΤΩΝ ΕΓΓΡΑΦΩΝ.

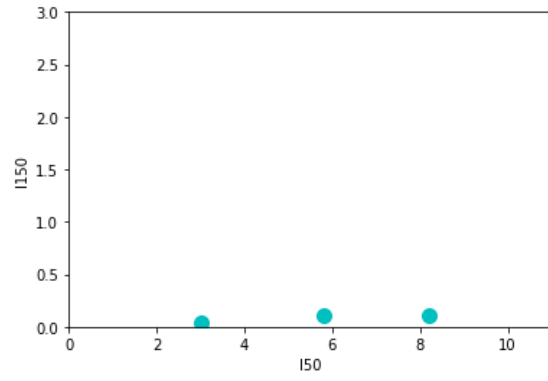
ΜΑΤΙ – ΣΚΟΥΠΑ – ΣΙΔΕΡΟ



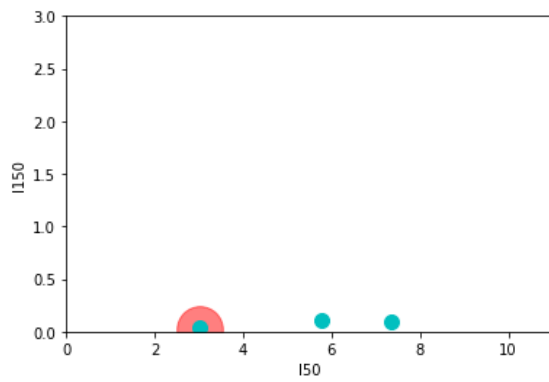
ΠΙΣΤΟΛΑΚΙ – ΜΑΤΙ – Η/Υ



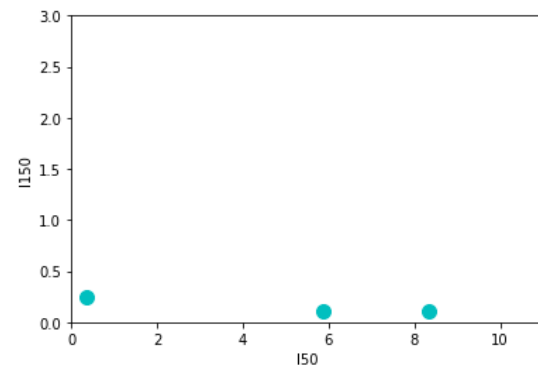
ΠΙΣΤΟΛΑΚΙ – ΜΑΤΙ – ΣΙΔΕΡΟ



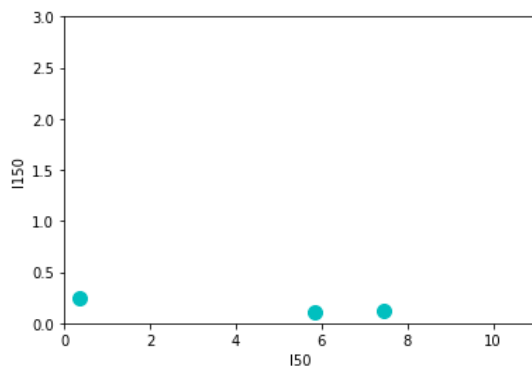
ΠΙΣΤΟΛΑΚΙ – ΜΑΤΙ – ΣΚΟΥΠΑ



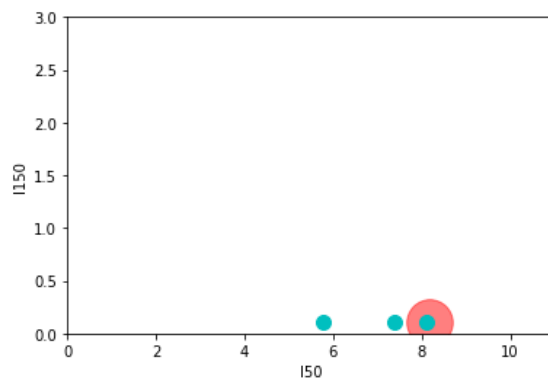
ΠΙΣΤΟΛΑΚΙ – ΣΙΔΕΡΟ – Η/Υ



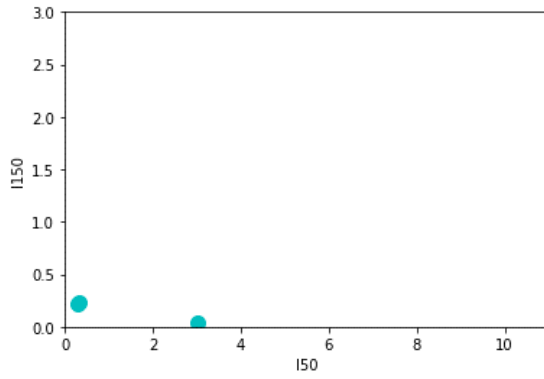
