



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ  
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

**Εφαρμογή τεχνικών συσταδοποίησης σε δεδομένα από  
έξυπνους μετρητές για το σχεδιασμό προγράμματος  
απόκρισης ζήτησης**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

Άγγελος Δ. Πιπεργιάς

**Επιβλέπων:** Δημήτριος Ασκούνης  
Καθηγητής ΕΜΠ

Αθήνα, Μάρτιος 2023





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ ΚΑΙ  
ΣΥΣΤΗΜΑΤΩΝ ΑΠΟΦΑΣΕΩΝ

**Εφαρμογή τεχνικών συσταδοποίησης σε δεδομένα από  
έξυπνους μετρητές για το σχεδιασμό προγράμματος  
απόκρισης ζήτησης**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Άγγελος Δ. Πιπεργιάς

**Επιβλέπων:** Δημήτριος Ασκούνης,  
Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 9<sup>η</sup> Μαρτίου 2023.

.....

Δημήτριος Ασκούνης

Καθηγητής ΕΜΠ

.....

Ιωάννης Ψαρράς

Καθηγητής ΕΜΠ

.....

Χρυσόστομος Δούκας

Αναπληρωτής Καθηγητής ΕΜΠ

Αθήνα, Μάρτιος 2023

.....  
Άγγελος Δ. Πιπεργιάς

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

Copyright © Άγγελος Πιπεργιάς, 2023

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

## Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον Σωτήρη Πελέκη, που από την αρχή της έρευνας μέχρι τη παράδοση αυτής της εργασίας με υποστήριξε με τις γνώσεις του, τις σκέψεις του και τις συμβουλές του.

Στην οικογένειά μου και τους φίλους μου: σας ευχαριστώ που είστε δίπλα μου. Είναι πολλές οι φορές που χάνομαι και νιώθω ευγνώμων που έχω εσάς, να μου δείχνετε το δρόμο.

## Περιεχόμενα

Ευχαριστίες.....	5
Σύνοψη-Περίληψη.....	8
Abstract.....	9
1. Κατάλογος Εικόνων.....	10
2. Κατάλογος Πινάκων.....	11
3. Εισαγωγή.....	12
3.1.1 Πλαίσιο.....	12
3.1.2 Σκοπός, ερωτήματα και στόχοι.....	13
3.1.3 Οριοθέτηση.....	13
3.1.4 Διάρθρωση της εργασίας.....	14
4. Βιβλιογραφική ανασκόπηση – Προγράμματα Απόκρισης Ζήτησης.....	15
4.1 Κατηγοριοποίηση Προγραμμάτων Απόκρισης Ζήτησης.....	15
4.1.1 Time Of Use (TOU).....	15
4.1.2 Critical Peak Pricing (CPP).....	17
4.1.3 Real Time Pricing (RTP).....	18
4.2 Επιλογή συμμετεχόντων για προγράμματα DR χρησιμοποιώντας δεδομένα από έξυπνους μετρητές.....	20
4.2.1 Ανάλογα με το σχήμα της κατανάλωσης.....	20
4.2.2 Ανάλογα με την εντροπία.....	21
4.3 Επιλογή συμμετεχόντων για προγράμματα DR με βάση τη χρήση.....	22
4.3.1 Εμπορικά φορτία.....	22
4.3.2 Οικιακά φορτία.....	23
5. Βιβλιογραφική Ανασκόπηση – Συσταδοποίηση σε δεδομένα από ηλεκτρικά φορτία	25
5.1 Προεπεξεργασία Δεδομένων.....	25
5.1.1 Διαχείριση κενών τιμών.....	25
5.2 Μείωση διαστατικότητας και επιλογή χαρακτηριστικών.....	25
5.2.1 Επιλογή συχνότητας δειγματοληψίας.....	26
5.2.2 Εξαγωγή χαρακτηριστικών (Feature Extraction).....	26
5.3 Μετρικές Απόστασης.....	27
5.3.1 Ευκλείδεια απόσταση.....	27
5.3.2 Απόσταση με δυναμική παραμόρφωση του χρόνου (Dynamic Time Warping)	27
5.4 Αλγόριθμοι.....	28
5.4.1 Αλγόριθμοι που βασίζονται στη διχοτόμηση (Partition Based Algorithms)	28
5.4.2 Ιεραρχικοί αλγόριθμοι (Hierarchical Algorithms).....	29
5.4.3 Αλγόριθμοι που βασίζονται στην πυκνότητα (Density Based Algorithms)	29
5.5 Μετρικές Επικύρωσης.....	30
5.5.1 Silhouette Score.....	30

5.5.2	Davies – Bouldin Index (DBI).....	30
5.5.3	Peak Match Score .....	31
6.	Μεθοδολογία .....	33
6.1	Σύνολο Δεδομένων.....	33
6.2	Ζητούμενο .....	33
6.3	Αρχική Ανάλυση Δεδομένων .....	34
6.4	Προεπεξεργασία των δεδομένων .....	36
6.4.1	Χειρισμός απομακρυσμένων τιμών (outliers).....	36
6.4.2	Χειρισμός missing data – Data imputation.....	36
6.4.3	Κανονικοποίηση Ημερήσιων Προφίλ.....	37
6.5	Επιλογές σχετικές με τη συσταδοποίηση.....	37
6.5.1	Ένα ή πολλά προφίλ ανά καταναλωτή.....	37
6.5.2	Εποχικός διαχωρισμός του συνόλου δεδομένων.....	38
6.5.3	Αλγόριθμοι και μετρικές απόστασης .....	38
6.5.4	Αριθμός συστάδων .....	39
6.6	Σύγκριση διαφορετικών επιλογών – Νέες μετρικές επικύρωσης .....	39
6.6.1	Μεθοδολογία σύγκρισης.....	39
6.6.2	Μετρικές επικύρωσης .....	40
6.6.3	ML Flow .....	40
6.7	Ορισμός νέων μετρικών επικύρωσης.....	42
6.7.1	Peak Performance Score.....	42
6.7.2	Χρονική χαλάρωση κατά 1 βήμα.....	44
6.8	Επιλογή των ιδανικότερων συστάδων για απόκριση ζήτησης.....	44
7.	Αποτελέσματα.....	46
7.1	Αποτελέσματα σύγκρισης μεταξύ αλγορίθμων και μετρικών απόστασης .....	46
7.2	Περιγραφή των συστάδων .....	49
7.3	Ανάλυση εντροπίας .....	57
8.	Συμπεράσματα – Προτάσεις .....	60
8.1	Προτάσεις ανά συστάδα .....	60
8.1.1	Χρήσεις φορτίων και μέση εντροπία ανά συστάδα .....	60
8.1.2	Κατηγορικός χαρακτηρισμός της εντροπίας .....	62
8.1.3	Τελικές προτάσεις .....	63
8.2	Ολοκληρωμένο σύστημα προγνώσεων, συσταδοποίησης και DR.....	65
8.2.1	Περιγραφή και παράδειγμα .....	65
8.2.2	Διάγραμμα του συστήματος .....	68
8.2.3	Recommendation Engine .....	69
8.2.4	Οντότητες που αλληλοεπιδρούν με το σύστημα .....	69
9.	Μελλοντική Εργασία .....	70
10.	Κατάλογος Βιβλιογραφίας .....	72

## Σύνοψη-Περίληψη

Τα τελευταία χρόνια οι αναλυτές στον κλάδο της ενέργειας παρακολουθούν με όλο και αυξανόμενο ενδιαφέρον την απόκριση ζήτησης (DR) σαν ένα εργαλείο για την καλύτερη διαχείριση του ηλεκτρικού δικτύου τις ώρες αιχμής και την διατήρηση της ισορροπίας μεταξύ παραγωγής και κατανάλωσης. Ωστόσο, η εφαρμογή προγραμμάτων απόκρισης ζήτησης είναι ακόμα περιορισμένη, ειδικά όσον αφορά τους οικιακούς καταναλωτές. Όσο το κόστος του εξοπλισμού που επιτρέπει σε σπίτια και επιχειρήσεις να συμμετέχουν στην απόκριση ζήτησης (πχ. έξυπνοι μετρητές, έξυπνοι ελεγκτές και συσκευές) μικραίνει και η τεχνολογία εξελίσσεται, αυξάνεται η ανάγκη για το σχεδιασμό και την εφαρμογή προγραμμάτων που θα εντάξουν τους καταναλωτές που ενδιαφέρονται.

Αυτή η εργασία εξερευνά τη χρήση τεχνικών συσταδοποίησης για την υποστήριξη του σχεδιασμού και της εφαρμογής ενός προγράμματος απόκρισης ζήτησης για ένα δίκτυο με εμπορικούς και οικιακούς καταναλωτές. Στόχος του προγράμματος είναι να μεταβάλλει την καταναλωτική συμπεριφορά των συμμετεχόντων με σκοπό την αντιμετώπιση δύο προβλημάτων στον συγχρονισμό παραγωγής και κατανάλωσης στο ηλεκτρικό δίκτυο: α) την αντίστροφη ροή φορτίου, που συμβαίνει όταν η παραγωγή από φωτοβολταϊκά στο τοπικό δίκτυο είναι μεγαλύτερη από την κατανάλωση και β) την κορύφωση της ζήτησης στο Ιταλικό εθνικό δίκτυο, που συνήθως συμβαίνει τις ώρες του απογεύματος. Για το στάδιο της συσταδοποίησης, δοκιμάζονται τρεις διαφορετικοί αλγόριθμοι, ο k-means, ο k-medoids και ένα ιεραρχικός αλγόριθμος, σε συνδυασμό με δύο μετρικές απόστασης, την ευκλείδεια και μία απόσταση constrained Dynamic Time Warping (DTW). Τα αποτελέσματά τους συγκρίνονται με μετρικές επικύρωσης, μεταξύ των οποίων υπάρχει και μία που δημιουργήθηκε για αυτή την εργασία, το Peak Performance Score. Στη συνέχεια, ο καλύτερος συνδυασμός παραμέτρων εφαρμόζεται στο διαχωρισμό των ημερήσιων προφίλ του συνόλου δεδομένων σε συστάδες και κάθε συστάδα χαρακτηρίζεται σε σχέση με το σχήμα της, την μέση εντροπία και τη χρήση των φορτίων που ανήκουν σε αυτή (οικιακά, εμπορικά κτλ.). Αυτά τα χαρακτηριστικά χρησιμοποιούνται για το διαχωρισμό των συστάδων που θα μπορούσαν να βοηθήσουν στην επίτευξη των στόχων του προγράμματος και την επιλογή τύπων προγραμμάτων που θα ταίριαζαν σε κάθε συστάδα. Τέλος, η εργασία προτείνει την υλοποίηση ενός ψηφιακού συστήματος το οποίο χρησιμοποιεί προγνωστικά μοντέλα, συσταδοποίηση και ένα μηχανισμό υπολογισμού της ζήτησης βασισμένο στην τιμή για την παραγωγή ημερήσιων, εξατομικευμένων προτάσεων και τιμολογιακών πολιτικών για κάθε καταναλωτή που συμμετέχει στο πρόγραμμα απόκρισης ζήτησης.

Λέξεις κλειδιά: απόκριση ζήτησης, prosumers, αντίστροφη ροή φορτίου, συσταδοποίηση, Dynamic Time Warping, επιλογή καταναλωτών, εντροπία φορτίου



## Abstract

Recently, there's been growing interest in demand response (DR) as a tool from the management of peak demand and the balance of generation and consumption in the electrical grid. However, the implementation of demand response programs is still quite limited, especially when it comes to residential consumers. As the cost of equipment that can help homes and businesses participate in demand response (i.e. smart meters, controllers and devices) decreases, there is a growing need for the design and implementation of programs to get interested consumers involved with DR.

This thesis explores the use of clustering techniques to aid in the design and implementation of a Demand Response (DR) program for a network of commercial and residential prosumers. The goal of the program is to shift participant's consumption behaviors to mitigate two issues with demand and generation timing in the electrical grid: a) reverse power flow, that occurs when generation from solar panels in the local grid exceeds consumption and b) system wide peak demand, that typically occurs during hours of the late afternoon. For the clustering stage, three popular algorithms for electrical load clustering, namely k-means, k-medoids and a hierarchical clustering algorithm, along with two different distance metrics, Euclidean and constrained Dynamic Time Warping (DTW), are evaluated using different validation metrics. The best configuration is employed to divide the dataset's daily load profiles into clusters and each cluster is analyzed in terms of load shape, mean entropy and distribution of load profiles from each load type. These characteristics are then used to distinguish the clusters that would be most likely to aid with the DR program's objectives and to select DR program structures that would fit each cluster. Finally, the thesis proposes the design of a DR system that uses forecasting, clustering and a demand projection engine based on price to produce daily, individualized DR recommendations and pricing structures for each customer participating in the program.

Apart from leading to the implementation of a possible DR program for the network analyzed, the methodology followed in this thesis along with the proposed digital system can hopefully help with the design and implementation of more DR programs in the future.

Keywords: Demand Response, prosumers, reverse power flow, clustering, Dynamic Time Warping, customer selection, load entropy

## 1. Κατάλογος Εικόνων

Εικόνα 1 – Παράδειγμα τιμολογιακής πολιτικής προγραμμάτων ΤΟΥ .....	15
Εικόνα 2 - άθροισμα παραγωγής και κατανάλωσης ενέργειας στο εικονικό δίκτυο .....	33
Εικόνα 3 - Heatmap με αριθμό μετρήσεων ανά μήνα και ανά smart meter.....	34
Εικόνα 4 - Bar plots με αριθμό δειγμάτων ανά μέρα της εβδομάδας και ώρα της ημέρας για τον μετρητή BBB6004 .....	35
Εικόνα 5 – Line Plot της κατανάλωση ενέργειας για μεγάλο χρονικό διάστημα για τον μετρητή BBB6030 – βοηθητικό για εντοπισμό outliers .....	35
Εικόνα 6 - λίστα του ML Flow με όλα τα πειράματα που εκτελέστηκαν για την επιλογή του καλύτερου configuration .....	41
Εικόνα 7 - κατάταξη πειραμάτων με βάση τη μετρική Relaxed Peak Performance .....	41
Εικόνα 8 - γραφική παράσταση που παρήγαγε το MLFlow για τη διερεύνηση του συσχετίσης μεταξύ των μετρικών Relaxed Peak Performance Score και Silhouette Score DTW .....	42
Εικόνα 9 - γραφική παράσταση που παρήγαγε το MLFlow για τη σύγκριση των αποτελεσμάτων της συσταδοποίησης με διάφορους συνδυασμούς παραμέτρων .....	47
Εικόνα 10 - Τα σχήματα των συστάδων που προέκυψαν για αλγόριθμο k-means, μετρική απόστασης DTW και k=14. Με διακεκομμένη γραμμή είναι το κέντρο της κάθε συστάδας και με αχνό πορτοκαλί τα δείγματα που την απαρτίζουν .....	48
Εικόνα 11 – Ποσοστά επί των συνολικών δειγμάτων που καλύπτει η κάθε συστάδα.....	49
Εικόνα 12 - Εντροπία για κάθε μετρητή στο σύνολο δεδομένων. Οι μετρητές έχουν ταξινομηθεί ανάλογα με την εντροπία (χαμηλή εντροπία αριστερά, υψηλή δεξιά).....	58
Εικόνα 13 - Μέση εντροπία για κάθε συστάδα .....	58
Εικόνα 14 – Ποσοστά προφίλ κάθε συστάδας που αντιστοιχούν στις διαφορετικές χρήσεις (περίπτωση που κάθε φορτίου ανήκει σε μία μόνο συστάδα) .....	61
Εικόνα 15 – Μέση εντροπία κάθε συστάδας (περίπτωση που κάθε φορτίο ανήκει σε μία συστάδα).....	62
Εικόνα 16 – κατηγορικός χαρακτηρισμός της εντροπίας .....	63
Εικόνα 17 - Πρόγνωση ημερήσιας κατανάλωσης για μετρητή BBB6168 .....	66
Εικόνα 18 - Συστάδα στην οποία αντιστοιχεί η πρόβλεψη για τον μετρητή BBB6168 .....	67
Εικόνα 19 - UML διάγραμμα του συστήματος απόκρισης ζήτησης .....	69

## 2. Κατάλογος Πινάκων

Πίνακας 1 - Μετρικές επικύρωσης που χρησιμοποιήθηκαν για τον έλεγχο των μεθόδων clustering .....	40
Πίνακας 2 - συνδυασμοί παραμέτρων με τις καλύτερες επιδόσεις στις μετρικές Relative Peak Performance Score και Silhouette DTW .....	47
Πίνακας 3 – προτεινόμενα προγράμματα για κάθε συστάδα και χαρακτηριστικά που βοήθησαν στην επιλογή .....	63

## 3. Εισαγωγή

### 3.1.1 Πλαίσιο

Με την εγκατάσταση όλο και περισσότερων ανανεώσιμων πηγών ενέργειας (ΑΠΕ), η παραγωγή στο δίκτυο ηλεκτρικής ενέργειας γίνεται όλο και πιο αποκεντρωμένη, πράγμα που δυσκολεύει την πρόβλεψη και διαχείριση της. Παράλληλα, η ζήτηση στο δίκτυο αυξάνεται χρόνο με το χρόνο. Αυτές οι δύο πραγματικότητες καθιστούν τη διατήρηση της ισορροπίας μεταξύ κατανάλωσης και παραγωγής στο δίκτυο ηλεκτρικής ενέργειας όλο και πιο δύσκολη, ειδικά τις ώρες αιχμής, που η κατανάλωση στο δίκτυο κορυφώνεται.

Ειδικά για τη διατήρηση της ισορροπίας παραγωγής και κατανάλωσης τις ώρες αιχμής, συχνά απαιτείται η λειτουργία μεγάλων μονάδων παραγωγής ενέργειας με ορυκτά καύσιμα. Αυτές οι μονάδες παράγουν ρεύμα ακριβό και ρυπογόνο, ενώ συχνά πρέπει να λειτουργούν συνέχεια, ακόμα και όταν η κατανάλωση θα μπορούσε να καλυφθεί και χωρίς τη συνεισφορά τους, γιατί πρέπει η λειτουργία τους να είναι συνεχόμενη. Παράλληλα, όσο η κατανάλωση αυξάνεται, νέες επενδύσεις σε μονάδες παραγωγής και γραμμές διανομής γίνονται απαραίτητες μόνο για την αναγκαιότητα να λειτουργεί το δίκτυο τις ώρες αιχμής.

Με βάση τα παραπάνω γίνεται ξεκάθαρο ότι το φαινόμενο κορύφωσης της κατανάλωσης τις ώρες αιχμής οδηγεί σε αύξηση του κόστους συντήρησης του δικτύου ηλεκτρικής ενέργειας και της τιμής του ρεύματος, πέρα από τις επιπτώσεις που έχει στο περιβάλλον. Για την αντιμετώπιση αυτών των προβλημάτων απαιτείται ένας μηχανισμός που θα επέτρεπε τη μείωση της κατανάλωσης τις ώρες αιχμής και άρα την επιπεδοποίηση της καμπύλης κατανάλωσης. Αυτός ο μηχανισμός είναι η απόκριση ζήτησης.

Η απόκριση ζήτησης είναι μία διαδικασία κατά την οποία οι διαχειριστές του δικτύου ηλεκτρικής ενέργειας ή άλλες οντότητες (πχ. πάροχοι) προσφέρουν κίνητρα στους καταναλωτές ώστε να τους παρακινήσουν να αλλάξουν τις καταναλωτικές τους συνήθειες. Όταν εφαρμόζεται σωστά μπορεί να βοηθήσει την καλύτερη διαχείριση της κατανάλωσης τις ώρες αιχμής (peak demand) και τη διατήρηση ισορροπίας κατανάλωσης – παραγωγής στο ηλεκτρικό δίκτυο [1].

Η απόκριση ζήτησης είναι μία τεχνολογία που βρίσκεται ακόμα σε αρχικά στάδια ανάπτυξης, ειδικά όσον αφορά τους οικιακούς καταναλωτές. Ένα από τα βασικά εμπόδια στην εξέλιξή της είναι ότι δεν υπάρχουν πολλά δεδομένα από πραγματικά προγράμματα. Εάν υπήρχαν τέτοια δεδομένα θα ήταν πιο εύκολο για τους αναλυτές της αγοράς ενέργειας να δείξουν τα οφέλη που μπορεί να έχει η απόκριση ζήτησης τόσο για το δίκτυο όσο και για τους καταναλωτές, πράγμα απαραίτητο για να ενδιαφερθεί η κοινή γνώμη, να γίνουν επενδύσεις σε εργατώρες και εξοπλισμό και η απόκριση ζήτησης να μπει επιτέλους στα σπίτια των καταναλωτών.

Σε αυτή την πρώιμη φάση, είναι κρίσιμο τα προγράμματα απόκρισης ζήτησης που εφαρμόζονται να είναι προσεκτικά σχεδιασμένα, ώστε να έχουν τα καλύτερα δυνατά αποτελέσματα. Πρώτο μέλημα για να συμβεί αυτό είναι η προσεκτική επιλογή συμμετεχόντων

και δεύτερο ο σχεδιασμός τιμολογιακών πολιτικών που θα δώσουν στους συμμετέχοντες αρκετό κίνητρο να αλλάξουν την κατανάλωσή τους με τρόπους επιθυμητούς για το δίκτυο.

### 3.1.2 Σκοπός, ερωτήματα και στόχοι

Σε αυτή την εργασία, σκοπός είναι ο σχεδιασμός ενός προγράμματος απόκρισης ζήτησης για ένα δίκτυο με prosumers (καταναλωτές και παραγωγούς ενέργειας).

Βασικό ερευνητικό ερώτημα είναι το ποια χαρακτηριστικά καθιστούν κάποιους καταναλωτές πιο ιδανικούς από άλλους για συμμετοχή σε προγράμματα απόκρισης ζήτησης. Επίσης διερευνάται το αν κάποιες δομές προγραμμάτων είναι πιο ιδανικές για συγκεκριμένες ομάδες καταναλωτών.

Τα παραπάνω ερωτήματα βρίσκουν απάντηση μέσα από τη βιβλιογραφική έρευνα. Στο πρακτικό κομμάτι της εργασίας, τα συμπεράσματα της βιβλιογραφικής έρευνας χρησιμοποιούνται σε συνδυασμό με τεχνικές ανάλυσης δεδομένων για τον προσδιορισμό των ομάδων καταναλωτών που θα βοηθούσαν περισσότερο στην επίτευξη των στόχων του προγράμματος απόκρισης ζήτησης και τον σχεδιασμό προγραμμάτων με συγκεκριμένη τιμολογιακή δομή για αυτές τις ομάδες καταναλωτών.

### 3.1.3 Οριοθέτηση

Η ανάλυση που ακολουθεί γίνεται με διαθέσιμα δεδομένα μόνο τις μετρήσεις από έξυπνους μετρητές και κάποια βασικά χαρακτηριστικά για τις εγκαταστάσεις στις οποίες είναι συνδεδεμένοι (μέγιστη ισχύς που αναγράφεται στο συμβόλαιο, χρήση του φορτίου, αν υπάρχει παραγωγή ή όχι). Δεν υπήρχαν άλλες διαθέσιμες πληροφορίες για τις συσκευές και το μίγμα κατανάλωσης κάθε εγκατάστασης, ούτε υπήρχε δυνατότητα επικοινωνίας με τους καταναλωτές και αποστολής ερωτηματολογίων. Τέτοιες πληροφορίες θα βοηθούσαν σίγουρα στον αρτιότερο σχεδιασμό ενός προγράμματος απόκρισης ζήτησης και θα μπορούσαν να συμπεριληφθούν σε κάποια μελλοντική ανάλυση.

Επίσης δεν γίνεται ανάλυση του πως η ευελιξία που θα πρόσφερε το προτεινόμενο πρόγραμμα θα μπορούσε να πουληθεί ως προϊόν στην αγορά ενέργειας, πράγμα σημαντικό για την οικονομική βιωσιμότητα του προγράμματος αλλά και τον ακριβή προσδιορισμό της τιμολογιακής πολιτικής.

Τέλος, τα προγράμματα που αναλύονται στο κεφάλαιο 4 και προτείνονται για τους καταναλωτές στο κεφάλαιο 8 είναι προγράμματα που βασίζονται στην τιμή (price based). Δεν έγινε λεπτομερής ανάλυση για προγράμματα που βασίζονται στα κίνητρα (incentive based). Αυτό έγινε γιατί στην αρχή της εργασίας υπήρχε ο στόχος του προσδιορισμού της τιμολογιακής πολιτικής του προγράμματος με χρήση ενισχυτικής μάθησης και μεταβλητή απόφασης την τιμή του ρεύματος για κάθε ώρα της ημέρας (περισσότερες πληροφορίες στο κεφάλαιο 9). Αυτή η ιδέα τελικά δεν υλοποιήθηκε λόγω πίεσης χρόνου και επειδή η έκταση της ανάλυσης με συσταδοποίηση είναι ήδη αρκετή για μία διπλωματική εργασία.

#### **3.1.4 Διάρθρωση της εργασίας**

Στο κεφάλαιο 2 αυτής της εργασίας παρουσιάζεται η βιβλιογραφική έρευνα σχετικά με τα προγράμματα απόκριση ζήτησης και στο κεφάλαιο 3 σχετικά με τις τεχνικές συσταδοποίησης. Ακολουθεί η παρουσίαση της μεθοδολογίας στο κεφάλαιο 4 και των αποτελεσμάτων της στο κεφάλαιο 5. Το κεφάλαιο 6 περιέχει συγκεκριμένες προτάσεις για προγράμματα απόκρισης ζήτησης που προκύπτουν από τα αποτελέσματα του κεφαλαίου 5, ενώ το κεφάλαιο 7 περιέχει κάποιες προτάσεις για μελλοντική έρευνα.

## 4. Βιβλιογραφική ανασκόπηση – Προγράμματα Απόκρισης Ζήτησης

### 4.1 Κατηγοριοποίηση Προγραμμάτων Απόκρισης Ζήτησης

Υπάρχουν δύο κύρια είδη προγραμμάτων απόκρισης ζήτησης: αυτά που βασίζονται στην τιμή (price or time based DR) και αυτά που βασίζονται στα κίνητρα (incentive based DR) [2]. Στα time based προγράμματα η τιμή του ρεύματος αλλάζει μέσα στη μέρα, με τους καταναλωτές να έχουν συμφέρον να καταναλώνουν περισσότερο ρεύμα τις περιόδους με χαμηλή τιμή και λιγότερο της περιόδους με υψηλή. Στα incentive based προγράμματα η τιμή του ρεύματος είναι σταθερή, αλλά οι διαχειριστές του προγράμματος ανταμείβουν τους καταναλωτές που καταφέρνουν να μειώσουν την προβλεπόμενη κατανάλωσή τους σε περιόδους που το σύστημα βρίσκεται υπό πίεση (πχ. η ζήτηση είναι πολύ μεγάλη και η παραγωγή ακριβή).

Παρακάτω παρουσιάζονται οι 3 βασικότεροι τύποι time based προγραμμάτων (TOU, CPP και RTP προγράμματα).

#### 4.1.1 Time Of Use (TOU)

Αυτά τα προγράμματα χωρίζουν το χρόνο σε περιόδους και προσφέρουν μία προκαθορισμένη τιμή ρεύματος για κάθε περίοδο [3]. Μία περίοδος μπορεί να διαρκεί από κάποιες ώρες μέχρι κάποιες μέρες. Ο aggregator επιλέγει τις τιμές κατά τον σχεδιασμό του προγράμματος, ώστε να είναι υψηλές τις ώρες που επιθυμεί μείωση στην κατανάλωση (peak hours) και χαμηλές τις υπόλοιπες ώρες (off peak hours). Αυτή είναι η φιλοσοφία του νυχτερινού τιμολογίου, που εφαρμόζεται ευρέως στην Ελλάδα για τους οικιακούς καταναλωτές. Η τιμή του ρεύματος είναι χαμηλότερη τις νυχτερινές ώρες για να παροτρύνει τους καταναλωτές να λειτουργούν κάποιες συσκευές τότε αντί για μέσα στη μέρα. Ως δεύτερο παράδειγμα, στην Εικόνα 1 υπάρχει η τιμολογιακή πολιτική για δύο διαφορετικά προγράμματα TOU από την Washington των ΗΠΑ (από το άρθρο [4])

TOU AND FLAT RATES FROM THE CHOSEN UTILITIES (SUMMER)- 2009 DATA

Company	Pattern	Period	TOU Rate (\$/kWh)	Flat Rate (\$/kWh)
BGE	On-peak	10am-8pm	0.18897	0.15062
	Shoulder	7am-10am 8pm-11pm	0.13140	
	Off-peak	11pm-7am	0.12116	
Dominion	On-peak	11am-10pm	0.15085	0.06507
	Off-peak	10pm-11am	0.01514	

Εικόνα 1 – Παράδειγμα τιμολογιακής πολιτικής προγραμμάτων TOU

Τα προγράμματα ΤΟΥ έχουν το πλεονέκτημα ότι επιτρέπουν στους καταναλωτές να κάνουν ένα σχεδιασμό στην κατανάλωση των διάφορων συστημάτων και συσκευών μέσα στη μέρα στην αρχή του προγράμματος και στη συνέχεια να ακολουθούν αυτό το σχεδιασμό [3]. Δεν χρειάζεται να είναι ευέλικτοι στο βαθμό που να ανταποκρίνονται σε real time σήματα ούτε να διαθέτουν έξυπνους ελεγκτές για αυτό το σκοπό. Επίσης μπορούν να υπολογίσουν την εξοικονόμηση που θα πετύχουν στον λογαριασμό του ρεύματος αν συμμετέχουν στο πρόγραμμα, κάτι που είναι πιο δύσκολο σε προγράμματα που έχουν πιο πολύπλοκη δομή και η τιμή για κάθε περίοδο της ημέρας δεν είναι προκαθορισμένη.

Στο [5] παρουσιάζεται μία μέθοδος για τον υπολογισμό της ενεργειακής και χρηματικής εξοικονόμησης που μπορεί να πετύχει μία οικία εξοπλισμένη με έξυπνες συσκευές εάν ενταχθεί σε ένα πρόγραμμα ΤΟΥ. Είναι σημαντικό να αναφερθεί ότι σύμφωνα με την οικονομική ανάλυση στο άρθρο, η επένδυση σε έξυπνες συσκευές δεν είναι αποδοτική για τον καταναλωτή, εάν οι τιμές του προγράμματος ΤΟΥ δεν διαφέρουν σημαντικά από τις στατικές τιμές.

Από τη μεριά του aggregator, τα ΤΟΥ προγράμματα είναι πιο απλά στο σχεδιασμό και τη διαχείριση (αφού δεν χρειάζεται να παρακολουθεί και να μεταβάλλει την τιμή ή να επικοινωνεί καθημερινά με τους καταναλωτές) και φαίνεται να είναι πιο δημοφιλή για τους καταναλωτές σε σχέση με προγράμματα δυναμικής προσαρμογής της τιμής (Real Time Pricing ή RTP) [6]. Ωστόσο, ενδέχεται να δημιουργούν έντονες νέες κορυφές στη ζήτηση, όταν η τιμή πέφτει από υψηλό σε χαμηλότερο επίπεδο [4], [5]. Αυτό το φαινόμενο περιορίζει τον αριθμό των καταναλωτών που μπορούν να ενταχθούν στο ίδιο πρόγραμμα ΤΟΥ (ή προγράμματα με παρόμοιες τιμές και περιόδους), αφού μεγάλη συμμετοχή καταναλωτών στο ίδιο πρόγραμμα θα δημιουργούσε μία νέα κορυφή για το σύστημα διανομής.

Στο [3] παρουσιάζονται αποτελέσματα από την εφαρμογή προγραμμάτων ΤΟΥ σε διαφορετικά μέρη του κόσμου. Τα προγράμματα αυτά πέτυχαν κατά μέσω όρο μείωση της ζήτησης κατά την ώρα του peak demand κατά περίπου 10%. Άλλες πηγές εκτιμούν την αναμενόμενη μείωση στο peak demand για απλά ΤΟΥ προγράμματα πιο χαμηλά, κοντά στο 5% [7]. Το 5% μπορεί να φαίνεται μικρό αλλά δεν είναι αμελητέο: στο [8] οι Rosenzweig και συνεργάτες εκτιμούν ότι μία μείωση της κατανάλωσης μόνο κατά 2-5% τις ώρες αιχμής θα μείωνε την spot τιμή της ηλεκτρικής ενέργειας κατά 50% ή και περισσότερο. Ωστόσο υπάρχουν και ΤΟΥ προγράμματα που δεν πετυχαίνουν τον στόχο τους: στο [3] αναφέρονται προγράμματα που δεν πέτυχαν αξιόλογα αποτελέσματα, όπως ένα πρόγραμμα στην Βόρεια Ιταλία, το οποίο δεν κατάφερε να μειώσει την κατανάλωση κατά τις ώρες του peak demand. Για την ακρίβεια, 75,6% των υποσταθμών της περιοχής που εφαρμόστηκε στο πρόγραμμα κατέγραψαν αύξηση στην κατανάλωση αυτές τις ώρες.

Συμπερασματικά, τα ΤΟΥ προγράμματα είναι απλά στο σχεδιασμό και στην εφαρμογή, τόσο για τους καταναλωτές όσο και για τους aggregators, και ως εκ τούτου είναι δημοφιλή με τους καταναλωτές και καλές επιλογές για εγκαταστάσεις που δεν διαθέτουν αυτόματους ελεγκτές για τον χειρισμό των ευέλικτων φορτίων. Ωστόσο, έχουν γενικά μεικτά αποτελέσματα και έχει



παρατηρηθεί ότι δημιουργούν έντονες νέες ακμές στη ζήτηση, όταν η τιμή του ρεύματος πέφτει.

#### 4.1.2 Critical Peak Pricing (CPP)

Σε αντίθεση με τα προγράμματα TOU, στα οποία η τιμή της ενέργειας για κάθε ώρα της ημέρας ή εβδομάδας είναι προκαθορισμένη, η τιμή στα CPP προγράμματα μπορεί να αλλάξει για κάποιες ώρες, όταν προβλέπεται ότι θα υπάρχει μεγάλη συμφόρηση στο δίκτυο, όπως όταν κάνει πολύ ζέστη ή πολύ κρύο [3]. Οι περίοδοι που συμβαίνει αυτή η αλλαγή στην τιμή ονομάζονται γεγονότα ή events. Οι καταναλωτές συνήθως ενημερώνονται για τα γεγονότα την προηγούμενη μέρα (από εφημερίδες, social media, μηνύματα στο κινητό κ.α.) και κινητοποιούνται να μειώσουν την κατανάλωσή τους κατά τη διάρκεια ενός γεγονότος από τις πολύ υψηλές τιμές, που είναι συνήθως υψηλότερες συγκριτικά με τις τιμές εκτός αιχμής απ' ό,τι είναι σε ένα πρόγραμμα TOU [9].

Η αρχή και το τέλος ενός event καθορίζονται από τον aggregator και συνήθως είναι περιορισμένες από τους όρους του προγράμματος, όπως και ο συνολικός αριθμός των γεγονότων που μπορούν να συμβούν σε ένα χρόνο [2]. Για παράδειγμα, η Dominion Virginia Power, που εφαρμόζει ένα πρόγραμμα CPP στις ΗΠΑ, μπορεί να πραγματοποιήσει μέχρι 25 CPP γεγονότα το χρόνο και αυτά τα γεγονότα δεν μπορούν να διαρκούν περισσότερο από 5 ώρες [2].

Από τα παραπάνω βλέπουμε ότι ένα πρόγραμμα CPP δεν είναι λύση για απόκριση ζήτησης σε καθημερινή βάση [3], αλλά σε εξαιρετικές περιπτώσεις που η κατανάλωση είναι πολύ μεγάλη ή οι συνθήκες τέτοιες που κινδυνεύει η αξιοπιστία του συστήματος. Για να αντιμετωπίσουν αυτή την αδυναμία, οι aggregators συχνά συνδυάζουν τα CPP γεγονότα με διαφορετικές στάθμες τιμών τις υπόλοιπες ώρες, όταν δεν υπάρχει κάποιο γεγονός. Με μία τέτοια λύση ουσιαστικά έχουμε ένα CPP πρόγραμμα να τρέχει «πάνω» από ένα TOU.

Από τη μεριά ενός καταναλωτή, τα CPP προγράμματα είναι γενικά εύκολα στην εφαρμογή, αφού δεν απαιτούν προσπάθεια σε καθημερινό επίπεδο, παρά μόνο τις φορές που υπάρχει ένα DR event. Μάλιστα έχει παρατηρηθεί ότι τα νοικοκυριά είναι πιο πρόθυμα να μειώσουν σημαντικά την κατανάλωσή τους σε ένα DR γεγονός που συμβαίνει μία στο τόσο, απ' ό,τι στις ώρες αιχμής ενός TOU προγράμματος, που απαιτεί καθημερινή ενασχόληση [3]. Ωστόσο, λόγω της περιστασιακής φύσης των CPP προγραμμάτων, οι καταναλωτές συνήθως δεν κερδίζουν σημαντική μείωση κόστους από τη συμμετοχή τους σε αυτά, ενώ αν για κάποιο λόγο δεν ενημερωθούν για κάποιο γεγονός, χρεώνονται με πολύ υψηλές τιμές.

Η περιστασιακή φύση των CPP προγραμμάτων περιορίζει το όφελος που μπορούν να προσφέρουν και για τις υπηρεσίες (εταιρίες ηλεκτρικής ενέργειας, aggregators). Επειδή δεν εφαρμόζονται καθημερινά, η μείωση που προσφέρουν στο συνολικό ενεργειακό κόστος ή τους αέριους ρύπους δεν είναι μεγάλη [3]. Ωστόσο, όσον αφορά την μείωση στην κατανάλωση τις ώρες αιχμής, τα προγράμματα CPP όταν συμβαίνει ένα γεγονός είναι πιο αποτελεσματικά από

τα προγράμματα TOU [7]. Αυτό επιβεβαιώνεται από τα αποτελέσματα προγραμμάτων που είναι ήδη σε εφαρμογή.

Στο [10] γίνεται μία περίληψη 15 πειραματικών προγραμμάτων απόκρισης ζήτησης. Ανάμεσα στα CPP προγράμματα, παρατηρήθηκε μείωση της κατανάλωσης κατά 13-20% κατά τη διάρκεια γεγονότων DR σε σπίτια που δεν υπήρχαν τεχνολογίες DLC (συστήματα που επιτρέπουν την αυτόματη μείωση της κατανάλωσης κατά τη διάρκεια ενός γεγονότος, επηρεάζοντας φορτία όπως κλιματιστικά), ενώ η μείωση ήταν ακόμα πιο μεγάλη όταν υπήρχαν τεχνολογίες DLC, από 27 μέχρι 44%. Οι συγγραφείς του [7] επιβεβαιώνουν τα παραπάνω. Υποστηρίζουν ότι όταν υπάρχουν συστήματα DLC, ένα CPP πρόγραμμα μπορεί να οδηγήσει σε τουλάχιστον 30% μείωση στην κατανάλωση κατά τη διάρκεια ενός event, ενώ θεωρούν ότι παρόμοιες μειώσεις μπορούν να επιτευχθούν και χωρίς τέτοια συστήματα, εφόσον γίνει προσεκτική επιλογή των καταναλωτών που θα συμμετέχουν στο πρόγραμμα και τους προσφερθεί σημαντική υποστήριξη.

Οι παραπάνω μειώσεις στην κατανάλωση είναι αρκετά μεγαλύτερες από αυτές του 5-10%, που, όπως αναφέρθηκε στην προηγούμενη υποενότητα, αναμένονται σε TOU προγράμματα. Στο [7] ανιχνεύονται δύο πιθανοί λόγοι για αυτή τη διαφορά. Ο πρώτος έχει να κάνει με την τιμή. Γενικά, όπως αναφέρθηκε παραπάνω, η τιμή του ρεύματος κατά τη διάρκεια ενός γεγονότος CPP είναι υψηλότερη συγκριτικά με τις τιμές εκτός αιχμής απ' ότι είναι σε ένα πρόγραμμα TOU [9]. Αυτό σημαίνει ότι και οι καταναλωτές έχουν μεγαλύτερο κίνητρο να μειώσουν την κατανάλωσή τους. Ο δεύτερος λόγος αφορά τη συχνότητα των στιγμών που οι συμμετέχοντες σε ένα πρόγραμμα καλούνται να μειώσουν την κατανάλωση τους. Ενδεχομένως να είναι πιο εύκολο για έναν καταναλωτή να μειώσει απότομα την κατανάλωσή του 25-30 φορές το χρόνο, γνωρίζοντας ότι αυτό δεν θα το κάνει καθημερινά και ότι μετά θα μπορεί να επιστρέψει στις συνήθειές του, απ' ότι να αλλάξει τελείως τις συνήθειες του, όπως θα χρειαζόταν σε ένα πρόγραμμα TOU.

Συμπερασματικά, τα προγράμματα CPP είναι γενικά εύκολα στην εφαρμογή και ιδανικά για τη διασφάλιση της αξιοπιστίας και ασφάλειας του ΣΗΕ όταν προβλέπεται μεγάλη συμφόρηση, αφού μπορούν να προσφέρουν μεγάλες μειώσεις στην κατανάλωση όταν αυτό απαιτείται. Ωστόσο, επειδή δεν απαιτούν την απόκριση ζήτησης καθημερινά άλλα περιστασιακά, δεν προσφέρουν μεγάλες εξοικονομήσεις σε κόστος ή αέριους ρύπους. Για αυτούς τους σκοπούς απαιτούνται προγράμματα που έχουν καθημερινό χαρακτήρα, όπως τα TOU ή τα RTP (Real Time Pricing), που θα παρουσιαστούν παρακάτω.

#### 4.1.3 Real Time Pricing (RTP)

Στα προγράμματα Real Time Pricing η τιμή του ρεύματος αλλάζει μέσα στη μέρα ανάλογα με την χονδρική τιμή του ρεύματος και τις συνθήκες που επικρατούν στο δίκτυο. Σκοπός τους είναι να κινητοποιούν τους καταναλωτές να μειώνουν την κατανάλωσή τους τις ώρες που η ζήτηση στο δίκτυο είναι υψηλή και η παραγωγή ακριβή, κάτι που αντικατοπτρίζεται στις υψηλές τιμές χονδρικής.

Οι τιμές σε ένα RTP πρόγραμμα συνήθως ανακοινώνονται στους καταναλωτές μία ώρα πριν, αλλά υπάρχουν και προγράμματα στα οποία ανακοινώνονται την προηγούμενη μέρα [3]. Ειδικά στην πρώτη περίπτωση, που ένας καταναλωτής δεν έχει το χρόνο να προγραμματίσει το πότε μέσα στη μέρα θα χρησιμοποιήσει κάθε συσκευή, είναι πολύ σημαντικό να υπάρχει κάποιος μηχανισμός ώστε οι τιμές να ανακοινώνονται άμεσα αλλά και κάποιος έξυπνος ελεγκτής (Energy Management Controller – EMC) που αντιδρά στις διαφορετικές τιμές και ανάλογα και με τις προτιμήσεις του καταναλωτή ελέγχει τα διάφορα φορτία στο κτήριο.

Τα προγράμματα RTP επιτρέπουν στους παρόχους να μεταβάλλουν την τιμή πώλησης του ρεύματος ώστε να αντικατοπτρίζει καλύτερα τις συνθήκες στο δίκτυο και την σχέση παραγωγής – ζήτησης της ηλεκτρικής ενέργειας [3]. Αυτό είναι ιδιαίτερα συμφέρον τόσο για τον πάροχο όσο και για το κοινωνικό σύνολο όταν οδηγεί σε μείωση της συνολικής κατανάλωσης τις ώρες αιχμής, που το ρεύμα είναι γενικά πιο ακριβό και ρυπογόνο.

Ωστόσο, η εφαρμογή προγραμμάτων RTP για οικιακούς καταναλωτές είναι ακόμα πολύ περιορισμένη. Υπάρχουν δύο κύριοι λόγοι για αυτό. Από τη μία πολλά σπίτια δεν διαθέτουν τον εξοπλισμό για να μπορούν να ακολουθήσουν ένα πρόγραμμα RTP. Η χρήση έξυπνων συσκευών και ελεγκτών είναι ακόμα περιορισμένη, ενώ σε πολλές περιοχές τα σπίτια δεν διαθέτουν καν έξυπνους μετρητές. Παράλληλα, οι περισσότεροι οικιακοί καταναλωτές προτιμούν να αποφεύγουν επενδύσεις που ενέχουν ρίσκο, και βλέπουν την υποχρέωση να ανταποκρίνονται σε real time σήματα τιμής ως βάρος [11]. Αυτό είναι λογικό δεδομένου ότι χωρίς αυτόματους ελεγκτές και έξυπνες συσκευές η χειροκίνητη προσαρμογή της κατανάλωσης σε τιμές που αλλάζουν τόσο συχνά είναι δύσκολη, πολύπλοκη και χρονοβόρα.

Παρόμοια είναι τα εμπόδια για τους εμπορικούς καταναλωτές, με τη διαφορά ότι εκείνοι έχουν γενικά περισσότερους πόρους για να επενδύσουν στις σχετικές τεχνολογίες και για να αναλύσουν την οικονομική βιωσιμότητα της επένδυσης. Παρόλα αυτά, όπως θα δούμε στην ενότητα 4.2, η μικρή μεταβλητότητα των εμπορικών φορτίων υποδεικνύει ότι μπορεί να είναι πιο συμφέρον να ενταχθούν σε ένα πρόγραμμα με συγκεκριμένη τιμολόγηση κάθε μέρα (TOU) απ' ότι σε ένα πρόγραμμα με δυναμική τιμολόγηση (RTP).