ΠΙΣΤΟΛΑΚΙ – ΣΚΟΥΠΑ – Η/Υ



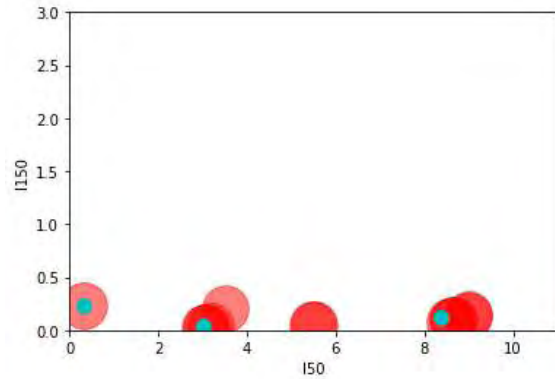
ΠΙΣΤΟΛΑΚΙ – ΣΚΟΥΠΑ – ΣΙΔΕΡΟ



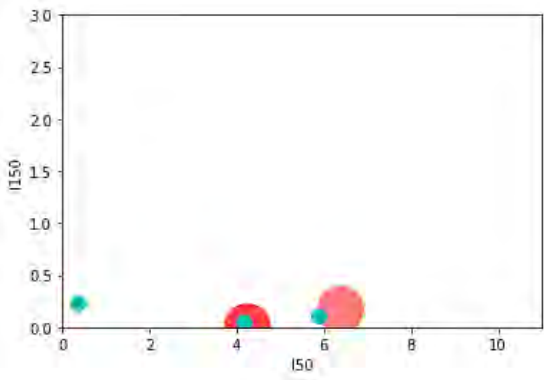
ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – Η/Υ



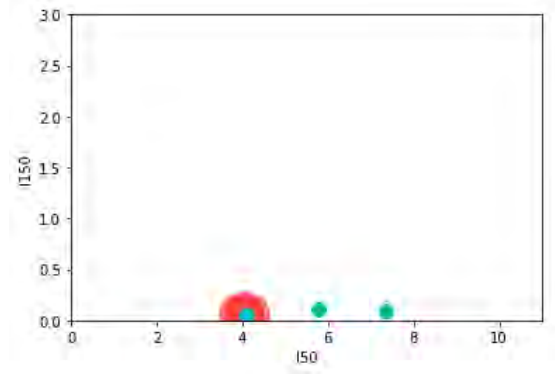
ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – ΜΑΤΙ



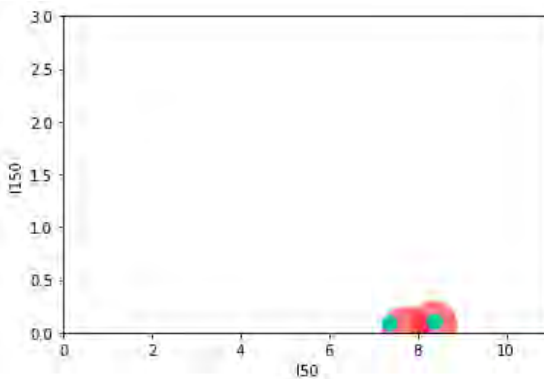
ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – ΣΙΔΕΡΟ