Επειδή η δυναμική τιμολόγηση έχει ακόμα μικρή εφαρμογή δεν υπάρχουν πολλά αποτελέσματα από πραγματικά προγράμματα για να εκτιμήσουμε την επίδρασή της στην κατανάλωση τις ώρες αιχμής, όπως υπήρχαν για τα προγράμματα TOU και CPP. Δύο πιλοτικά προγράμματα που εφαρμόστηκαν σε οικιακούς καταναλωτές στην Washington [12] και στο Chicago [13] το 2007 και το 2009 αντίστοιχα, έδωσαν ελπιδοφόρα αποτελέσματα. Στο πρώτο εγκαταστάθηκαν στα σπίτια των καταναλωτών ελεγκτές που τους επέτρεπαν να καθορίσουν αυτόματες προσαρμογές της θέρμανσης ανάλογα με την τιμή του ρεύματος και τη θερμοκρασία. Αυτά τα συστήματα, μαζί με κάποιες πηγές ευελιξίας στο τοπικό δίκτυο (αντλίες ύδρευσης, γεννήτριες) κατάφεραν να αποδώσουν 5-20% μείωση στη κατανάλωση τις ώρες αιχμής (ανάμεσα στους συμμετέχοντες του προγράμματος). Στο δεύτερο πρόγραμμα δεν δόθηκαν στους καταναλωτές αυτόματοι ελεγκτές αλλά κάποιες λάμπες που άλλαζαν χρώμα ανάλογα με το επίπεδο τιμής του ρεύματος. Ο έλεγχος των διάφορων συσκευών στο σπίτι ήταν

δική τους ευθύνη. Παρόλα αυτά, το αποτέλεσμα ήταν και πάλι σημαντικό: μείωση στην κατανάλωση τις ώρες αιχμής που κυμαίνονταν στο διάστημα 5-14%.

## 4.2 Επιλογή συμμετεχόντων για προγράμματα DR χρησιμοποιώντας δεδομένα από έξυπνους μετρητές

Πριν από την εφεύρεση και την ευρεία εγκατάσταση έξυπνων μετρητών οι καταναλωτές επιλέγονταν για προγράμματα απόκρισης ζήτησης με βάση τη συνολική τους κατανάλωση. Αυτοί που είχαν τη μεγαλύτερη κατανάλωση ήταν και αυτοί που τελικά συμμετείχαν σε αυτά τα προγράμματα, γιατί σε γενικές γραμμές όσο πιο μεγάλη ήταν η κατανάλωσή τόσο μεγαλύτερη ήταν και η ευελιξία που μπορούσαν να παρέχουν στο δίκτυο [14]. Αυτό είχε ως αποτέλεσμα τα περισσότερα προγράμματα απόκρισης ζήτησης να απευθύνονται σε βιομηχανικά ή μεγάλα εμπορικά φορτία.

Πλέον, οι έξυπνοι μετρητές μετρούν και καταγράφουν την κατανάλωση των φορτίων με συχνότητα λεπτών ή και δευτερολέπτων και αυτό δίνει στους υπεύθυνους για προγράμματα DR πληροφορία για το πως η κατανάλωση ενός φορτίου αυξομειώνεται μέσα στη μέρα, καθώς και πως διαφέρει από μέρα σε μέρα και από εποχή σε εποχή. Αυτή η πληροφορία επιτρέπει στους αναλυτές των έξυπνων δικτύων να κάνουν πιο λεπτομερείς αναλύσεις και να επιλέγουν τους συμμετέχοντες για προγράμματα DR με μεγαλύτερη σιγουριά για την ευελιξία που μπορούν να προσφέρουν στο δίκτυο.

Δύο χαρακτηριστικά που λαμβάνονται υπόψη στην επιλογή καταναλωτών για προγράμματα DR είναι το πως μεταβάλλεται η κατανάλωσή τους μέσα στη μέρα, άρα το σχήμα της γραφικής παράστασης της κατανάλωσής τους συναρτήσει του χρόνου, και η εντροπία ή μεταβλητότητα στην κατανάλωση. Το πως αυτά τα χαρακτηριστικά επηρεάζουν τη διαδικασία επιλογής αναλύεται παρακάτω.

### 4.2.1 Ανάλογα με το σχήμα της κατανάλωσης

Η γραφική παράσταση της κατανάλωσης ενός φορτίου συναρτήσει του χρόνου για την περίοδο μίας μέρας ονομάζεται ημερήσιο προφίλ φορτίου (daily load profile). Ένα Στις περισσότερες μελέτες για την επιλογή φορτίων για προγράμματα απόκρισης ζήτησης γίνεται συσταδοποίηση πάνω στα ημερήσια προφίλ, ώστε να χωριστούν οι καταναλωτές σε ομάδες με προφίλ διαφορετικού σχήματος (πχ. [15], [16]). Στη συνέχεια με βάση τα σχήματα μπορούν να γίνουν κάποιες υποθέσεις για την ευελιξία που μπορεί να προσφέρει η κάθε ομάδα στο σύστημα.

Το πιο σημαντικό χαρακτηριστικό που συνήθως διαχωρίζει τα σχήματα μεταξύ τους είναι η ύπαρξη χαρακτηριστικών κορυφών (peaks) μέσα στη μέρα. Το που βρίσκονται οι κορυφές στο προφίλ ενός καταναλωτή είναι σημαντικό στοιχείο για το αν θα επιλεγεί για ένα πρόγραμμα DR. Ο λόγος είναι ότι τις ώρες που υπάρχουν κορυφές ο καταναλωτής επιβαρύνει περισσότερο το δίκτυο σε σχέση με άλλες ώρες μέσα στη μέρα. Επομένως αλλαγές στη συμπεριφορά του

τις ώρες που κανονικά θα είχε κορυφή μπορούν να προσφέρουν μεγαλύτερη μείωση στο φόρτο του συστήματος απ' ότι άλλες ώρες, που η κατανάλωσή του είναι γενικά μικρή.

#### 4.2.2 Ανάλογα με την εντροπία

Όταν αναφερόμαστε στη μεταβλητότητα ενός φορτίου, εννοούμε το κατά πόσο η καταναλωτική συμπεριφορά του φορτίου είναι σταθερή στο χρόνο και εύκολα προβλέψιμη (οπότε λέμε ότι έχει μικρή μεταβλητότητα) ή άστατη και δύσκολα προβλέψιμη (οπότε λέμε ότι έχει μεγάλη μεταβλητότητα). Ένας άλλος όρος που χρησιμοποιείται συχνά αντί της μεταβλητότητας είναι η εντροπία.

Στη βιβλιογραφία υπάρχουν διάφοροι τρόποι μέτρησης της εντροπίας. Στο [17] γίνεται διαχωρισμός των ημερήσιων προφίλ από τα φορτία του συνόλου δεδομένων με χρήση k-means. Ένα φορτίο του οποίου τα ημερήσια προφίλ ανήκουν σε πολλές διαφορετικές συστάδες θεωρείται πιο απρόβλεπτο και άρα έχει υψηλή εντροπία. Στο [18] και στο [19] αυτή η λογική εκφράζεται μαθηματικά με το παρακάτω άθροισμα:

$$S_n = \sum_{i=1}^K (p(C_i) * \log(p(C_i))),$$

όπου K ο αριθμός των συστάδων, C<sub>i</sub> το κέντρο της συστάδας i και p(C<sub>i</sub>) η πιθανότητα μία καμπύλη φορτίου από το φορτίο n να ανήκει στην συστάδα με κέντρο C<sub>i</sub>.

Γενικά οι έρευνες έχουν δείξει [20] ότι καταναλωτές με μικρή εντροπία είναι κατάλληλοι για incentive-based προγράμματα (πχ. με direct control) γιατί είναι πιο εύκολο να προβλεφθεί και να προγραμματιστεί η κατανάλωσή τους. Απ' την άλλη, καταναλωτές με μεγάλη εντροπία ταιριάζουν σε price based προγράμματα, γιατί έχουν μεγαλύτερη ευελιξία ως προς την κατανάλωσή τους και μπορούν να πάρουν real time αποφάσεις πιο εύκολα.

Με βάση τα παραπάνω μπορούμε να κάνουμε την υπόθεση ότι και μεταξύ των price based προγραμμάτων, αυτά που ακολουθούν την ίδια πολιτική τιμολόγησης κάθε μέρα (προγράμματα TOU) θα ευνοούν περισσότερο καταναλωτές με μικρή μεταβλητότητα, γιατί μπορούν, ευκολότερα από τους καταναλωτές με μεγάλη μεταβλητότητα, να προγραμματίσουν αναλυτικά τις εργασίες τους ανάλογα με την τιμή και να μετά να ακολουθούν αυτό τον προγραμματισμό κάθε μέρα. Απ' την άλλη, προγράμματα με περιστασιακό χαρακτήρα (προγράμματα CPP) ή που η τιμή αλλάζει από ώρα σε ώρα (προγράμματα RTP) θα ευνοούν περισσότερο τους καταναλωτές με μεγάλη μεταβλητότητα, γιατί είναι πιο ευέλικτοι στην καταναλωτική τους συμπεριφορά από αυτούς με μικρή μεταβλητότητα. Ωστόσο, χρειάζεται να διεξαχθούν περισσότερα προγράμματα και να συλλεχθούν περισσότερα δεδομένα από τον πραγματικό κόσμο ώστε να ελεγχθούν οι παραπάνω υποθέσεις.

### 4.3 Επιλογή συμμετεχόντων για προγράμματα DR με βάση τη χρήση

Στην ενότητα 4.2 εξηγήθηκε το πως δύο χαρακτηριστικά που μπορούν να εξαχθούν από τις μετρήσεις κατανάλωσης ενός φορτίου, το σχήμα του προφίλ κατανάλωσης και η εντροπία, επηρεάζουν την ευελιξία που μπορεί να προσφέρει το φορτίο σε ένα πρόγραμμα απόκρισης ζήτησης και ποιες δομές προγραμμάτων (TOU, RTP κλπ.) του ταιριάζουν.

Σε αυτή την ενότητα διερευνάται το πως επηρεάζει τα παραπάνω ένα τρίτο χαρακτηριστικό: η χρήση του φορτίου, δηλαδή ο σκοπός για τον οποίο χρησιμοποιείται. Έμφαση δίνεται στα εμπορικά και τα οικιακά φορτία, γιατί αυτά αφορά και το πρακτικό κομμάτι αυτής της εργασίας. Θα δούμε πως το αν ένα φορτίο είναι εμπορικό ή οικιακό επηρεάζει τα χαρακτηριστικά της ενότητας 4.2 (το σχήμα του προφίλ κατανάλωσης και την εντροπία) καθώς και άλλα χαρακτηριστικά που είναι χρήσιμα για την διερεύνηση του αν ένας καταναλωτής είναι καλός υποψήφιος για ένα πρόγραμμα απόκρισης ζήτησης, όπως το συνολικό μέγεθος της κατανάλωσης και το μείγμα της κατανάλωσης (ποσοστό της κατανάλωσης για κάθε τελική χρήση, όπως θέρμανση, φωτισμό κλπ.).

#### 4.3.1 Εμπορικά φορτία

Ο όρος εμπορικά φορτία (Commercial loads) αναφέρεται στις ηλεκτρικές εγκαταστάσεις επιχειρήσεων που παρέχουν υπηρεσίες. Τέτοια φορτία είναι πχ. τα κτήρια εμπορικών καταστημάτων, γραφεία, ξενοδοχεία, εστιατόρια κλπ. [21] Παρακάτω παρουσιάζονται 4 κριτήρια που πρέπει να λαμβάνονται υπόψη κατά το σχεδιασμό προγραμμάτων απόκρισης ζήτησης για εμπορικά φορτία.

**Σχήμα της κατανάλωσης:** επειδή τα φορτία στα εμπορικά κτήρια (θέρμανση, φωτισμός κλπ) λειτουργούν περισσότερο τις ώρες εργασίας, το τυπικό σχήμα του προφίλ κατανάλωσης ενός εμπορικού κτηρίου εμφανίζει τη μεγαλύτερη κατανάλωση το πρωί και το απόγευμα. Η κατανάλωση εκτός των ωρών εργασίας είναι γενικά μικρότερη [22]. Το ίδιο ισχύει και για την κατανάλωση τα Σαββατοκύριακα [22], αφού τότε πολλές επιχειρήσεις κλείνουν ή λειτουργούν με μειωμένο ωράριο και λιγότερο προσωπικό. Αυτό έχει ως αποτέλεσμα η μείωση στην συνολική κατανάλωση που μπορεί να προσφέρει ένα εμπορικό κτήριο τις καθημερινές και τις ώρες εργασίας να είναι μεγαλύτερη από τη μείωση που μπορεί να προσφέρει τα Σαββατοκύριακα και εκτός των ωρών εργασίας. Αυτή η πραγματικότητα πρέπει να λαμβάνεται υπόψη κατά την επιλογή καταναλωτών για ένα πρόγραμμα DR που αποσκοπεί στην μεταβολή της ζήτησης συγκεκριμένες ώρες της ημέρας και μέρες της εβδομάδας.

**Εντροπία:** γενικά τα εμπορικά φορτία είναι λιγότερο μεταβλητά από τα οικιακά [23]. Αυτό σημαίνει ότι ενδεχομένως είναι πιο εύκολο για έναν aggregator να προβλέψει την καταναλωτική τους συμπεριφορά τους και άρα και την ευελιξία που μπορούν να προσφέρουν στο δίκτυο σε ένα DR γεγονός. Επίσης, με βάση τα συμπεράσματα της ενότητας 4.2.2, τα εμπορικά φορτία είναι πιο ιδανικά από τα οικιακά για προγράμματα TOU και λιγότερο για προγράμματα RTP και CPP.

**Μέγεθος της κατανάλωσης:** Όσο μεγαλύτερη είναι συνολικά η κατανάλωση ενός κτηρίου, τόσο μεγαλύτερο είναι δυνητικά το όφελος που μπορεί να προσφέρει στο δίκτυο, εάν ενταχθεί σε ένα πρόγραμμα απόκρισης ζήτησης. Ως παράδειγμα, ας θεωρήσουμε ότι ένα μεγάλο πανεπιστήμιο και ένα μικρό εστιατόριο θέλουν να συμμετέχουν σε ένα πρόγραμμα απόκρισης ζήτησης. Εάν υποθέσουμε ότι τα δύο φορτία είναι παρόμοια σε όλους τους παράγοντες που επηρεάζουν την ελαστικότητά τους πέρα από το μέγεθος (σχήμα κατανάλωσης, ποσοστό ευέλικτου φορτίου κλπ.), ο διαχειριστής του προγράμματος θα είχε μεγαλύτερο όφελος από την ένταξη του μεγαλύτερου φορτίου (δηλαδή του πανεπιστημίου) στο πρόγραμμα, γιατί θα μπορούσε να προσφέρει μεγαλύτερη μείωση στην συνολική κατανάλωση όταν αυτό χρειάζεται. Αυτός είναι ο λόγος που πριν να γίνουν διαθέσιμα δεδομένα από έξυπνους μετρητές οι καταναλωτές επιλέγονταν για προγράμματα απόκρισης ζήτησης πρώτα με βάση τη συνολική τους κατανάλωση [14].

**Μείγμα κατανάλωσης:** Όσον αφορά το ποσοστό της ενέργειας που καταναλώνεται για κάθε τελική χρήση, προγράμματα απόκρισης ζήτησης για εμπορικά κτήρια έχουν δείξει ότι το μεγαλύτερο μέρος της ευελιξίας προέρχεται από τα συστήματα θέρμανσης, ψύξης, εξαερισμού και φωτισμού [24]. Όσο περισσότερη ενέργεια καταναλώνει ένα κτήριο για αυτές τις χρήσεις, τόσο περισσότερη ευελιξία θα μπορούσε θεωρητικά να προσφέρει.

Τα παραπάνω κριτήρια σε συνδυασμό με τις δυνατότητες για τοπική παραγωγή και αποθήκευση ενέργειας χρησιμοποιήθηκαν στο [25], σε στατιστικά για πάνω από 80 εκατομμύρια καταναλωτές στην Ευρώπη, με σκοπό την εύρεση ομάδων καταναλωτών που θα είχαν το μεγαλύτερο όφελος από τεχνολογίες DER (Distributed Energy Resources), δηλαδή κατανεμημένης παραγωγής (distributed production), κατανεμημένης αποθήκευσης (distributed storage) και απόκρισης ζήτησης (DR). Οι συγγραφείς κατέληξαν ότι ανάμεσα στα εμπορικά κτήρια πιο ιδανικά για συμμετοχή σε προγράμματα απόκρισης ζήτησης ήταν: πανεπιστήμια και σχολεία, νοσοκομεία, εμπορικά κέντρα και γραφεία.

### 4.3.2 Οικιακά φορτία

Τα οικιακά φορτία έχουν γενικά μικρότερη ποσότητα κατανάλωσης ανά εγκατάσταση σε σχέση με τα εμπορικά ή τα βιομηχανικά. Αυτό σημαίνει ότι προγράμματα DR με οικιακούς καταναλωτές χρειάζονται πολλούς περισσότερους συμμετέχοντες από προγράμματα με εμπορικούς ή βιομηχανικούς καταναλωτές για να προσφέρουν το ίδιο όφελος στο δίκτυο, γεγονός που τα καθιστά πιο πολύπλοκα στη διαχείριση. Ωστόσο, δεδομένου ότι τα οικιακά φορτία έχουν τη μεγαλύτερη ευθύνη για το φαινόμενο της κορύφωσης της ζήτησης τις ώρες αιχμής [26], η επιτυχής ένταξη τους σε προγράμματα DR μπορεί να επιφέρει μεγάλα οικονομικά και περιβαλλοντικά οφέλη [27].

Παρακάτω παρουσιάζονται κάποια κριτήρια που πρέπει να λαμβάνονται υπόψη κατά το σχεδιασμό προγραμμάτων απόκρισης ζήτησης για οικιακά φορτία.

**Σχήμα της κατανάλωσης:** το σχήμα του προφίλ κατανάλωσης ενός σπιτιού εξαρτάται από τον τρόπο ζωής των κατοίκων του. Σπίτια με εργαζόμενους εμφανίζουν γενικά μεγαλύτερη κατανάλωση τις πρωινές και τις απογευματινές ώρες, δηλαδή πριν και μετά τις ώρες εργασίας, ενώ σπίτια με συνταξιούχους ή μικρά παιδιά που μένουν στο σπίτι τις περισσότερες ώρες της ημέρας μπορεί να εμφανίζουν πολύ διαφορετική συμπεριφορά. Περισσότερες πληροφορίες για το πως ο τρόπος ζωής των κατοίκων ενός σπιτιού επηρεάζει την κατανάλωση υπάρχουν στα άρθρα [15], [28].

**Εντροπία:** τα οικιακά φορτία είναι συνήθως πιο μεταβλητά από τα εμπορικά ή τα βιομηχανικά φορτία, πράγμα που σημαίνει ότι είναι πιο δύσκολο η κατανάλωση ενός σπιτιού να προβλεφθεί με ακρίβεια [23]. Ωστόσο μεγαλύτερη μεταβλητότητα σημαίνει ότι τα οικιακά φορτία είναι και πιο ευέλικτα στην καταναλωτική τους συμπεριφορά από εμπορικά κτήρια ή βιομηχανίες όπου η παραγωγική διαδικασία θέτει πολλούς περιορισμούς στην ποσότητα και τον χρονισμό της κατανάλωσης. Με βάση τα συμπεράσματα της ενότητας 4.2.2, αυτό σημαίνει ότι τα οικιακά φορτία είναι πιο ιδανικά από τα εμπορικά για προγράμματα RTP και CPP και λιγότερο για προγράμματα TOU.

**Μείγμα κατανάλωσης:** έρευνες έχουν δείξει ότι τα πιο ευέλικτα φορτία σε μία κατοικία, όπως και σε ένα εμπορικό κτήριο, είναι τα συστήματα θέρμανσης, ψύξης και εξαερισμού (HVAC systems), ακολουθούμενα από τους θερμοσίφωνες και κάποιες smart συσκευές που μπορούν να τεθούν σε λειτουργία οποιαδήποτε ώρα της ημέρας (πχ. smart πλυντήρια). Όσο περισσότερη ενέργεια καταναλώνει ένα σπίτι για αυτές τις χρήσεις, τόσο περισσότερη ευελιξία θα μπορούσε θεωρητικά να προσφέρει.



## 5. Βιβλιογραφική Ανασκόπηση – Συσταδοποίηση σε δεδομένα από ηλεκτρικά φορτία

Βασικό κομμάτι της παρούσας εργασίας είναι η εφαρμογή αλγορίθμων συσταδοποίησης (clustering) σε δεδομένα κατανάλωσης και παραγωγής ηλεκτρικής ενέργειας. Η συσταδοποίηση είναι μία διαδικασία που μπορεί να χωριστεί σε 4 στάδια: την προεπεξεργασία των δεδομένων (preprocessing), την μείωση του μεγέθους των δεδομένων και την επιλογή των χαρακτηριστικών πάνω στα οποία θα γίνει η συσταδοποίηση (feature selection), το στάδιο που χωρίζονται τα δεδομένα (clustering stage) και τέλος τον έλεγχο της επίδοσης της συσταδοποίησης (validation stage)[20]. Παρακάτω παρουσιάζονται συνοπτικά οι μέθοδοι και τα εργαλεία που συναντώνται στη βιβλιογραφία, για κάθε ένα από τα παραπάνω στάδια, με έμφαση σε δεδομένα από ηλεκτρικά φορτία.

### 5.1 Προεπεξεργασία Δεδομένων

Τα δεδομένα που προέρχονται από έξυπνους μετρητές δεν είναι πάντα αξιόπιστα. Συχνά έχουν επηρεαστεί από εξωγενείς παράγοντες όπως θόρυβο ή περιόδους που λόγω καιρού ή βλάβης η επικοινωνία με τη βάση δεδομένων έχει κοπεί. Ως αποτέλεσμα, τα δεδομένα συχνά περιέχουν κενές τιμές (null values), ή τιμές που λόγω θορύβου είναι πολύ μεγάλες σε σχέση με τις υπόλοιπες (απομακρυσμένες τιμές - outliers). Αν τα δεδομένα δοθούν σε κάποιο αλγόριθμο συσταδοποίησης σε αυτή τη μορφή, τα αποτελέσματα θα είναι αλλοιωμένα. Για τον παραπάνω λόγο συχνά το πρώτο βήμα σε μία διαδικασία συσταδοποίησης είναι ο καθαρισμός των δεδομένων [20]

#### 5.1.1 Διαχείριση κενών τιμών

Για τη διαχείριση των κενών στα περισσότερα άρθρα επιλέγεται η διαγραφή των ημερών που έχουν κενές τιμές (πχ. [15] ή που έχουν συνεχόμενες κενές τιμές πάνω από ένα όριο (πχ. [29]).

Αν ωστόσο οι κενές τιμές στο σύνολο δεδομένων είναι πολλές ή η συμπλήρωση των κενών απαραίτητη για άλλους λόγους, μπορεί να γίνει αντικατάστασή τους με μεθόδους data imputation. [30], [31]

### 5.2 Μείωση διαστατικότητας και επιλογή χαρακτηριστικών

Οι περισσότεροι αλγόριθμοι συσταδοποίησης δυσκολεύονται να διαχωρίσουν τα δείγματα όσο περισσότερα είναι τα χαρακτηριστικά τους. Αν κάθε δείγμα αναπαρίσταται με ένα διάνυσμα, διανύσματα με περισσότερες διαστάσεις οδηγούν σε μεγαλύτερη πολυπλοκότητα, υψηλότερους χρόνους εκτέλεσης και συχνά συστάδες που δεν διαφέρουν ικανοποιητικά μεταξύ τους [32]. Αυτό το πρόβλημα συχνά αναφέρεται στη βιβλιογραφία ως «κατάρρα της διαστατικότητας» (curse of dimensionality) [33].

Παρακάτω παρουσιάζονται κάποιες μέθοδοι που μπορούν να μειώσουν τις διαστάσεις δεδομένων ηλεκτρικής ενέργειας για εφαρμογές συσταδοποίησης, οδηγώντας σε γρηγορότερους χρόνους εκτέλεσης και καλύτερα αποτελέσματα.

### 5.2.1 Επιλογή συχνότητας δειγματοληψίας

Σε προβλήματα συσταδοποίησης για ηλεκτρικά φορτία, η συχνότερη περίπτωση είναι τα χαρακτηριστικά που δίνονται στον αλγόριθμο να είναι μετρήσεις κατανάλωσης ηλεκτρικής ενέργειας μέσα στη μέρα. Αυτές μπορεί να είναι ανά ώρα, δίνοντας δείγματα με 24 διαστάσεις, ανά μισή ώρα, δίνοντας σε δείγματα με 48 διαστάσεις, ανά δεκαπεντάλεπτο, δίνοντας δείγματα με 92 διαστάσεις και ούτω καθ' εξής.

Η επιλογή της συχνότητας που θα εμφανίζονται τα δείγματα περιορίζεται από τη μία από τη μετρητή που χρησιμοποιείται, αλλά σε περίπτωση μετρητή που λαμβάνει μετρήσεις με μεγάλη συχνότητα (πχ. ανά λεπτό), η συχνότητα που θα έχουν τα δεδομένα που θα δοθούν στον αλγόριθμο συσταδοποίησης, και επομένως οι διαστάσεις των διανυσμάτων εισόδου, μπορεί να αποφασιστεί από τον αναλυτή. Αυτό γίνεται με επαναδειγματοληψία (resampling) των δεδομένων.

Η επιλογή της συχνότητας επαναδειγματοληψίας είναι κρίσιμη: όσο μεγαλώνει το σχήμα της καμπύλης κατανάλωσης γίνεται πιο ακριβές, αλλά οι διαστάσεις πληθαίνουν και η γενίκευση με το διαχωρισμό των δεδομένων σε συστάδες γίνεται δυσκολότερη.

### 5.2.2 Εξαγωγή χαρακτηριστικών (Feature Extraction)

Η μέθοδοι εξαγωγής χαρακτηριστικών (feature extraction methods) απεικονίζουν το χώρο των δειγμάτων σε έναν άλλο χώρο με μικρότερες διαστάσεις, με τρόπο που όσο γίνεται διατηρεί την πληροφορία που είναι σημαντική για τη συσταδοποίηση. Αυτό μπορεί να γίνει με δύο τρόπους: την εξαγωγή χαρακτηριστικών με βάση τη γνώση ειδικών στον τομέα της αγοράς ηλεκτρικής ενέργειας (feature definition based on expert knowledge) ή τη χρήση μεθόδων συμπίεσης δεδομένων (data compression) που χρησιμοποιούνται γενικότερα σε προβλήματα ανάλυσης δεδομένων[20].

Κάποιες μέθοδοι συμπίεσης που έχουν εφαρμοστεί σε δεδομένα κατανάλωσης είναι: η Principal Component Analysis (PCA) [34], που δημιουργεί νέα χαρακτηριστικά που είναι γραμμικοί συνδυασμοί των αρχικών και κρατάει αυτά με τη μεγαλύτερη διακύμανση, η Symbolic Aggregate Approximation (SAX) [35], που χωρίζει μία συνεχή χρονοσειρά σε διακριτά κομμάτια και αντικαθιστά κάθε κομμάτι με ένα σύμβολο, και η Discrete Wavelet Transform (DWT) [36]. Επίσης έχουν χρησιμοποιηθεί και autoencoders [37], δηλαδή νευρωνικά δίκτυα που μπορούν να μάθουν αποτελεσματικές κωδικοποιήσεις χρονοσειρών ή άλλων δεδομένων.

### 5.3 Μετρικές Απόστασης

Μία σημαντική απόφαση που πρέπει να ληφθεί πριν συσταδοποίηση ενός συνόλου δεδομένων είναι το ποια μετρική απόστασης θα χρησιμοποιηθεί για την σύγκριση μεταξύ των δειγμάτων. Η κάθε μετρική απόστασης έχει τα δυνατά και τα αδύναμα της σημεία, και η επιλογή συνήθως βασίζεται πάνω στη φύση των δεδομένων (πχ. αν είναι κατηγορικά ή αριθμητικά, αν είναι δεδομένα χρονοσειρά ή όχι κλπ.) και τον αλγόριθμο που θα μπορούσε να χρησιμοποιηθεί (δεν λειτουργούν όλοι οι αλγόριθμοι με όλες τις μετρικές απόστασης).

Παρακάτω αναλύονται δύο μετρικές απόστασης που μπορούν να χρησιμοποιηθούν για τη συσταδοποίηση δεδομένων από ηλεκτρικά φορτία, η ευκλείδεια απόσταση, που είναι η πιο δημοφιλής επιλογή στη βιβλιογραφία, και η απόσταση με χρήση Dynamic Time Waring, που είναι γενικά ιδανικότερη από την ευκλείδεια για συσταδοποίηση χρονοσειρών και χρησιμοποιείται στο πειραματικό κομμάτι αυτής της εργασίας.

#### 5.3.1 Ευκλείδεια απόσταση

Η ευκλείδεια απόσταση για δύο διανύσματα κατανάλωσης  $p$  και  $q$  με  $n$  διαστάσεις δίνεται από τον ίδιο τύπο που δίνει την απόσταση μεταξύ δύο σημείων σε  $n$ -διάστατο χώρο:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2}$$

Είναι η πιο δημοφιλής μετρική απόστασης για δεδομένα κατανάλωσης από ηλεκτρικά φορτία [20]. Οι κυριότερος λόγος για αυτό είναι ότι υπολογίζεται πολύ γρήγορα και χρησιμοποιείται συνήθως μαζί με τον πιο δημοφιλή αλγόριθμο συσταδοποίησης, τον K-means.

Ωστόσο, έχει το μειονέκτημα ότι δεν λαμβάνει υπόψη κάποια συσχέτιση μεταξύ των διαφορετικών τιμών ενός διανύσματος. Αυτό είναι προβληματικό για την περίπτωση των δεδομένων από ηλεκτρικά φορτία, γιατί αυτά συνήθως οργανώνονται σε χρονοσειρές, δηλαδή διανύσματα στα οποία τα διαφορετικά χαρακτηριστικά συσχετίζονται χρονικά. Στην πράξη, αυτό σημαίνει ότι η ευκλείδεια απόσταση δεν αποτυπώνει πληροφορία για χαρακτηριστικά σχήματα που μπορεί να εμφανίζονται στα προφίλ κατανάλωσης, πάρα μόνο για τη διαφορά που υπάρχει στο μέγεθος της κατανάλωσης δύο προφίλ για τις ίδιες χρονικές στιγμές.

#### 5.3.2 Απόσταση με δυναμική παραμόρφωση του χρόνου (Dynamic Time Waring)

Η απόσταση με χρήση Dynamic Time Waring (DTW) [38] χρησιμοποιείται για την μέτρηση της ομοιότητας μεταξύ χρονοσειρών. Ο αλγόριθμος που κάνει τη μέτρηση πρώτα επιτρέπει τη χρονική μετατόπιση και παραμόρφωσή δύο χρονοσειρών ώστε να είναι ιδανικά ευθυγραμμισμένες και στην συνέχεια υπολογίζει την απόσταση σε αυτή την ευθυγράμμιση.

Στη συσταδοποίηση προφίλ φορτίου, η χρονική μετατόπιση και παραμόρφωση των χρονοσειρών επιτρέπει η απόσταση με χρήση DTW να αποτυπώνει πληροφορία και για το σχήμα των προφίλ (σε αντίθεση με την ευκλείδεια απόσταση, όπως είδαμε στην ενότητα 5.3.1). Ωστόσο, στην απλή μορφή του ο αλγόριθμος αγνοεί πλήρως το χρονισμό των σχημάτων, κάτι

που στις περισσότερες εφαρμογές για τη συσταδοποίηση προφίλ φορτίου δεν είναι επιθυμητό, γιατί ο στόχος είναι να προκύψουν συστάδες με κορυφές σε διαφορετικές χρονικές στιγμές μέσα στη μέρα.

Ένα από τα βασικά μειονεκτήματα του DTW είναι ο μεγάλη υπολογιστική πολυπλοκότητα σε σχέση με άλλες μετρικές απόστασης και ειδικά την ευκλείδεια απόσταση. Αυτή η πολυπλοκότητα κάνει τη χρήση του απαγορευτική σε περιπτώσεις που το σύνολο δεδομένων είναι πολύ μεγάλο και υπάρχει ανάγκη η συσταδοποίηση να γίνεται γρήγορα. Το παραπάνω πρόβλημα μπορεί να αντιμετωπιστεί με παραλλαγές του αλγόριθμου που λειτουργούν προσεγγιστικά και με περιορισμούς στις χρονικές παραμορφώσεις (πχ. [39]). Ένας τέτοιος περιορισμός είναι το Sakoe-Chiba radius [40], που χρησιμοποιείται στο πειραματικό κομμάτι αυτής της εργασίας.

## 5.4 Αλγόριθμοι

Πολλοί διαφορετικοί αλγόριθμοι συσταδοποίησης έχουν εφαρμοστεί σε δεδομένα από ηλεκτρικά φορτία. Παρακάτω παρουσιάζονται οι βασικότεροι εξ' αυτών, χωρισμένοι σε κατηγορίες ανάλογα με τον τρόπο λειτουργίας τους.

### 5.4.1 Αλγόριθμοι που βασίζονται στη διχοτόμηση (Partition Based Algorithms)

Οι αλγόριθμοι που βασίζονται στη διχοτόμηση ξεκινούν διαχωρίζοντας τυχαία τον χώρο των δειγμάτων σε  $K$  διαφορετικές συστάδες. Στη συνέχεια προσαρμόζουν αυτές τις συστάδες επαναληπτικά, τοποθετώντας κάθε δείγμα στην κοντινότερη συστάδα και υπολογίζοντας ξανά το κέντρο της. Τελικό σκοπός είναι τα κέντρα των διαφορετικών συστάδων να έχουν μεγάλη απόσταση μεταξύ τους και τα δείγματα μέσα σε κάθε συστάδα μικρή. Στο τελικό αποτέλεσμα συνήθως παίζει σημαντικό ρόλο ο αρχικός διαχωρισμός του χώρου, επομένως είναι συμφέρον οι partition based αλγόριθμοι να τρέχουν με πολλές διαφορετικές αρχικές καταστάσεις και να επιλέγεται αυτή που οδήγησε στο καλύτερο αποτέλεσμα [20].

Ο πιο δημοφιλής partition-based αλγόριθμος γενικότερα [41] αλλά και για δεδομένα από ηλεκτρικά φορτία [20] είναι ο  $K$ -means. Αν και είναι απλός, γρήγορος αλγόριθμος και χειρίζεται καλά μεγάλα σύνολα δεδομένων με υψηλή διαστατικότητα, είναι ευαίσθητος σε απομακρυσμένες τιμές (outliers) και θορυβώδη δεδομένα [41]. Παραλλαγές του  $K$ -means, όπως ο  $K$ -medians και ο  $K$ -medoids, μπορούν να αντιμετωπίσουν αυτά τα προβλήματα με αντάλλαγμα το μεγαλύτερο υπολογιστικό κόστος [20].

Ένα πρόβλημα στη χρήση των παραπάνω αλγόριθμων είναι η ανάγκη επιλογής του αριθμού των συστάδων, δηλαδή του  $k$ , πριν την εκτέλεση τους. Αυτή είναι μια σημαντική απόφαση γιατί ένα μικρό  $k$  μπορεί να οδηγήσει σε συστάδες που δεν είναι αρκετά αντιπροσωπευτικές του συνόλου δεδομένων, ενώ ένα μεγάλο  $k$  σε αποτελέσματα που είναι πολύπλοκα και όχι αρκετά γενικά [20]. Το πρόβλημα της επιλογής του  $k$  μπορεί να λυθεί χρησιμοποιώντας τον adaptive  $k$ -means. Στον adaptive  $k$ -means ο αριθμός των συστάδων προσαρμόζεται κατά την εκτέλεση,

με βάση το παρακάτω κριτήριο: κάθε δείγμα πρέπει να απέχει το πολύ μία οριακή απόσταση  $d_{min}$  από το κέντρο της συστάδας του. Ο αλγόριθμος συγχωνεύει ή διαχωρίζει τις συστάδες ώστε αυτό το κριτήριο να ικανοποιείται για όλα τα δείγματα [42].

#### 5.4.2 Ιεραρχικοί αλγόριθμοι (Hierarchical Algorithms)

Η βασική ιδέα των ιεραρχικών αλγορίθμων είναι η τοποθέτηση του συνόλου δεδομένων σε ένα ιεραρχικό δέντρο από συστάδες. Αυτό μπορεί να γίνει με δύο τρόπους [41]. Τον σωρευτικό (agglomerative) και τον διχαστικό (divisive). Στον πρώτο τρόπο κάθε δείγμα είναι αρχικά μία συστάδα και σε κάθε επανάληψη οι δύο πιο κοντινές συστάδες ενοποιούνται μέχρι να υπάρχει μόνο μία συστάδα. Στον δεύτερο ακολουθείται η αντίστροφη διαδικασία. Όλα τα δείγματα ανήκουν αρχικά σε μία συστάδα και σε κάθε επανάληψη η πιο αραιή συστάδα διαχωρίζεται μέχρι κάθε δείγμα να ανήκει στη δική του συστάδα.

Ένα πλεονέκτημα των ιεραρχικών αλγορίθμων σε σχέση με αυτούς που βασίζονται στη διχοτόμηση είναι ότι η επιλογή του αριθμού των συστάδων δεν είναι απαραίτητη πριν την εκτέλεση του αλγορίθμου. Αφού έχει εκτελεστεί ο αλγόριθμος και υπάρχει το πλήρες δενδροδιάγραμμα, ο αναλυτής μπορεί να πάρει τα αποτελέσματα της συσταδοποίησης για όποιο αριθμό συστάδων επιθυμεί [20].

Πέρα από το παραπάνω πλεονέκτημα, οι ιεραρχικοί αλγόριθμοι είναι γενικά απλοί στον ορισμό και την κατανόηση και δεν έχουν περιορισμούς στη μετρική απόστασης που χρησιμοποιείται για τον υπολογισμό των αποστάσεων [20], κάτι που δεν ισχύει για όλους τους αλγόριθμους, όπως αναφέρεται παραπάνω για τους partition based αλγόριθμους. Από την άλλη, έχουν γενικά μεγάλη χρονική πολυπλοκότητα [20] [41] και είναι ευαίσθητοι σε απομακρυσμένες τιμές [20].

#### 5.4.3 Αλγόριθμοι που βασίζονται στην πυκνότητα (Density Based Algorithms)

Οι αλγόριθμοι που βασίζονται στην πυκνότητα λειτουργούν με τη βασική ιδέα ότι αν απεικονίσουμε τα δείγματα στο χώρο δειγμάτων, περιοχές με μεγάλη πυκνότητα (δηλαδή μεγάλη συγκέντρωση δειγμάτων) πρέπει να ανήκουν στην ίδια συστάδα [41]. Αυτοί οι αλγόριθμοι λειτουργούν καλύτερα σε δεδομένα όπου περιοχές με υψηλή πυκνότητα είναι χωρισμένες από περιοχές με χαμηλή πυκνότητα, ώστε ο διαχωρισμός των συστάδων να είναι ξεκάθαρος. Αυτή είναι γενικά οι εικόνα και όταν απεικονίζουμε δεδομένα από ηλεκτρικά φορτία στο χώρο. Επομένως η χρήση αλγορίθμων βασισμένων στην πυκνότητα σε τέτοια δεδομένα έχει διαισθητική βάση.

Οι density based αλγόριθμοι είναι γενικά αποδοτικοί χρονικά [41] και μπορούν να σχηματίσουν συστάδες που το σχήμα τους στο χώρο δειγμάτων δεν είναι κυρτό [20], πράγμα που δυσκολεύει αλγορίθμους άλλων ειδών. Ωστόσο απαιτούν μεγάλη μνήμη όταν ο όγκος των δεδομένων είναι μεγάλος [41].

Ο πιο διαδεδομένος αλγόριθμος που βασίζεται στην πυκνότητα είναι ο DBSCAN [43]. Πέρα από τα πλεονεκτήματα που αναφέρθηκαν παραπάνω, είναι πολύ καλός στο να αναλύει δεδομένα με θόρυβο και ανωμαλίες και μάλιστα μπορεί να τις ανιχνεύσει κατά την εκτέλεση. Επίσης δεν απαιτεί από το χρήστη να επιλέξει τον αριθμό των συστάδων από πριν [20], χωρίς αυτό να σημαίνει ότι η παραμετροποίησή του είναι εύκολη. Η επιλογή τιμών για τις δύο πιο σημαντικές παραμέτρους, την ακτίνα της κάθε «γειτονιάς» και τον ελάχιστο αριθμό δειγμάτων σε κάθε γειτονιά επηρεάζει σε μεγάλο βαθμό τα αποτελέσματα [41].

Μία εναλλακτική που δεν απαιτεί την ίδια προσοχή στον ορισμό των παραμέτρων και δίνει καλύτερα αποτελέσματα από τον DBSCAN όταν οι πυκνότητες των συστάδων στο χώρο δειγμάτων δεν είναι ομοιόμορφες [20] είναι ο CFSFDP [44]. Στο συγκεκριμένο αλγόριθμο είναι απαραίτητος ο ορισμός του αριθμού των συστάδων πριν την εκτέλεση [41].

## 5.5 Μετρικές Επικύρωσης

Σε αυτή την ενότητα παρουσιάζονται κάποιες μετρικές επικύρωσης που χρησιμοποιούνται στη βιβλιογραφία για τον έλεγχο της συσταδοποίησης δεδομένων από ηλεκτρικά φορτία.

### 5.5.1 Silhouette Score

Το Silhouette Score [45] μετράει το πόσο παρόμοιο είναι ένα αντικείμενο με τα αντικείμενα της συστάδας στην οποία ανήκει σε σχέση με τα αντικείμενα της πιο κοντινής συστάδας στην οποία δεν ανήκει. Για ένα αντικείμενο  $i$  του συνόλου δεδομένων δίνεται από τον τύπο:

$$S_i = (b_i - a_i) / \max(b_i, a_i),$$

όπου  $a_i$  είναι ο μέσος όρος της απόστασης του  $i$  από τα αντικείμενα της συστάδας στην οποία ανήκει και  $b_i$  ο μέσος όρος της απόστασής του από τα αντικείμενα της πιο κοντινής συστάδας. Για την μέτρηση της απόστασης μπορεί να χρησιμοποιηθεί οποιαδήποτε μετρική, όπως ευκλείδεια ή DTW (ενότητα 5.3).

Από τον τύπο μπορεί να αποδειχτεί ότι το Silhouette Score παίρνει τιμές από -1 έως 1. Όσο μεγαλύτερη είναι η τιμή του, τόσο «καλύτερη» είναι και η τοποθέτηση του αντικειμένου  $i$  στη συστάδα του σε σχέση με την κοντινότερη συστάδα. Το συνολικό Silhouette Score για μία εφαρμογή συσταδοποίησης ισούται με το μέσο όρο των επιμέρους σκορ για κάθε αντικείμενο. Συνολικό Silhouette Score κοντά στο 1 σημαίνει ότι οι συστάδες που έχουν δημιουργηθεί είναι πυκνές και καλά διαχωρισμένες.

### 5.5.2 Davies – Bouldin Index (DBI)

Ο δείκτης Davies – Bouldin (DBI) είναι μία μετρική επικύρωσης για εφαρμογές συσταδοποίησης που παρουσιάστηκε για πρώτη φορά στο [46].

Ισούται με τον μέσο όρο επί όλων των συστάδων του λόγου του διασκορπισμού (scatter) εντός της συστάδας  $K$  προς την απόστασή της από την κοντινότερη συστάδα. Όσο πιο χαμηλή είναι

η τιμή του τόσο πιο κοντά είναι τα αντικείμενα της ίδιας συστάδας (χαμηλός διασκορπισμός) και τόσο καλύτερα διαχωρισμένες είναι οι συστάδες μεταξύ τους, επομένως χαμηλότερες τιμές υποδεικνύουν καλύτερη συσταδοποίηση.

### 5.5.3 Peak Match Score

Το Peak Match Score παρουσιάζεται για πρώτη φορά στο [29] σε μία προσπάθεια να δημιουργηθεί μία μετρική επικύρωσης προσαρμοσμένη στο πρόβλημα της συσταδοποίησης ηλεκτρικών φορτίων για απόκριση ζήτησης, όπου είναι σημαντικό οι καμπύλες κατανάλωσης να χωρίζονται σε συστάδες ανάλογα με το πότε εμφανίζουν χαρακτηριστικές ακμές, δηλαδή οι ακμές των κέντρων των συστάδων να συγχρονίζονται με τις ακμές των δειγμάτων που ανήκουν σε αυτές.

Ορίζεται ως ο μέσος όρος των επιμέρους σκορ  $m_i$  για κάθε δείγμα  $i$  του συνόλου δεδομένων:

$$peakMatchScore = \frac{1}{N} \sum_{i=1}^N m_i$$

Όπου  $N$  ο αριθμός των δειγμάτων και όπου το επιμέρους σκορ  $m_i$ , για ημερήσιες χρονοσειρές κατανάλωσης  $i$  με μετρήσεις ανά μισή ώρα, δηλαδή με 48 μετρήσεις:

$$m_i = \begin{cases} \frac{\langle l_i, c_i \rangle}{\sum_{k=1}^{48} l_i(k)}, \text{ αν } \sum_{k=1}^{48} l_i(k) > 0 \\ 1, \text{ αν } \sum_{k=1}^{48} l_i(k) = 0 \text{ και } \sum_{k=1}^{48} c_i(k) = 0 \\ 0, \text{ αλλιώς} \end{cases}$$

Όπου  $l_i$  ένα δυαδικό διάνυσμα με 48 διαστάσεις, όσες και το δείγμα  $i$ , που έχει 1 τις χρονικές στιγμές όπου το δείγμα  $i$  έχει κορυφή και 0 όλες τις υπόλοιπες,  $c_i$  ένα αντίστοιχο διάνυσμα για το κέντρο της συστάδας στην οποία καταχωρήθηκε το  $i$ , και  $\langle l_i, c_i \rangle$  εσωτερικό τους γινόμενο.

Για συντομία στην επεξήγηση του Peak Match Score αλλά και άλλων μετρικών στην ενότητα 6.7 αυτής της εργασίας, θα αναφερόμαστε στην τοποθέτηση ενός δείγματος  $i$  με μία ακμή  $p$  σε συστάδα που έχει κέντρο  $c$  με ακμή στην ίδια χρονική στιγμή με την  $p$  ως «**επιτυχή ανίχνευση**» της ακμής  $p$ . Ενώ αν το κέντρο  $c$  έχει ακμή  $p'$  σε χρονική στιγμή που το δείγμα  $i$  δεν έχει, θα λέμε ότι ο αλγόριθμος έκανε «**λανθασμένη ανίχνευση**» της ακμής  $p'$ .