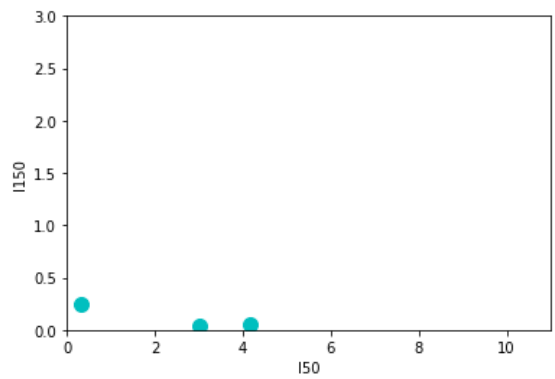
ΠΙΣΤΟΛΑΚΙ – ΤΟΣΤΙΕΡΑ – ΣΚΟΥΠΑ



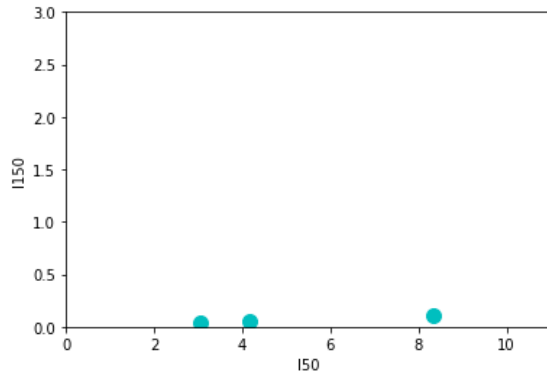
ΣΚΟΥΠΑ – ΣΙΔΕΡΟ – Η/Υ



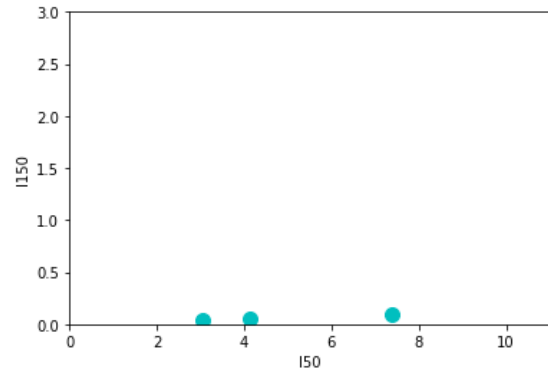
ΤΟΣΤΙΕΡΑ – ΜΑΤΙ – Η/Υ



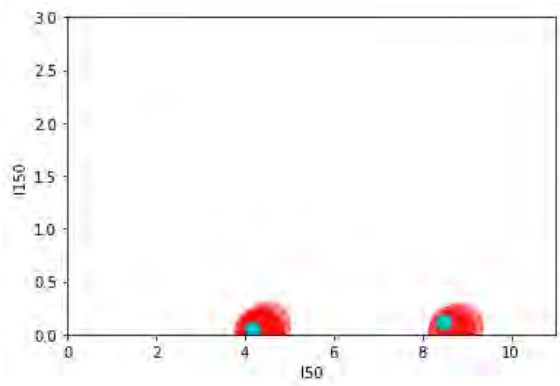
ΤΟΣΤΙΕΡΑ – ΜΑΤΙ – ΣΙΔΕΡΟ



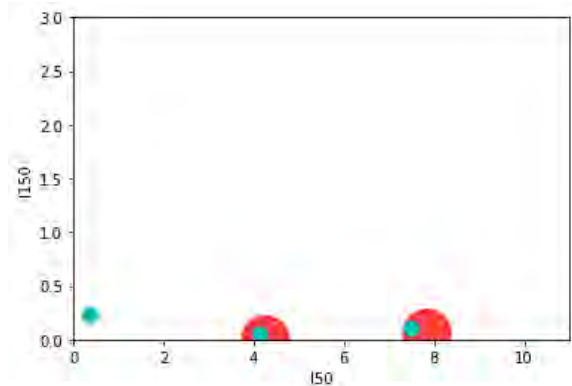
ΤΟΣΤΙΕΡΑ – ΜΑΤΙ – ΣΚΟΥΠΑ



ΤΟΣΤΙΕΡΑ – ΣΙΔΕΡΟ – Η/Υ



ΤΟΣΤΙΕΡΑ – ΣΚΟΥΠΑ – Η/Υ



ΤΟΣΤΙΕΡΑ – ΜΑΤΙ – ΣΙΔΕΡΟ