Από τον ορισμό του Peak Match Score βλέπουμε ότι είναι αύξον σε σχέση με τα επιμέρους σκορ  $m_i$  των δειγμάτων. Το επιμέρους σκορ  $m_i$  κάθε δείγματος παίρνει τιμές στο διάστημα  $[0,1]$  και είναι μεγαλύτερο όσο περισσότερες από τις ακμές του δείγματος  $i$  είναι

συγχρονισμένες με ακμές στο κέντρο  $c$  της συστάδας όπου ανήκει, δηλαδή για όσο περισσότερες από τις ακμές του  $i$  γίνεται «επιτυχής ανίχνευση» από τον αλγόριθμο.

Καταλαβαίνουμε λοιπόν ότι το Peak Match Score είναι **μεγαλύτερο όσο καλύτερα ο αλγόριθμος «ανιχνεύει» τις ακμές και των δειγμάτων  $i$  και συσταδοποιεί με βάση αυτές.**

Επίσης παρατηρούμε ότι όταν ένα δείγμα δεν έχει κάποια ακμή αλλά το κέντρο της συστάδας στην οποία ανήκει έχει (δηλαδή  $\sum_{k=1}^{48} l_i(k) = 0$  και  $\sum_{k=1}^{48} c_i(k) > 0$ , η τρίτη περίπτωση στο ...) το επιμέρους σκορ  $m_i$  για το δείγμα γίνεται 0. Δηλαδή, σε αυτή την περίπτωση το score «τιμωρεί» τον αλγόριθμο που κατέταξε το δείγμα σε συστάδα που έχει χαρακτηριστική ακμή, ενώ το δείγμα δεν έχει, δηλαδή για «λανθασμένη ανίχνευση» ακμής.

Τι γίνεται όμως όταν η συστάδα  $c_i$  έχει ακμές σε όλα τα σημεία που έχει και το δείγμα  $l_i$ , αλλά και σε άλλα σημεία πέρα από αυτά; Σε αυτή την περίπτωση, που εμπεριέχεται στην πρώτη του ορισμού (αν  $\sum_{k=1}^{48} l_i(k) > 0$ ), ο αλγόριθμος επιβραβεύεται με τέλειο επιμέρους σκορ για το δείγμα,  $m_i = 1$ , παρόλο που έκανε και κάποιες «λανθασμένες ανιχνεύσεις» ακμών. Υπάρχει λοιπόν το πρόβλημα ότι η μετρική Peak Match Score, όπως ορίστηκε στο [29], τιμωρεί τον αλγόριθμο για «λανθασμένη ανίχνευση» ακμής σε κάποιες περιπτώσεις, και σε κάποιες άλλες όχι.

Για τη λύση αυτού του προβλήματος και για τις ανάγκες αυτής της εργασίας δημιουργήθηκε μία νέα μετρική, το Peak Performance Score, που ορίζεται στην ενότητα 6.7.1.



## 6. Μεθοδολογία

### 6.1 Σύνολο Δεδομένων

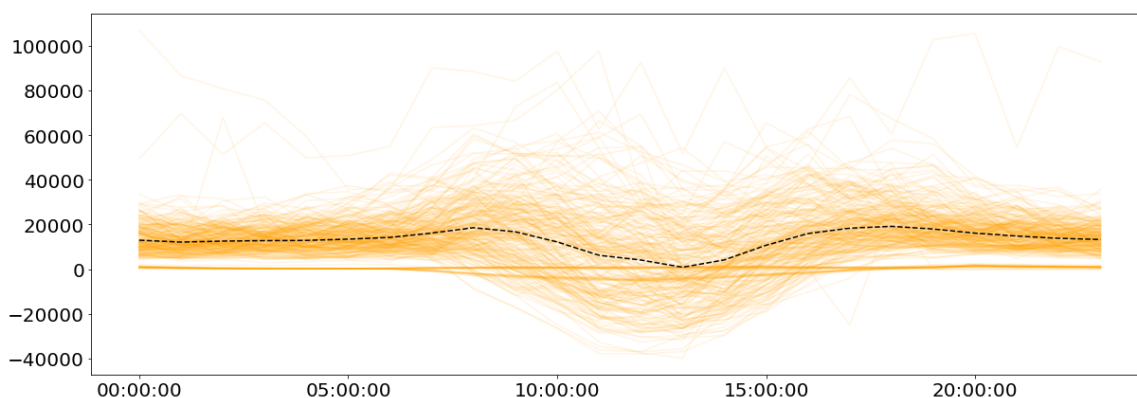
Τα δεδομένα που χρησιμοποιήθηκαν στο πειραματικό μέρος διατέθηκαν από την εταιρεία ASM Terni, μία δημοτική επιχείρηση που κατέχει και διαχειρίζεται το δίκτυο διανομής ηλεκτρικής ενέργειας στην κωμόπολη Terni της κεντρικής Ιταλίας. Προέρχονται από 51 έξυπνους μετρητές που η ASM αποφάσισε να ομαδοποιήσει σε ένα εικονικό δίκτυο, προκειμένου να διερευνήσει τις δυνατότητες των φορτίων του εικονικού δικτύου για απόκριση ζήτησης.

### 6.2 Ζητούμενο

Τα δύο σημαντικότερα προβλήματα που η ASM ευελπιστεί να λύσει μέσω της απόκρισης ζήτησης είναι:

A) το πρόβλημα της αντίστροφης ροής φορτίου (reverse power flow), κατά το οποίο η συνολική παραγωγή ενέργειας στο εικονικό δίκτυο είναι μεγαλύτερη από την συνολική κατανάλωση. Το φαινόμενο αυτό συμβαίνει νωρίς το μεσημέρι, όταν η παραγωγή από φωτοβολταϊκές εγκαταστάσεις κορυφώνεται, και όταν συμβαίνει υπάρχει έγχυση φορτίου από το δίκτυο που διαχειρίζεται η ASM στο δίκτυο που το τροφοδοτεί, δηλαδή το ιταλικό εθνικό δίκτυο. Αυτό είναι κάτι ανεπιθύμητο για την ASM, που επιθυμεί όλη η ενέργεια που παράγεται μέσα στο εικονικό δίκτυο να καταναλώνεται από τα φορτία του.

Στην εικόνα 2 βλέπουμε το άθροισμα παραγωγής και κατανάλωσης στο εικονικό δίκτυο για κάθε ημέρα στο σύνολο δεδομένων. Παρατηρούμε ότι αντίστροφη ροή φορτίου (τα κομμάτια του διαγράμματος που είναι αρνητικά) μπορούμε να έχουμε οποιαδήποτε στιγμή ανάμεσα στις 08:00 και στις 16:00. Ωστόσο, το φαινόμενο κορυφώνεται την περίοδο 11:00 με 14:00.



Εικόνα 2 - άθροισμα παραγωγής και κατανάλωσης ενέργειας στο εικονικό δίκτυο

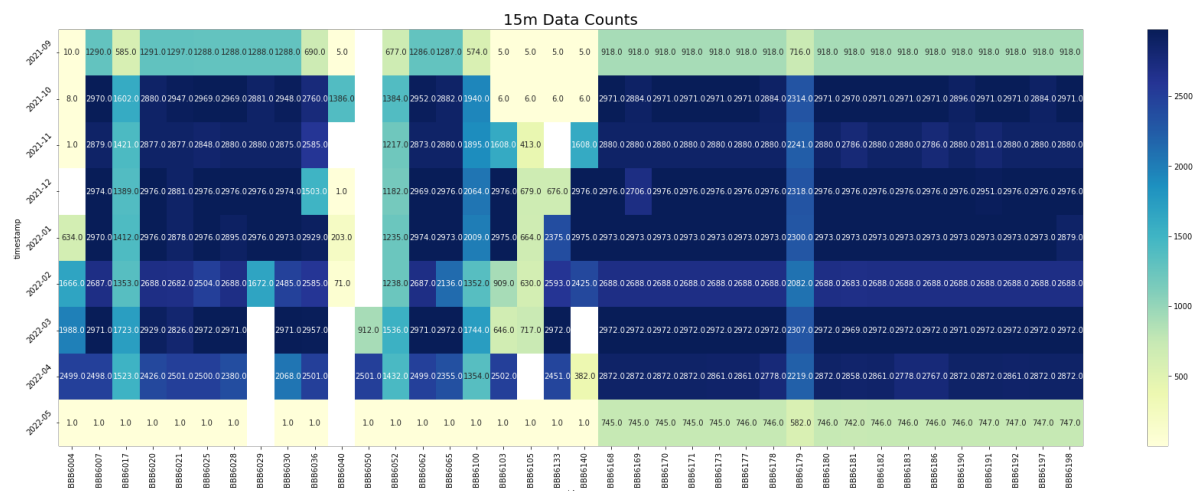
B) η μείωση στη ζήτηση της ηλεκτρικής ενέργειας τις ώρες αιχμής για το Ιταλικό εθνικό δίκτυο γενικότερα, που συνήθως είναι αργά το απόγευμα (17:00-19:00) .

Στην ανάλυση που ακολουθεί, εφαρμόζουμε τεχνικές συσταδοποίησης (clustering) στα δεδομένα κατανάλωσης, προκειμένου να εκτιμήσουμε ποια φορτία του εικονικού δικτύου θα ήταν προτιμότερο να ενταχθούν σε ένα πρόγραμμα απόκρισης ζήτησης. Μια τέτοια εκτίμηση εξαρτάται αρχικά από το αν μπορούν να συνεισφέρουν στη λύση των 2 παραπάνω προβλημάτων και κατά δεύτερον από το πόσο μεγάλη μπορεί να είναι αυτή η συνεισφορά, κάτι που, όπως αναλύεται στην ενότητα 4.2, εξαρτάται από πολλούς παράγοντες (μέγεθος του φορτίου, είδος χρήσης, προβλεψιμότητα στην καταναλωτική συμπεριφορά κ.α.).

### 6.3 Αρχική Ανάλυση Δεδομένων

Πριν την εφαρμογή αλγορίθμων συσταδοποίησης, ήταν σημαντικό να κάνουμε μία αρχική ανάλυση του συνόλου δεδομένων (initial data analysis), ώστε να αποκτήσουμε μία γενική εικόνα για τα χαρακτηριστικά του και να ανιχνεύσουμε προβλήματα προς αντιμετώπιση στο στάδιο της προ-επεξεργασίας, όπως απομακρυσμένες τιμές (outliers) και περιόδους που λείπουν (missing values). Σημαντικό κομμάτι αυτής της ανάλυσης ήταν η δημιουργία διαγραμμάτων που βοήθησαν στον εύκολο εντοπισμό των παραπάνω προβλημάτων.

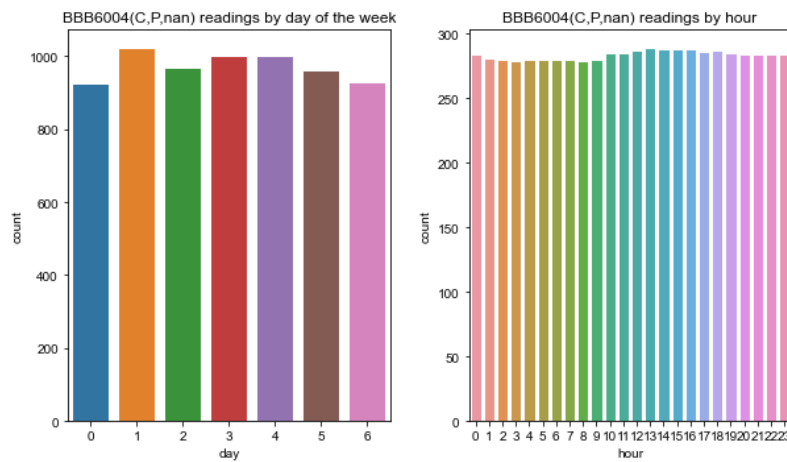
Το διάγραμμα στην **Εικόνα 3**, που εδώ έχει δημιουργηθεί για ένα υποσύνολο των δεδομένων, είναι ένα heatmap με τον αριθμό των samples για κάθε μήνα του χρόνου και μας έδωσε μία εικόνα για τους μήνες που είχαν τα περισσότερα διαστήματα χωρίς δεδομένα. Στον ψ άξονα έχουμε τους μήνες και στον χ τους έξυπνους μετρητές. Όπου υπάρχει κενό είναι γιατί δεν υπήρχε καμία μέτρηση.



**Εικόνα 3 - Heatmap με αριθμό μετρήσεων ανά μήνα και ανά smart meter**

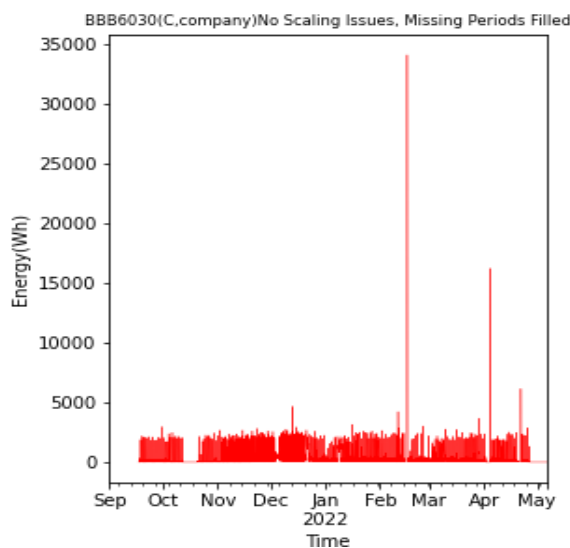
Διαγράμματα όπως αυτά στην **Εικόνα 4** που έγιναν για κάθε έξυπνο μετρητή ξεχωριστά, μας βοήθησαν να καταλάβουμε εάν υπήρχαν συγκεκριμένες μέρες της εβδομάδας ή ώρες της

ημέρας για τις οποίες είχαμε λιγότερα δεδομένα. Εδώ στον ψ άξονα έχουμε τον αριθμό των μετρήσεων και στον χ την ημέρα της εβδομάδας / ώρα της ημέρας.



**Εικόνα 4 - Bar plots με αριθμό δειγμάτων ανά μέρα της εβδομάδας και ώρα της ημέρας για τον μετρητή BBB6004**

Για τον γρήγορο εντοπισμό απομακρυσμένων τιμών βοήθησαν διαγράμματα όπως αυτό στην εικόνα Εικόνα 5 . Στον ψ άξονα έχουμε την κατανάλωση ενέργειας και στον χ το χρόνο. Εδώ φαίνονται τουλάχιστον 2 τιμές που είναι πολύ μεγαλύτερες από τις υπόλοιπες και χρειάζονται περαιτέρω διερεύνηση.



**Εικόνα 5 – Line Plot της κατανάλωση ενέργειας για μεγάλο χρονικό διάστημα για τον μετρητή BBB6030 – βοηθητικό για εντοπισμό outliers**

Τα παραπάνω διαγράμματα θα ήταν δύσκολο να γίνουν για ένα σύνολο δεδομένων με πολλά smart meters (πχ. εκατοντάδες ή χιλιάδες). Εκεί η ανάλυση θα βασιζόταν σε σωρευτικά

διαγράμματα και μετρικές. Επειδή στην περίπτωση μας είχαμε μόνο 50 smart meters αλλά και συνεχή επικοινωνία με τον πάροχο των δεδομένων, είχαμε τη δυνατότητα να κοιτάξουμε την κάθε περίπτωση ξεχωριστά και να ανακαλύψουμε προβλήματα που αφορούσαν μόνο συγκεκριμένους έξυπνους μετρητές.

## 6.4 Προεπεξεργασία των δεδομένων

Αφού εντοπίσαμε τις περιόδους που λείπανε δεδομένα και το πότε υπήρχαν τιμές outliers έπρεπε να επιλέξουμε κάποιες μεθόδους για την αντιμετώπισή τους, ώστε το σύνολο δεδομένων να είναι πλήρες και να μην έχει ακραίες τιμές που μπορούν να αλλοιώσουν τα αποτελέσματα της συσταδοποίησης. Οι παρακάτω διαδικασίες τοποθετήθηκαν στο pipeline της προ-επεξεργασίας των δεδομένων με τη σειρά που αναφέρονται.

### 6.4.1 Χειρισμός απομακρυσμένων τιμών (outliers)

Μαζί με το σύνολο δεδομένων είχαμε διαθέσιμη την μέγιστη ισχύ που αναγραφόταν στο σύμβολο κάθε καταναλωτή. Έτσι εφαρμόστηκε ο εξής κανόνας για όλες τις διαθέσιμες μετρήσεις: εάν μία μέτρηση ήταν μεγαλύτερη από την μέγιστη ισχύ στο σύμβολο του καταναλωτή επί το χρονικό διάστημα από την προηγούμενη μέτρηση, δηλαδή μεγαλύτερη από την μέγιστη ενέργεια που θα μπορούσε να καταναλώσει ο καταναλωτής μέσα σε αυτό το χρονικό διάστημα, τότε διαγράψαμε την μέτρηση και θεωρούσαμε ότι έλειπε.

### 6.4.2 Χειρισμός missing data – Data imputation

για περιόδους χωρίς δεδομένα που διαρκούν:

- λιγότερο από 2 ώρες ή περισσότερο από 2 ώρες αλλά κατά τη διάρκεια της νύχτας (από τις 00:00 μέχρι τις 06:00), συμπλήρωση των παρατηρήσεων μοιράζοντας την αμέσως επόμενη διαθέσιμη παρατήρηση για την κατανάλωση ενέργειας στα χρονικά διαστήματα που λείπουν
- περισσότερο από 2 ώρες και εκτός του διαστήματος 00:00 με 06:00, αφαίρεση της αντίστοιχης μέρας από το σύνολο δεδομένων του clustering

Η μέθοδος του διαμοιρασμού της επόμενης διαθέσιμης μέτρησης επιλέχθηκε αντί για άλλες μεθόδους (όπως αντικατάσταση με ιστορικό μέσο) γιατί είχαμε τη διαβεβαίωση ότι όταν είχαμε μία μέτρηση αυτή αντιστοιχούσε σε όλη την ενέργεια που καταναλώθηκε από το φορτίο στο διάστημα από την προηγούμενη μέτρηση.

Οι 2 ώρες επιλέχθηκαν ως χρονικό όριο για την εφαρμογή της παραπάνω μεθόδου με τη λογική ότι όσο το χρονικό παράθυρο μεγαλώνει τόσο λιγότερο ακριβής είναι η μέθοδος στην αποκατάσταση του ακριβούς σχήματος της καμπύλης και δεν θέλαμε να επηρεάσουμε τα αποτελέσματα της συσταδοποίησης με σχήματα κατανάλωσης που έχουν αλλοιωθεί.

Ο παραπάνω περιορισμός είχε ως εξαίρεση τις ώρες μέσα στη νύχτα, γιατί στο δικό μας σύνολο δεδομένων δεν παρατηρούσαμε μεγάλη δραστηριότητα μέσα στη νύχτα και ούτως ή άλλως οι ώρες της νύχτας δεν είχαν μεγάλο ενδιαφέρον για την επίτευξη των στόχων του προγράμματος απόκρισης ζήτησης. Επομένως, δεν θέλαμε να αφαιρέσουμε από το σύνολο δεδομένων ημέρες με σημαντική πληροφορία τις ώρες της ημέρας λόγω κενών μέσα στη νύχτα.

### 6.4.3 Κανονικοποίηση Ημερήσιων Προφίλ

Επειδή στο σύνολο δεδομένων είχαμε και μικρούς καταναλωτές (πχ. κατοικίες) και μεγάλους (πχ. το campus ενός πανεπιστημίου) και θέλαμε η συσταδοποίηση να γίνει με βάση μόνο το σχήμα των ημερήσιων προφίλ, δηλαδή να μην παραμορφώνονται τα αποτελέσματα από το ολικό μέγεθος της κατανάλωσης, επιλέξαμε να κανονικοποιήσουμε τα προφίλ πριν τη συσταδοποίηση. Αυτό έγινε με χρήση l2 νόρμας.

## 6.5 Επιλογές σχετικές με τη συσταδοποίηση

### 6.5.1 Ένα ή πολλά προφίλ ανά καταναλωτή

Μία σημαντική απόφαση που έπρεπε να ληφθεί πριν αρχίσει η διαδικασία του clustering είναι εάν τα δείγματα θα είναι οι μέσες καταναλώσεις των φορτίων ή όλα τα ημερήσια προφίλ. Στην πρώτη περίπτωση, κάθε φορτίο αντιστοιχεί σε ένα δείγμα στο σύνολο δεδομένων προς συσταδοποίηση, ενώ στη δεύτερη σε πολλά. Στη βιβλιογραφία χρησιμοποιούνται και οι 2 μέθοδοι.

Στο [47] οι συγγραφείς δημιουργούν ένα προφίλ μέσης κατανάλωσης για κάθε φορτίο (με την κατανάλωση σε κάθε ώρα να είναι ίση με τη μέση κατανάλωση για εκείνη την ώρα για την περίοδο μελέτης) και κάνουν συσταδοποίηση πάνω σε αυτά τα προφίλ. Αυτή η προσέγγιση μειώνει σε μεγάλο βαθμό τον όγκο των δεδομένων (αφού συνολικά έχουμε τόσα προφίλ όσα και τα φορτία), αλλά αγνοεί διαφορές που μπορεί να υπάρχουν στο timing και την ποσότητα της κατανάλωσης ενός σπιτιού από μέρα σε μέρα και από εποχή σε εποχή.

Μία άλλη προσέγγιση είναι η συσταδοποίηση χρησιμοποιώντας όλα τα ημερήσια προφίλ κατανάλωσης για όλα τα φορτία. Αυτή η προσέγγιση έχει το μεγάλο προτέρημα ότι δεν αγνοεί την εποχικότητα και την περιοδικότητα που μπορεί να εμφανίζεται στις καταναλώσεις ενός φορτίου, είναι ωστόσο πιο πολύπλοκη, αφού προφίλ καταναλώσεων του ίδιου φορτίου μπορούν να καταχωρηθούν σε διαφορετικές συστάδες, και προϋποθέτει πολύ μεγαλύτερο όγκο δεδομένων (αφού για 1 χρόνο δεδομένων, έχουμε 365 διαφορετικά προφίλ για κάθε φορτίο). Στο [18], [47] και στο [15], που ακολουθείται η παραπάνω τακτική, οι συγγραφείς θεωρούν ότι ένα φορτίο “ανήκει” στη συστάδα που περιέχει το μεγαλύτερο ποσοστό των ημερήσιων καμπύλων του.

Για την δική μας περίπτωση επιλέχθηκε συσταδοποίηση πάνω στις ημερήσιες καμπύλες φορτίου για όλο το δείγμα, όχι σε μία μέση καμπύλη για κάθε καταναλωτή. Εξάλλου, επειδή

το dataset περιέχει μόνο 50 smart meters, η συσταδοποίηση σε 50 δείγματα δεν θα ήταν πολύ χρήσιμη.

### 6.5.2 Εποχικός διαχωρισμός του συνόλου δεδομένων

Σε κάποια άρθρα οι συγγραφείς επιλέγουν να διαχωρίσουν τα δεδομένα ανάλογα με την εποχή, κυρίως σε δεδομένα χειμώνα και δεδομένα καλοκαιριού. Ο λόγος είναι ότι στα περισσότερα μέρη του κόσμου οι καταναλωτική συμπεριφορά σε σπίτια και επιχειρήσεις εξαρτάται σε μεγάλο βαθμό από τον καιρό[48]

Στη δίκη μας εργασία επιλέξαμε να μην κάνουμε τέτοιο διαχωρισμό, τουλάχιστον στην πρώτη ανάλυση των δεδομένων, γιατί θεωρήσαμε ότι μπορούμε να εξερευνήσουμε τον εποχικό χαρακτήρα των καταναλώσεων καλύτερα συσταδοποιώντας όλα τα δεδομένα και αναλύοντας τις διαφορές μεταξύ χειμώνα και καλοκαιριού εκ των υστέρων, στα αποτελέσματα. Αν δεν διαχωρίσουμε τα δεδομένα μπορούμε να ελέγξουμε πχ. αν η κατανομή των καταναλώσεων ενός έξυπνου μετρητή στις διάφορες συστάδες αλλάζει από τη μία εποχή στην άλλη ή αν κάποιες συστάδες περιέχουν περισσότερες καλοκαιρινές ημέρες απ' ότι χειμερινές, ή το αντίθετο.

Τέτοιες συγκρίσεις θα ήταν πιο δύσκολο να γίνουν αν διαχωρίζαμε εποχικά τις καταναλώσεις πριν τη συσταδοποίηση, γιατί τότε δεν θα είχαμε κοινές συστάδες για να μας διευκολύνουν στη σύγκριση.

### 6.5.3 Αλγόριθμοι και μετρικές απόστασης

Για την επιλογή της κατάλληλης μετρικής απόστασης και αλγορίθμου δοκιμάστηκαν πολλές από τις προτάσεις των ενοτήτων 5.3 και 5.4. Στον παρακάτω πίνακα φαίνονται αναλυτικά οι διαφορετικές επιλογές:

Αλγόριθμος	Μετρική
K-means	Ευκλείδεια, DTW με Sakoe-Chiba Radius = 1
K-medoids	Ευκλείδεια, DTW με Sakoe-Chiba Radius = 1
Agglomerative	Ευκλείδεια, DTW με Sakoe-Chiba Radius = 1

Για να καταλήξουμε στις τις παραπάνω επιλογές, πρώτα σκεφτήκαμε το πως θα θέλαμε να γίνει η συσταδοποίηση ιδανικά. Επιθυμούσαμε τη δημιουργία συστάδων που να διαχωρίζονται από δύο χαρακτηριστικά: τον χρονισμό των κορυφών και κοιλάδων των προφίλ τους (πχ. δύο προφίλ που εμφανίζουν κορυφή στις 2 το μεσημέρι θέλαμε να είναι στην ίδια συστάδα) καθώς και το σχήμα των προφίλ τους (πχ. θέλαμε προφίλ με δύο κορυφές μέσα στην μέρα να είναι σε διαφορετική συστάδα από προφίλ με μία κορυφή).

Η ευκλείδεια απόσταση είναι καλή στο να αποτυπώνει πληροφορία για το πρώτο χαρακτηριστικό αλλά δεν αποτυπώνει πληροφορία για το δεύτερο, το σχήμα του προφίλ. Από

την άλλη η απόσταση που χρησιμοποιεί Dynamic Time Warping, όταν δεν είναι κάπως περιορισμένη, αποτυπώνει πληροφορία για το σχήμα του προφίλ αλλά όχι για τον χρονισμό των κορυφών.

Χρειαζόμασταν επομένως μία μετρική απόστασης που να βρίσκεται κάπου ενδιάμεσα. Να έχει μία χρονική ελευθερία, ώστε να μπορεί να αποτυπώνει πληροφορία σχετικά με το σχήμα του προφίλ, αλλά όχι στο βαθμό που αγνοεί τελείως το χρόνο. Αυτή την ιδέα τελικά την υλοποιήσαμε χρησιμοποιώντας Dynamic Time Warping, επιτρέποντας όμως τη μετακίνηση κάθε σημείου της χρονοσειράς μόνο για μία χρονική μονάδα, δηλαδή μία ώρα. Πρακτικά αυτό σημαίνει ότι ορίσαμε το Sakoe-Chiba radius ίσο με 1 [40].

Η απόφαση αυτή λήφθηκε με τη λογική ότι μία χρονική χαλαρότητα της μίας ώρας δεν θα ήταν προβληματική αλλά μάλιστα επιθυμητή, γιατί η ουσία ήταν ο χωρισμός των κορυφών ανάλογα με την περίοδο της ημέρας (πρωί, μεσημέρι, απόγευμα κλπ.) και όχι τη συγκεκριμένη ώρα.

Εκτός από την απόσταση με DTW αποφασίστηκε να γίνουν τα πειράματα και με την ευκλείδεια απόσταση, επειδή είναι αυτή που χρησιμοποιείται στα περισσότερα άρθρα και θα αποτελούσε χρήσιμο benchmark. Εξάλλου έχει και το μεγάλο πλεονέκτημα της μικρής χρονικής πολυπλοκότητας.

Από αλγορίθμους επιλέχθηκε ο αρχικά k-means ως industry standard και επειδή έχει μικρές υπολογιστικές απαιτήσεις. Σε αυτόν προσθέσαμε τον k-medoids, που διαχειρίζεται καλά το θόρυβο και έχει το επιπλέον πλεονέκτημα ότι τα κέντρα των συστάδων του είναι πραγματικά δείγματα του συνόλου και όχι μέσοι όροι και έναν σωρευτικό ιεραρχικό αλγόριθμο με ward linkage λόγω της απλότητας και της συμβατότητάς του και με τις δύο μετρικές απόστασης.

Η σύγκριση των διαφορετικών επιλογών έγινε με συστηματικό τρόπο και αναλύεται περαιτέρω στην ενότητα 6.6.

#### 6.5.4 Αριθμός συστάδων

Στην συστηματική δοκιμή που αναλύεται στην ενότητα 6.6 δοκιμάστηκε ο διαχωρισμός των δεδομένων σε 3 μέχρι 20 συστάδες. Το ανώτερο όριο των 20 συστάδων γιατί επιθυμούσαμε ένα σχετικά μικρό σύνολο συστάδων που να γενικεύει καλά στο σύνολο δεδομένων, ώστε να είναι εύκολη η εξαγωγή συμπερασμάτων και η διαχείριση από έναν αναλυτή.

## 6.6 Σύγκριση διαφορετικών επιλογών – Νέες μετρικές επικύρωσης

### 6.6.1 Μεθοδολογία σύγκρισης

Για τη σύγκριση των διαφορετικών επιλογών και την επιλογή του αποτελέσματος που εξυπηρετούσε καλύτερα το πρόβλημα συνδυάσαμε τη σύγκριση των αποτελεσμάτων μέσω μετρικών επικύρωσης με οπτική θεώρηση των αποτελεσμάτων. Αυτό έγινε γιατί ενώ οι μετρικές επικύρωσης ποσοτικοποιούν την ποιότητα των αποτελεσμάτων, μειώνουν την πολύπλοκη διαδικασία της κρίσης του αν η συσταδοποίηση έγινε σωστά σε έναν μόνο αριθμό

και είναι επόμενο η εικόνα που δίνουν να είναι μερική. Η οπτική θεώρηση είναι χρήσιμη ως έλεγχος ότι η συσταδοποίηση λειτούργησε πράγματι με επιθυμητό τρόπο, δηλαδή ότι οι συστάδες που δημιουργήθηκαν διαφέρουν στα δύο χαρακτηριστικά που αναφέρθηκαν παραπάνω: τον χρονισμό των ακμών και το σχήμα της καμπύλης κατανάλωσης.

### 6.6.2 Μετρικές επικύρωσης

Στον παρακάτω πίνακα φαίνονται οι μετρικές επικύρωσης που χρησιμοποιήθηκαν για τη σύγκριση των αποτελεσμάτων:

Μετρικές επικύρωσης
Silhouette Score
Silhouette Score DTW
DB Index
Peak Match Score
Peak Performance Score

**Πίνακας 1 - Μετρικές επικύρωσης που χρησιμοποιήθηκαν για τον έλεγχο των μεθόδων clustering**

Οι τέσσερις πρώτες μετρικές αναφέρονται στην ενότητα 5.5. Η μετρική Peak Performance Score επινοήθηκε για τις ανάγκες της συγκεκριμένης εργασίας και ορίζεται παρακάτω.

Επίσης είναι σημαντικό να αναφέρουμε ότι για τον υπολογισμό των μετρικών που αναγνωρίζουν κορυφές (peak match score, peak performance score κλπ) ως κορυφές θεωρήθηκαν τοπικά μέγιστα με peak prominence  $\geq 0,2$ , δηλαδή με κατακόρυφη απόσταση από το χαμηλότερο προηγούμενο ή επόμενο τοπικό ελάχιστο του κανονικοποιημένου προφίλ κατανάλωσης πάνω από 0,2 [49].

### 6.6.3 ML Flow

Ένα σημαντικό κομμάτι της εργασίας ήταν η σύγκριση των διαφορετικών configurations για τη συσταδοποίηση (διαφορετικοί συνδυασμοί αλγορίθμου, μετρικής επικύρωσης και αριθμού συστάδων) προκειμένου να επιλέξουμε το configuration που οδηγεί στα καλύτερα αποτελέσματα, ή καλύτερα στα αποτελέσματα που εξυπηρετούν καλύτερα τις ανάγκες του προβλήματος. Αυτό σήμαινε ότι θα έπρεπε να εκτελέσουμε πολλά διαφορετικά πειράματα, και θα χρειαζόμασταν ένα τρόπο να αποθηκεύσουμε τα αποτελέσματά τους και αργότερα να τα συγκρίνουμε. Σε αυτή τη διαδικασία βοήθησε πολύ η χρήση του ML Flow.

Το ML Flow είναι μία open source πλατφόρμα που διευκολύνει στα διάφορα στάδια του κύκλου ζωής μίας εφαρμογής μηχανικής μάθησης. Στην παρούσα εργασία, χρησιμοποιήθηκε για την δημιουργία του pipeline επεξεργασίας των δεδομένων, την εκτέλεση των πειραμάτων, την αποθήκευση των αποτελεσμάτων, καθώς και τη σύγκριση τους με βάση τις διάφορες μετρικές του Πίνακας 1.



Στην Εικόνα 6 φαίνεται μέρος μίας λίστας με όλα τα πειράματα που εκτελέστηκαν για τις ανάγκες της εργασίας.

Showing 100 matching runs

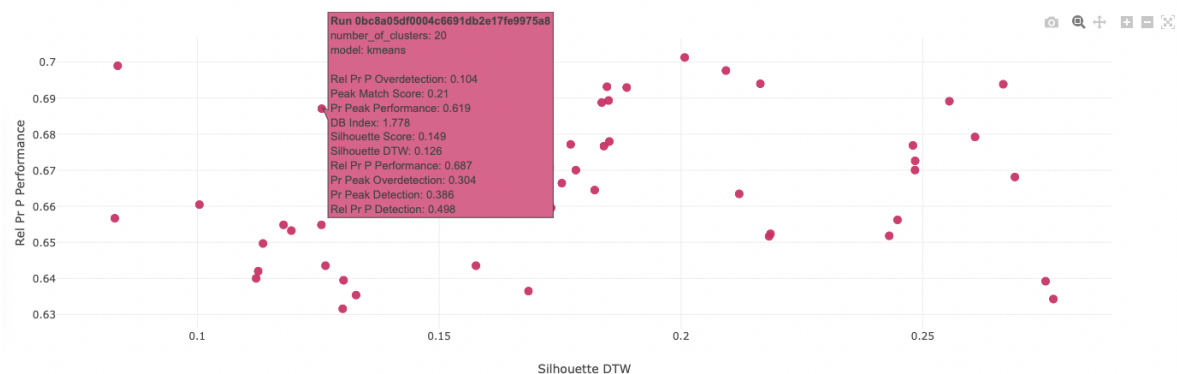
	Metrics <				Parameters <					
<input type="checkbox"/>	DB Index	Peak Match Sc	Rel Pr P Perfor	Silhouette DTW	distance_metri	in_csv	↓ k	model	number_of_clu	season
<input type="checkbox"/>	2.546	0.164	0.523	0.201	euclidian	/var/folders...	9	som	9	all
<input type="checkbox"/>	0.939	0.145	0.408	0.247	dtw	/var/folders...	9	agglomerat...	9	all
<input type="checkbox"/>	2.03	0.151	0.529	0.242	euclidian	/var/folders...	9	agglomerat...	9	all
<input type="checkbox"/>	2.636	0.133	0.527	0.239	dtw	/var/folders...	9	kmedioids	9	all
<input type="checkbox"/>	1.86	0.239	0.539	0.24	euclidian	/var/folders...	9	kmedioids	9	all
<input type="checkbox"/>	2.087	0.26	0.652	0.219	dtw	/var/folders...	9	kmeans	9	all
<input type="checkbox"/>	1.782	0.149	0.622	0.287	euclidian	/var/folders...	9	kmeans	9	all
<input type="checkbox"/>	2.471	0.194	0.551	0.221	euclidian	/var/folders...	8	som	8	all
<input type="checkbox"/>	0.795	0.148	0.408	0.269	dtw	/var/folders...	8	agglomerat...	8	all

**Εικόνα 6 - λίστα του ML Flow με όλα τα πειράματα που εκτελέστηκαν για την επιλογή του καλύτερου configuration**

Για κάθε πείραμα που εκτελέστηκε, το ML Flow συγκράτησε τις παραμέτρους της συσταδοποίησης (αλγόριθμος, μετρική απόστασης, αριθμός συστάδων) καθώς και τις τιμές των μετρικών επικύρωσης (validation metrics). Αυτό διευκόλυνε πολύ τη σύγκριση μεταξύ των διαφορετικών επιλογών, αφού μπορούσαμε, μέσω της γραφική διεπαφής του MLFlow, να κατατάξουμε τις διαφορετικές επιλογές με βάση την επίδοσή τους σε κάθε μετρική επικύρωσης (Εικόνα 7), καθώς και να παράγουμε διαγράμματα για να ελέγξουμε τη συσχέτιση μεταξύ παραμέτρων (πχ. του αριθμού των συστάδων) και επιδόσεων στις μετρικές επικύρωσης (πχ. Peak Performance Score). Ένα τέτοιο διάγραμμα φαίνεται στην Εικόνα 8.

	Metrics	Parameters <					
<input type="checkbox"/>	↓ Rel Pr P Perfc	distance_metri	in_csv	k	model	number_of_clu	season
<input type="checkbox"/>	0.701	euclidian	/var/folders...	15	kmeans	15	all
<input type="checkbox"/>	0.701	-	-	-	kmeans	15	-
<input type="checkbox"/>	0.699	euclidian	/var/folders...	20	som	20	all
<input type="checkbox"/>	0.699	-	-	-	som	20	-
<input type="checkbox"/>	0.698	euclidian	/var/folders...	14	kmeans	14	all
<input type="checkbox"/>	0.698	-	-	-	kmeans	14	-
<input type="checkbox"/>	0.694	euclidian	/var/folders...	13	kmeans	13	all

**Εικόνα 7 - κατάταξη πειραμάτων με βάση τη μετρική Relaxed Peak Performance**



**Εικόνα 8 - γραφική παράσταση που παρήγαγε το MLFlow για τη διερεύνηση του συσχετίσης μεταξύ των μετρικών Relaxed Peak Performance Score και Silhouette Score DTW**

## 6.7 Ορισμός νέων μετρικών επικύρωσης

Όπως αναφέρεται στην ενότητα 5.5, το Peak Match Score ήταν μία καλή προσπάθεια να ποσοτικοποιηθεί η ικανότητα μίας μεθόδου συσταδοποίησης για ηλεκτρικά φορτία να συσταδοποιεί με βάση τις ακμές, δηλαδή να τοποθετεί τα δείγματα που έχουν ακμές σε συστάδες που έχουν τις ίδιες ακμές, δηλαδή ακμές τις ίδιες χρονικές στιγμές. Ωστόσο, όπως ήταν ορισμένο, τιμωρούσε μία «λανθασμένη ανίχνευση» μία ακμής (βλ. ενότητα 0) μόνο όταν το δείγμα δεν είχε καθόλου ακμές.

Παρακάτω ορίζεται η μετρική Peak Performance Score, εμπνευσμένη από το Peak Match Score αλλά προσαρμοσμένη ώστε να λαμβάνει υπόψη τόσο την «επιτυχή ανίχνευση» ακμών όσο και τη «λανθασμένη», σε όλες τις περιπτώσεις.

### 6.7.1 Peak Performance Score

Το Peak Performance Score ορίζεται ως ο μέσος όρος των επιμέρους σκορ  $m_i$  για τα δείγματα του συνόλου δεδομένων:

$$peakPerformanceScore = \frac{1}{N} \sum_{i=1}^N m_i$$

Όπου  $N$  είναι ο αριθμός των δειγμάτων στο σύνολο δεδομένων και το επιμέρους σκορ για κάθε δείγμα  $i$ ,  $m_i$ , είναι:

$$m_i = \begin{cases} 1, & \text{αν } \sum_{k=1}^{24} l_i(k) = 0 \text{ και } \sum_{k=1}^{24} c_i(k) = 0 \\ \frac{\langle l_i, c_i \rangle}{\max\{\sum_{k=1}^{24} l_i(k), \sum_{k=1}^{24} c_i(k)\}}, & \text{αλλιώς} \end{cases}$$

Όπου  $l_i$  ένα δυαδικό διάνυσμα με 24 διαστάσεις, μία για κάθε ώρα της ημέρας, που έχει 1 τις ώρες όπου το δείγμα  $i$  έχει κορυφή και 0 όλες τις υπόλοιπες,  $c_i$  το αντίστοιχο διάνυσμα για το κέντρο της συστάδας στην οποία καταχωρήθηκε το  $i$ , και  $\langle l_i, c_i \rangle$  εσωτερικό τους γινόμενο.

Σε αντίθεση με το Peak Match Score, το Peak Performance Score τιμωρεί τη «λανθασμένη ανίχνευση» ακμής (βλ. ενότητα 0) σε όλες τις περιπτώσεις:

1. Αν το δείγμα  $i$  δεν έχει καμία ακμή αλλά η συστάδα στην οποία έχει καταχωρηθεί έχει, τότε το επιμέρους σκορ  $m_i$  γίνεται 0 και το συνολικό σκορ μειώνεται, όπως και στο Peak Match Score
2. Αν το δείγμα έχει ακμές αλλά η συστάδα στην οποία έχει καταχωρηθεί από τον αλγόριθμο έχει περισσότερες, τότε  $\sum_{k=1}^{24} c_i(k) > \sum_{k=1}^{24} l_i(k)$ , ο παρονομαστής του  $m_i$  γίνεται  $\max\{\sum_{k=1}^{24} l_i(k), \sum_{k=1}^{24} c_i(k)\} \rightarrow \sum_{k=1}^{24} c_i(k)$ , το επιμέρους σκορ  $m_i$  μειώνεται, και άρα το ίδιο και το συνολικό σκορ. Αυτό δεν γίνεται στο Peak Match Score. Ένα παράδειγμα με 5 διαστάσεις αντί για 24 για απλότητα: έστω  $l_i = [0,0,0,1,0]$  το δυαδικό διάνυσμα του δείγματος και  $c_i = [0,1,0,1,0]$  του κέντρου της συστάδας στην οποία έχει καταχωρηθεί. Σε αυτή την περίπτωση για το Peak Match Score  $m_i = 1$ , δηλαδή ο αλγόριθμος επιβραβεύτηκε για την «επιτυχή ανίχνευση» της ακμής στην θέση 3 αλλά δεν τιμωρήθηκε για την «λανθασμένη ανίχνευση» ακμής στην θέση 1. Αντίθετα, για το Peak Performance Score  $m_i = 0.5$ , δηλαδή ο αλγόριθμος πήρε τους μισούς πόντους, γιατί «ανίχνευσε επιτυχώς» τη μία ακμή στο δείγμα αλλά έκανε και «λανθασμένη ανίχνευση» μίας δεύτερης ακμής.

Το Peak Performance Score παίρνει τιμές από 0 έως 1. Αν πάρει την τιμή 1 αυτό σημαίνει ότι ο αλγόριθμος έκανε «επιτυχή ανίχνευση» όλων των ακμών των δειγμάτων (δηλαδή όλα τα δείγματα τοποθετήθηκαν σε συστάδες που είχαν ακμές τις ίδιες ώρες με αυτά) και ακόμη ότι δεν έκανε καμία «λανθασμένη ανίχνευση» ακμής (δεν υπήρξε δείγμα που να τοποθετήθηκε σε συστάδα που να είχε ακμή σε ώρα που το δείγμα δεν είχε). Σε αυτή την περίπτωση ο αλγόριθμος θα ήταν τέλειος όσον αφορά τη συσταδοποίηση με βάση τον χρονισμό των ακμών.

Από την άλλη, αν πάρει την τιμή 0 σημαίνει ότι δεν έγινε καμία «επιτυχής ανίχνευση» ακμής και ότι όλες οι ανιχνεύσεις που έγιναν ήταν «λανθασμένες ανιχνεύσεις». Σε αυτή την περίπτωση ο αλγόριθμος δεν συσταδοποιεί καθόλου με βάση το χρονισμό των ακμών.

### 6.7.2 Χρονική χαλάρωση κατά 1 βήμα

Το Peak Performance Score εφαρμόστηκε και με χρονική χαλάρωση ενός βήματος (δηλαδή μίας ώρας) στο «ταίριασμα» των ακμών. Για παράδειγμα, κατά τον υπολογισμό αυτής της μετρικής, εάν ένα δείγμα είχε κορυφή στις 1 μ.μ. και ο αλγόριθμος το κατέτασσε σε συστάδα με κορυφή στις 2 μ.μ. θα θεωρούσαμε ότι οι ακμές ταιριάζουν.

Για να ξεχωρίσουμε την πιο «χαλαρή» μετρική από την προηγούμενη προσθέσαμε το επίθετο “relaxed” στο όνομά της (Peak Performance Score -> Relaxed Peak Performance Score).

## 6.8 Επιλογή των ιδανικότερων συστάδων για απόκριση ζήτησης

Μετά από την εκτέλεση των πειραμάτων και την επιλογή του βέλτιστου συνδυασμού αλγορίθμου και παραμέτρων για τη συσταδοποίηση, έγινε μία ανάλυση των συστάδων προκειμένου να επιλεγθούν αυτές με τις καλύτερες προοπτικές να βοηθήσουν στην επίτευξη των στόχων του προγράμματος απόκρισης ζήτησης.

Αυτή η ανάλυση έγινε λαμβάνοντας υπόψη 3 χαρακτηριστικά που επηρεάζουν το αν μία ομάδα καταναλωτών αξίζει να ενταχθεί σε ένα πρόγραμμα DR και τη δομή που συμφέρει να έχει το πρόγραμμα αυτό: το σχήμα του κέντρου της συστάδας (χαρακτηριστικό προφίλ κατανάλωσης), τη χρήση των φορτίων (οικιακά, εμπορικά κλπ.) και την μέση εντροπία. Το πως αυτά τα χαρακτηριστικά χρησιμοποιούνται στη βιβλιογραφία για την επιλογή των καλύτερων υποψηφίων για ένα πρόγραμμα απόκρισης ζήτησης αναλύεται στην ενότητα 0. Για τη δική μας περίπτωση:

- **Σχήμα χαρακτηριστικού προφίλ:** όπως αναφέρθηκε στην ενότητα 6.2, μας αφορούσαν καταναλωτές που μπορούν να βοηθήσουν α) στη μείωση του reverse power flow, κάτι που μπορεί να γίνει με **αύξηση** της κατανάλωσης τις ώρες που η παραγωγή κορυφώνεται, κάτι που όπως είδαμε στην ενότητα 6.2 γίνεται κυρίως στις 11 το πρωί με 2 το μεσημέρι, και β) στην μείωση της κατανάλωσης τις ώρες αιχμής δηλαδή συνήθως στις 5-7 το απόγευμα. Για αυτό ψάξαμε για ομάδες καταναλωτών (συστάδες) που έχουν χαρακτηριστικά προφίλ με κορυφή α) λίγο πριν τις 11 το πρωί ή λίγο μετά τις 2, ώστε να τους παροτρύνουμε να μεταφέρουν αυτή την κορυφή στο διάστημα 11-2 (πχ. αν συνήθως μαγειρεύουν στις 3 το μεσημέρι να αρχίσουν να μαγειρεύουν νωρίτερα, στο διάστημα 11-2) ή β) στις 5-7 το απόγευμα, ώστε να τους παροτρύνουμε να μειώσουν την κατανάλωση αυτές τις ώρες, είτε ακυρώνοντας κάποιες εργασίες (peak shaving) είτε

προγραμματίζοντάς τες νωρίτερα ή αργότερα μέσα στη μέρα (peak shifting - πχ. αν ζεσταίνουν το νερό για μπάνιο μόλις γυρνάνε από τη δουλειά, να το κάνουν νωρίτερα, το μεσημέρι)

- **Χρήση των φορτίων:** το σύνολο δεδομένων που είχαμε στη διάθεσή μας περιείχε τόσο οικιακά όσο και εμπορικά φορτία. Αφού βρήκαμε ποιες συστάδες μας ενδιέφεραν περισσότερο από το σχήμα της κατανάλωσης των κέντρων τους, έπρεπε να αποφασίσουμε ποιοι τύποι προγραμμάτων απόκρισης ζήτησης (ενότητα 4.1) θα τους ταίριαζαν περισσότερο. Όπως είδαμε στην ενότητα 4.3, τα εμπορικά φορτία είναι ιδανικότερα για προγράμματα που ακολουθούν τον ίδιο τρόπο τιμολόγησης κάθε μέρα και απαιτούν αναλυτικό προγραμματισμό των διάφορων εργασιών (άρα Time Of Use (TOU) προγράμματα), ενώ τα οικιακά φορτία, επειδή οι εργασίες τους γενικά δεν είναι τόσο κρίσιμες και χρονικά περιορισμένες όσο των εμπορικών, μπορούν να αποδώσουν καλά σε προγράμματα με περιστασιακά DR γεγονότα (πχ. προγράμματα Critical Peak Pricing (CPP)) είτε και προγράμματα με δυναμική τιμολόγηση (Real Time Pricing (RTP)), εφόσον υπάρχει η κατάλληλη υλική υποδομή ώστε η διαχείριση των φορτίων να γίνεται αυτόματα (smart συσκευές, controllers κλπ.)
- **Εντροπία:** ένα άλλο χαρακτηριστικό που επηρεάζει το ποιος είναι ο ιδανικότερος τύπος προγράμματος για έναν καταναλωτή ή μία ομάδα καταναλωτών είναι η εντροπία. Όπως είδαμε στην ενότητα 4.2, καταναλωτές με μικρή εντροπία ταιριάζουν περισσότερο σε incentive based προγράμματα ή προγράμματα με προβλέψιμη τιμολογική πολιτική (TOU), ενώ καταναλωτές με μεγάλη εντροπία αποδίδουν καλύτερα σε προγράμματα CPP ή RTP. Για τις ανάγκες τις εργασίας μετρήσαμε την εντροπία για κάθε φορτίο στο σύνολο δεδομένων, με βάση τον τύπο στην ενότητα 4.2.2, καθώς και τη μέση εντροπία για κάθε συστάδα. Αυτές οι πληροφορίες ήταν χρήσιμες στην εξαγωγή συγκεκριμένων προτάσεων για απόκριση ζήτησης.

## 7. Αποτελέσματα

### 7.1 Αποτελέσματα σύγκρισης μεταξύ αλγορίθμων και μετρικών απόστασης

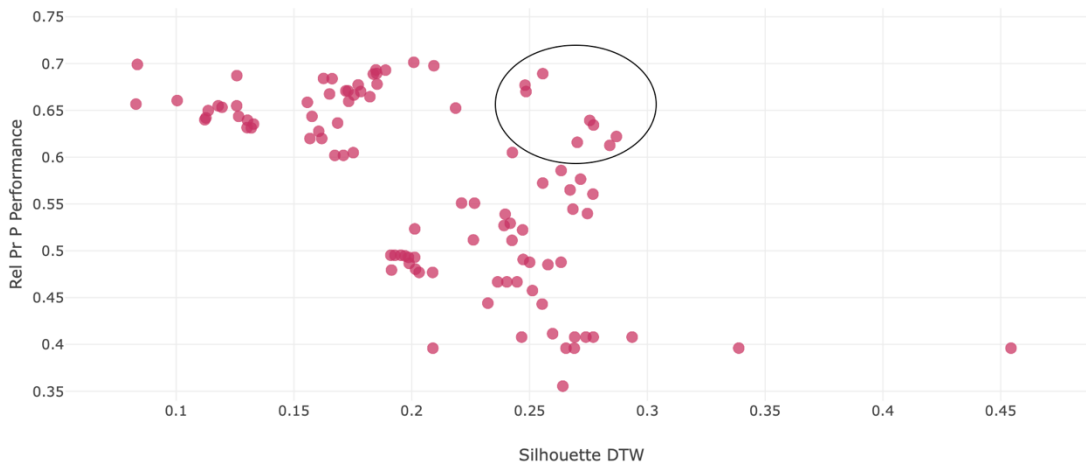
Όπως αναφέρουμε στην ενότητα 6.5.3, για να συγκρίνουμε τις διαφορετικές συστάδες καταναλωτών και να επιλέξουμε τις κατάλληλες για απόκριση ζήτησης ήταν σημαντικό αυτές να διαφέρουν μεταξύ τους όσον αφορά το σχήμα τους και ακόμη τα δείγματα που ανήκουν σε κάθε μία να έχουν κύρια ακμή περίπου την ίδια ώρα. Θέλαμε δηλαδή αποτελέσματα στα οποία πχ. τα δείγματα που έχουν μία κύρια ακμή το πρωί να ανήκουν όλα στην ίδια συστάδα και αυτή η συστάδα να είναι διαφορετική από αυτή στην οποία ανήκουν τα δείγματα με ακμή το μεσημέρι.

Για να ποσοτικοποιήσουμε το πόσο καλά συσταδοποιούσε ο κάθε αλγόριθμος με βάση τις ακμές χρησιμοποιήσαμε τις μετρικές Peak Match Score, Peak Performance Score και Relaxed Peak Performance Score, ενώ για να ελέγξουμε το αν οι συστάδες που προκύπτουν διαφέρουν επαρκώς μεταξύ τους τις μετρικές DB Score, Silhouette Score και Silhouette Score DTW (με Sakoe-Chiba Radius = 1).

Αφού τρέξαμε δοκιμές παρατηρήσαμε ότι τα πειράματα με μετρική απόστασης DTW τα πήγαν καλύτερα στις μετρικές επικύρωσης με χρονική χαλάρωση μίας ώρας (Relaxed Peak Performance Score και Silhouette Score DTW) σε σχέση με τις αυστηρές, πράγμα αναμενόμενο αφού το dynamic time warping λειτουργεί και αυτό με χρονική χαλάρωση.

Τελικά αποφασίσαμε να επικεντρωθούμε στις μετρικές με χρονική χαλάρωση γιατί, όπως αναφέραμε στην ενότητα 6.7.2, δεν επιθυμούσαμε τη δημιουργία συστάδων με κύριες ακμές για κάθε διαφορετική ώρα της ημέρας αλλά για γενικότερες χρονικές περιόδους (ξημερώματα, πρωί, μεσημέρι κλπ.)

Με τη βοήθεια του ML Flow δημιουργήσαμε το διάγραμμα της εικόνας Εικόνα 9, που έχει στον ψ άξονα το Relaxed Peak Performance Score και στον χ το Silhouette Score DTW και τοποθετεί στο χώρο όλα τα πειράματα με διαφορετικό συνδυασμό παραμέτρων. Μας ενδιέφεραν περισσότερο οι συνδυασμοί με καλή επίδοση και στις δύο μετρικές, δηλαδή αυτοί στην πάνω δεξιά περιοχή του χώρου, που είναι κυκλωμένη στη φωτογραφία.



**Εικόνα 9 - γραφική παράσταση που παρήγαγε το MLFlow για τη σύγκριση των αποτελεσμάτων της συσταδοποίησης με διάφορους συνδυασμούς παραμέτρων**

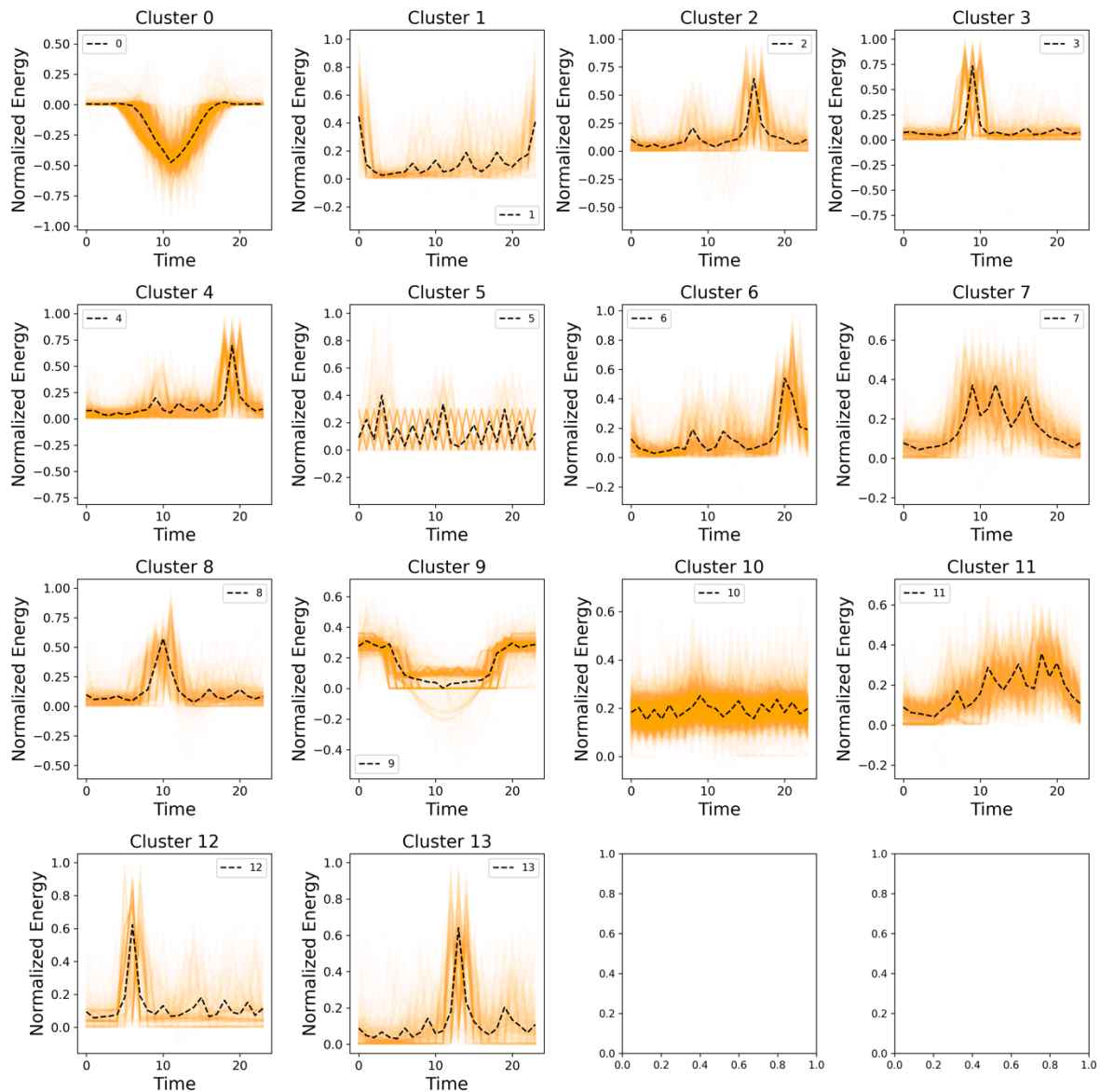
Ο Πίνακας 2 περιέχει όλους τους συνδυασμούς παραμέτρων που ανήκαν σε αυτή την περιοχή.

**Πίνακας 2 - συνδυασμοί παραμέτρων με τις καλύτερες επιδόσεις στις μετρικές Relative Peak Performance Score και Silhouette DTW**

Αλγόριθμος	Μετρική Απόστασης	Αριθμός συστάδων
K-means	DTW	6
K-means	DTW	8
K-means	DTW	9
K-means	DTW	14
K-means	Euclidean	8
K-means	Euclidean	9
K-medoids	DTW	13
K-medoids	DTW	14

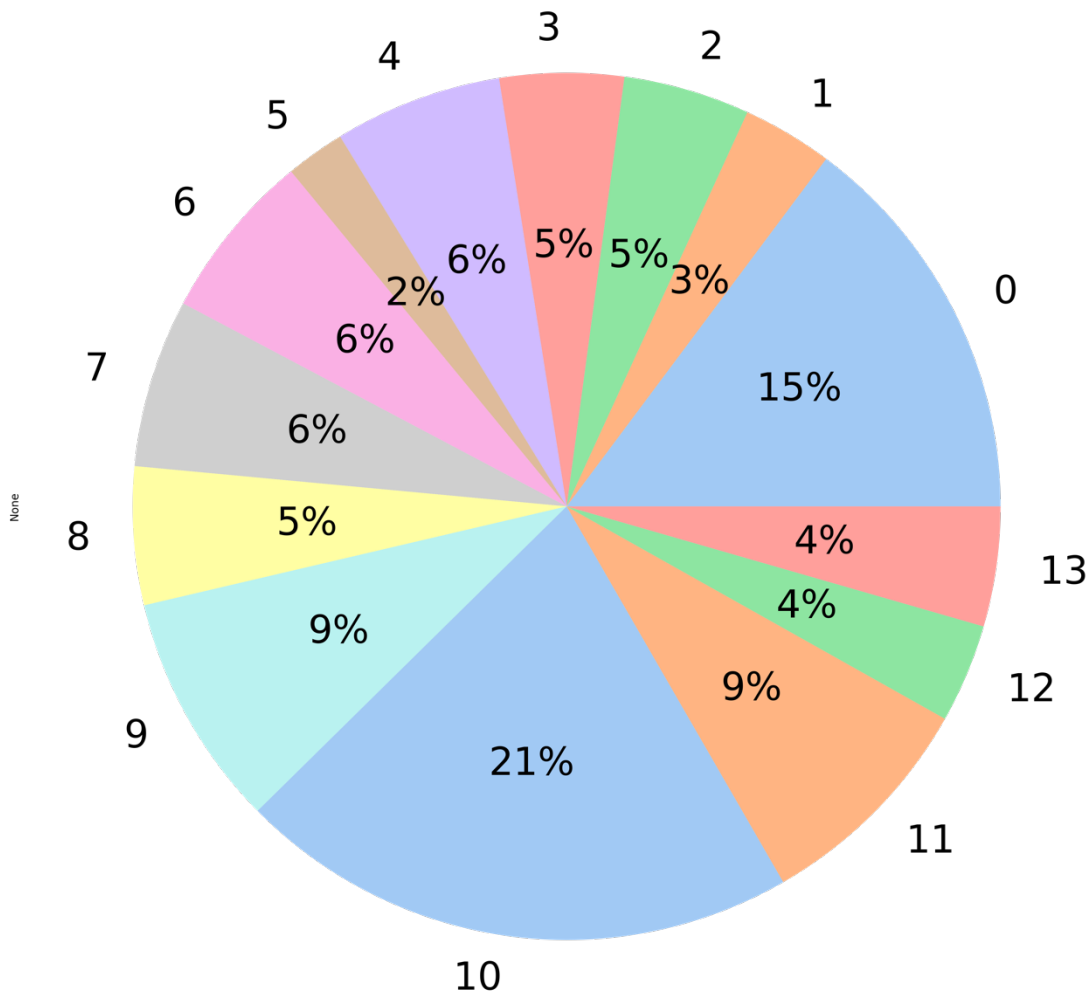
Τελικά μετά από οπτική θεώρηση των συστάδων αποφασίσαμε να συνεχίσουμε την ανάλυση με αλγόριθμο K-means, μετρική απόστασης DTW και αριθμό συστάδων = 14.

Στην Εικόνα 10 βλέπουμε τις συστάδες που προέκυψαν για τον παραπάνω συνδυασμό παραμέτρων, ενώ στην Εικόνα 11 μία πίτα με το ποσοστό των συνολικών δειγμάτων που καλύπτει κάθε συστάδα.



**Εικόνα 10 - Τα σχήματα των συστάδων που προέκυψαν για αλγόριθμο k-means, μετρική απόστασης DTW και  $k=14$ . Με διακεκομμένη γραμμή είναι το κέντρο της κάθε συστάδας και με αχνό πορτοκαλί τα δείγματα που την απαρτίζουν**





**Εικόνα 11 – Ποσοστά επί των συνολικών δειγμάτων που καλύπτει η κάθε συστάδα**

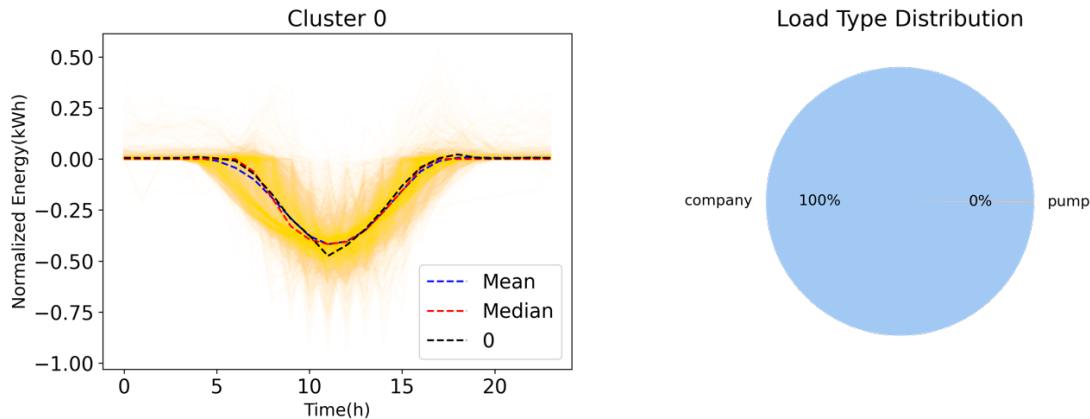
## 7.2 Περιγραφή των συστάδων

Παρακάτω γίνεται μία περιγραφή των συστάδων με βάση το σχήμα του χαρακτηριστικού προφίλ τους και τις χρήσεις των φορτίων που ανήκουν σε αυτές.

Είναι σημαντικό να επαναλάβουμε ότι η συσταδοποίηση έγινε επί των ημερήσιων καμπυλών φορτίου και όχι επί της μέσης κατανάλωσης. Επομένως ένα φορτίο μπορεί να ανήκει σε

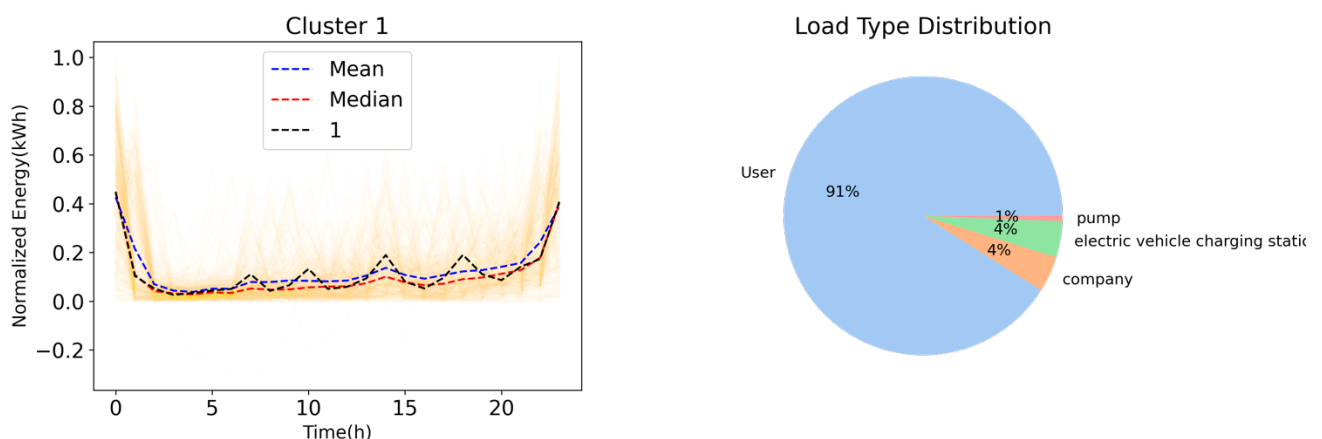
διαφορετική συστάδα ανάλογα με τη μέρα. Στη συνέχεια, όταν αναλύονται οι συστάδες και γίνονται υποθέσεις για τη χρησιμότητα των μελών τους σε ένα πρόγραμμα DR, κάθε μέλος θεωρείται ότι είναι ένα φορτίο τις μέρες που η κατανάλωσή του ταιριάζει με το προφίλ της συστάδας. Τις μέρες που η κατανάλωση ταιριάζει με κάποια άλλη συστάδα, οι υποθέσεις αυτές δεν θα ισχύουν. Αυτή η ιδιαιτερότητα λήφθηκε υπόψιν και αναλύθηκε περαιτέρω στο κεφάλαιο 7.

**Cluster 0:**



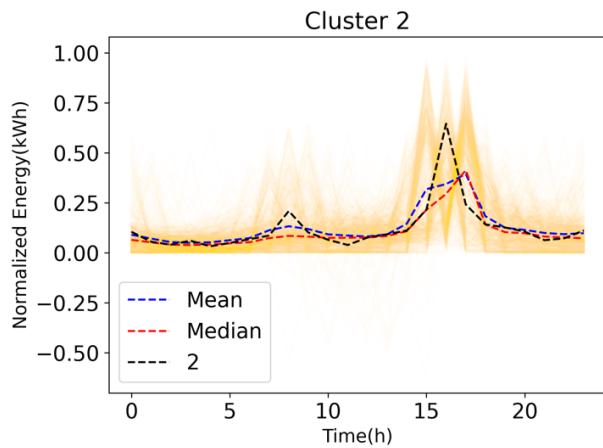
Περιέχει τα ημερήσια προφίλ εγκαταστάσεων που χρησιμοποιούνται μόνο για παραγωγή ενέργειας, αφού αν υπάρχει κατανάλωση είναι πολύ μικρή και προέρχεται μάλλον από μικρά βοηθητικά φορτία. Οι ώρες που υπάρχει παραγωγή είναι ίδιες με τις ώρες που υπάρχει ηλιοφάνεια, οπότε είναι εμφανές ότι τα φορτία είναι φωτοβολταϊκές εγκαταστάσεις και επειδή η παραγωγή των φωτοβολταϊκών δεν μπορεί να ελεγχθεί, δεν μπορούν να χρησιμοποιηθούν για να παρέχουν ελαστικότητα στο δίκτυο.

**Cluster 1:**

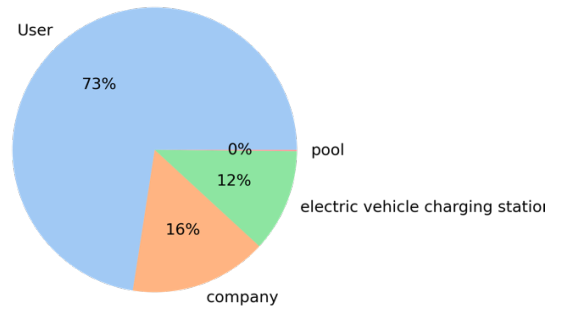


Περιέχει κυρίως καταναλώσεις από οικιακούς χρήστες (κατά 88%) που κορυφώνονται περίπου τα μεσάνυχτα. Δεδομένου ότι τα μεσάνυχτα συνήθως δεν υπάρχει συμφόρηση στο δίκτυο, δεν υπάρχει λόγος τα φορτία που ανήκουν σε αυτόν τον cluster να ενταχθούν σε πρόγραμμα DR.

**Cluster 2:**

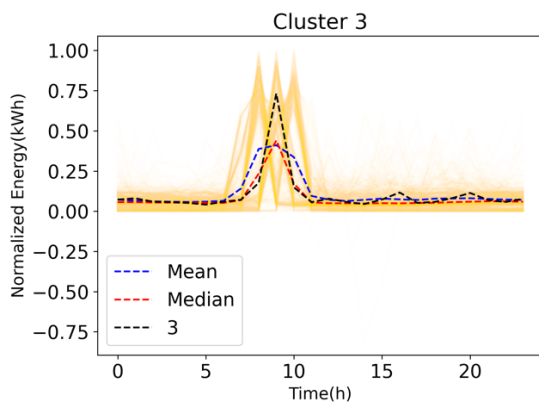


Load Type Distribution

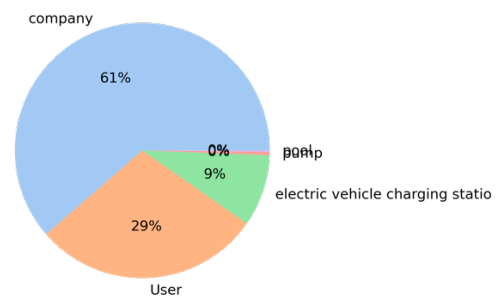


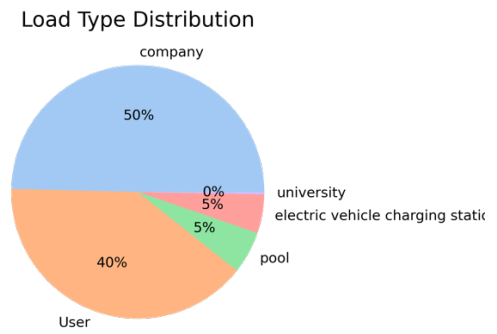
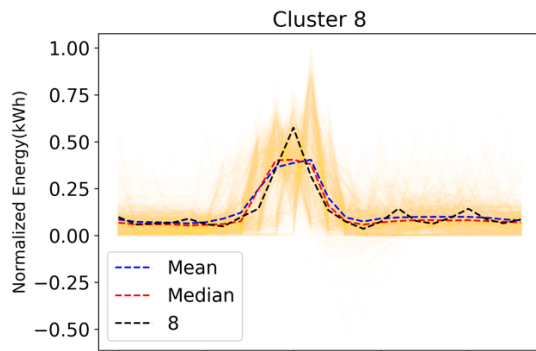
Οι καταναλώσεις που ανήκουν σε αυτόν τον cluster αντιστοιχούν κυρίως σε οικιακούς χρήστες (κατά 73%) και χαρακτηρίζονται από μία κορυφή νωρίς το απόγευμα, στο διάστημα 15:00 με 17:00. Θα ήταν συμφέρον για το δίκτυο αυτή η κορυφή να μετακινείται νωρίτερα, πιο κοντά στο μεσημέρι και μακριά από το απόγευμα, για να βοηθάει ταυτόχρονα και στην μείωση της αντίστροφης ροής φορτίου και στην αποσυμφόρηση του εθνικού δικτύου τις ώρες που η ζήτηση κορυφώνεται (από τις 17:00 και μετά).

**Cluster 3&8:**



Load Type Distribution

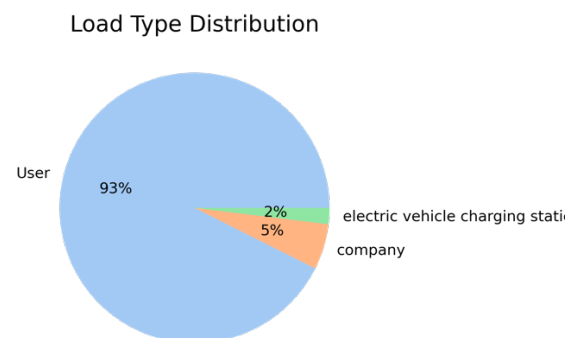
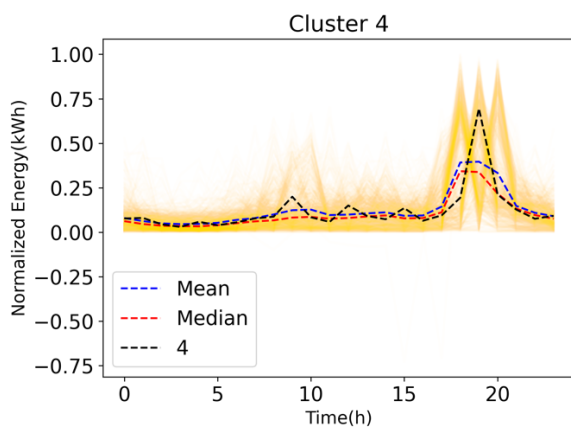


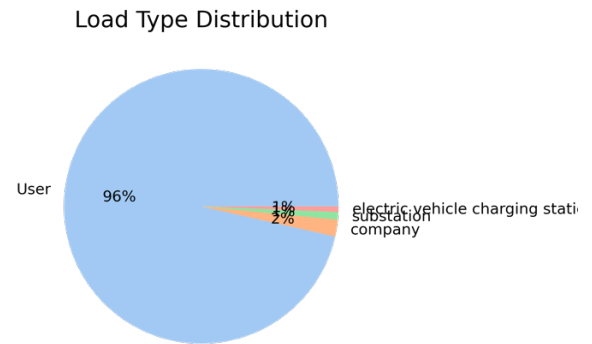
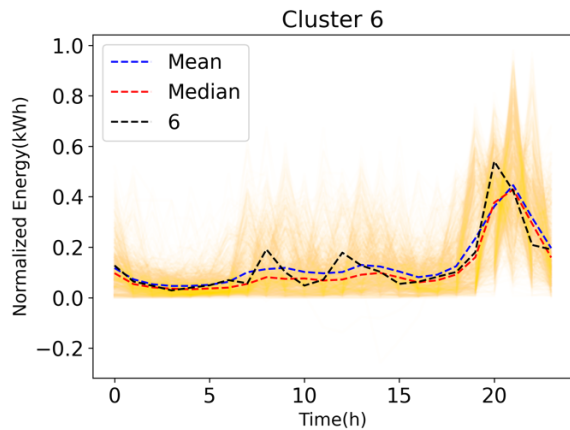


Αναλύουμε τους 2 αυτούς clusters μαζί, γιατί, έχουν παρόμοιο σχήμα, με κορυφή τις πρωινές ώρες (ο 3 από τις 8 μέχρι τις 10 και ο 8 από τις 9 μέχρι τις 11). Διαφορά υπάρχει στις χρήσεις των φορτίων που ανήκουν σε αυτούς, με τον cluster 8 να περιέχει ποσοστιαία περισσότερες καταναλώσεις από οικιακά φορτία (40% έναντι 29%) και επίσης κάποιες καταναλώσεις από τη δημόσια πισίνα.

Όπως είδαμε στην ενότητα 6.2, τις ώρες 11:00 με 14:00 είναι που κορυφώνεται η αντίστροφη ροή φορτίου. Γι' αυτό το λόγο, τα φορτία που ανήκουν στους clusters 3 και 8 μπορούν να συνεισφέρουν στην μείωση του φαινομένου αν μετακυλίσουν την κορυφή στην κατανάλωσή τους αργότερα, στο διάστημα 11:00 με 14:00.

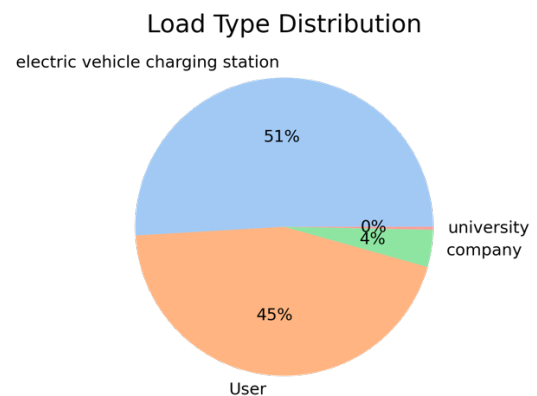
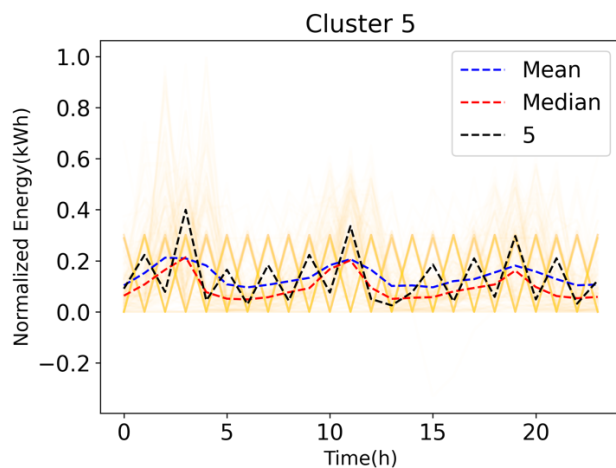
#### Cluster 4&6:





Περιέχουν προφίλ κατανάλωσης με κορυφή αργά το απόγευμα - νωρίς το βράδυ (από τις 18:00 μέχρι τις 22:00) και αποτελούνται κατά μεγάλο ποσοστό (93 και 96% αντίστοιχα) από καταναλώσεις οικιακών χρηστών, πράγμα αναμενόμενο αφού αύξηση της κατανάλωσης αργά το απόγευμα γίνεται κυρίως σε σπίτια τα οποία είναι άδεια κατά τις ώρες εργασίας και γεμίζουν όταν οι κάτοικοι γυρίζουν από τη δουλειά. Επειδή η ζήτηση στο ιταλικό δίκτυο συχνά κορυφώνεται αργά το απόγευμα (ενότητα 6.2), τα φορτία που ανήκουν στον cluster μπορούν να ενταχθούν σε ένα πρόγραμμα DR με σκοπό τη μείωση της κατανάλωσης αυτές τις ώρες, κάνοντας μετακύλιση φορτίου (πιο πριν ή πιο μετά, ανάλογα με τις ανάγκες του δικτύου και τους περιορισμούς του κάθε χρήστη).

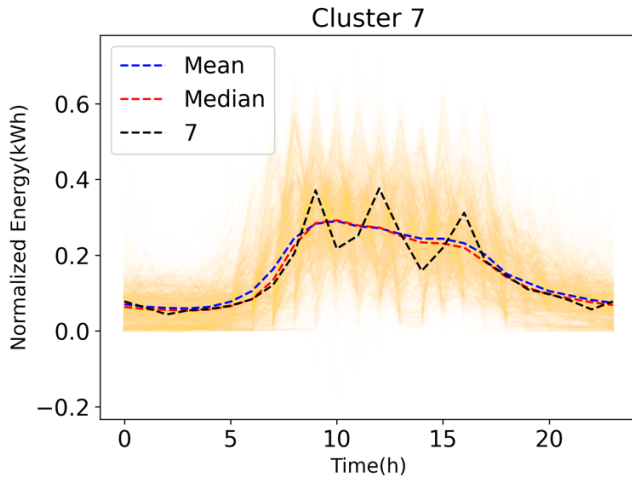
**Cluster 5:**



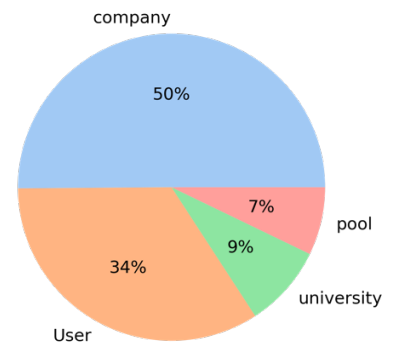
Αντιστοιχεί σε καταναλώσεις που δεν έχουν μία μεγάλη κορυφή μέσα στην μέρα, αλλά πολλές κορυφές παρόμοιου μεγέθους. Περιέχει προφίλ από οικιακούς χρηστές και μεγάλο μέρος των προφίλ ενός φορτιστή ηλεκτρικών αυτοκινήτων (28% των καταναλώσεων του φορτιστή ανήκουν στον cluster). Επειδή ο cluster δεν χαρακτηρίζεται από μία συγκεκριμένη κορυφή, είναι πιο δύσκολο να εξαχθούν συμπεράσματα για το πως θα μπορούσαν να χρησιμοποιηθούν

τα φορτία που ανήκουν σε αυτόν σε ένα πρόγραμμα DR. Για πιο ασφαλή συμπεράσματα χρειάζεται ανάλυση του κάθε φορτίου ξεχωριστά.

**Cluster 7:**

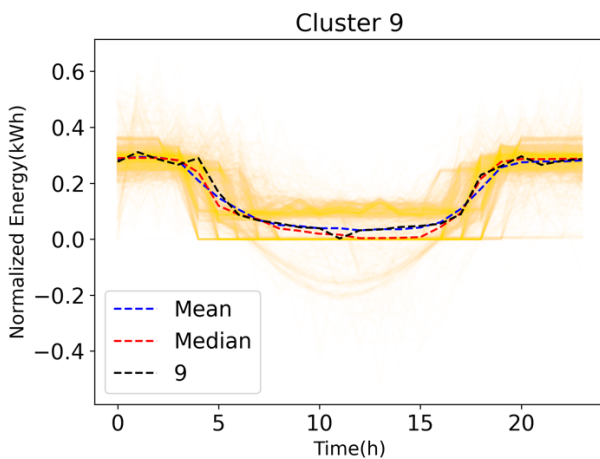


Load Type Distribution

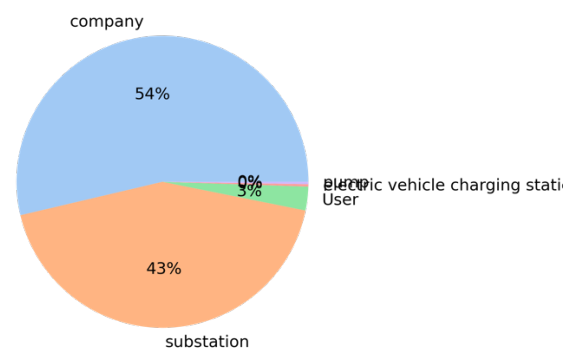


Αποτελείται από ημερήσια προφίλ που έχουν μεγάλη κατανάλωση τις ώρες εργασίας (περίπου από 7:00 μέχρι 19:00), χωρίς να εμφανίζουν κάποια απότομη κορυφή όπως άλλοι clusters, αλλά πολλές μικρότερες κορυφές (το κέντρο του cluster έχει 3 κορυφές). Τα μέλη του cluster θα μπορούσαν να βοηθήσουν στην μείωση του φαινομένου της έκχυσης φορτίου, αυξάνοντας την κατανάλωση τις ώρες που η παραγωγή είναι μεγάλη, αλλά και στην μείωση της συμφόρησης του δικτύου νωρίς το απόγευμα, μεταφέροντας την κατανάλωση από αυτές τις ώρες, λίγο νωρίτερα. Είναι αξιοσημείωτο ότι ο cluster περιέχει κυρίως καταναλώσεις εταιριών (50%) και κάποιων εγκαταστάσεων με πολλές δυνατότητες για παροχή ευελιξίας, πχ. το τοπικό πανεπιστήμιο και μία δημόσια πισίνα.

**Cluster 9:**

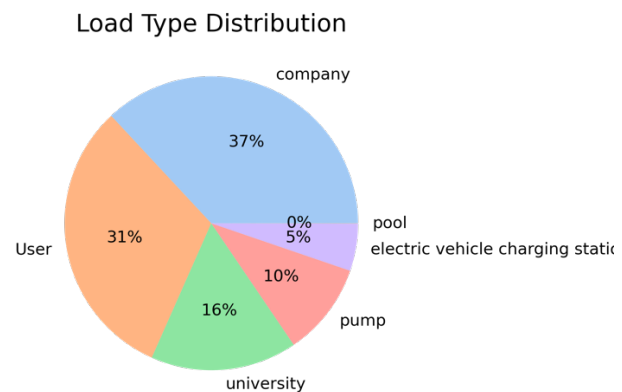
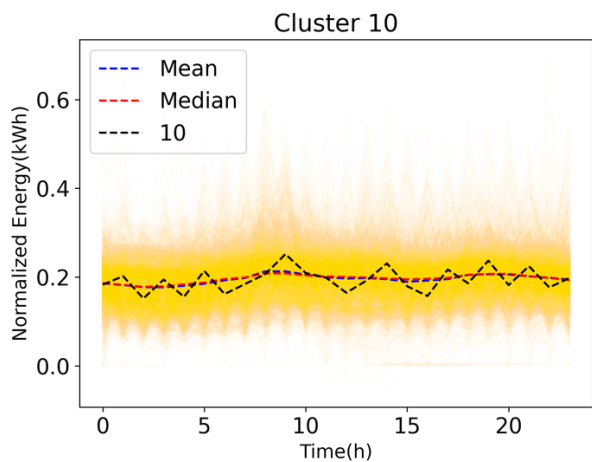


Load Type Distribution



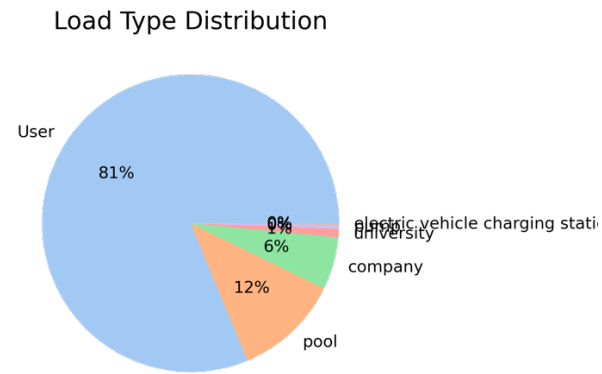
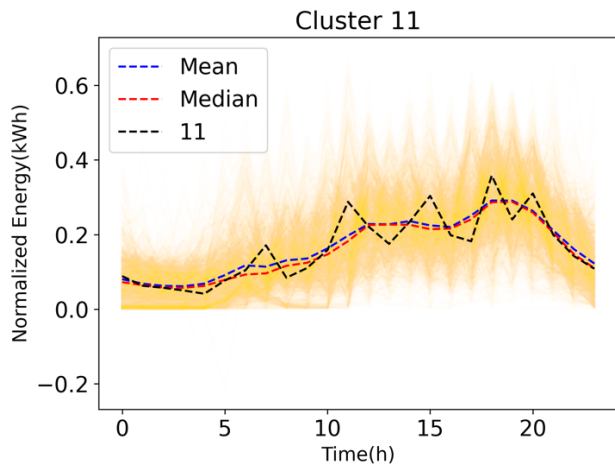
Αντιστοιχεί σε μικτά φορτία, δηλαδή εγκαταστάσεις που διαθέτουν και κατανάλωση και παραγωγή ενέργειας. Αυτό φαίνεται από το πως μειώνεται η συνολική κατανάλωση τις ώρες που έχει ήλιο (6:00 με 20:00). Επειδή στα ημερήσια προφίλ που ανήκουν σε αυτόν τον cluster έχουμε σύμπλεξη της παραγωγής και της κατανάλωσης, είναι δύσκολο να εξαχθούν συμπεράσματα για το πως μπορεί να μεταβληθεί η κατανάλωση των φορτίων για τους στόχους ενός προγράμματος DR. Για να γίνει αυτό χρειάζεται ένας τρόπος να χωριστεί κάθε ημερήσιο προφίλ σε ένα προφίλ κατανάλωσης και ένα προφίλ παραγωγής.

### Cluster 10:



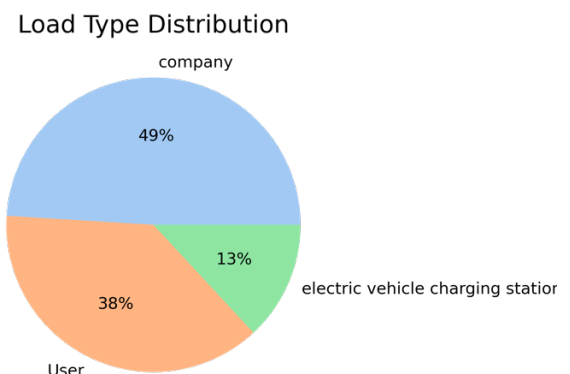
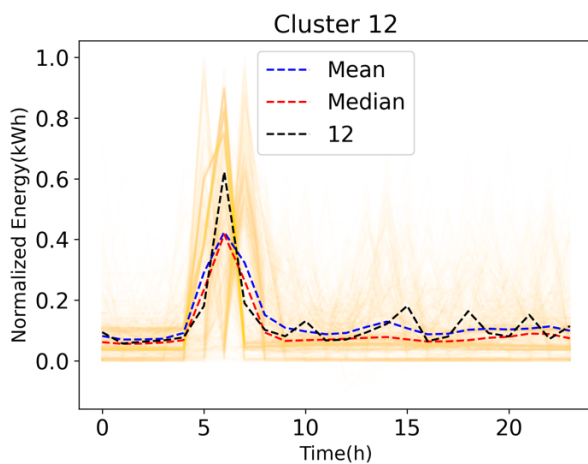
Αυτός ο cluster δεν εμφανίζει κάποια κορυφή και κατά πάσα πιθανότητα αντιστοιχεί σε μέρες που το φορτίο δεν χρησιμοποιείται, πχ. όταν ένα σπίτι είναι άδειο ή μία επιχείρηση είναι κλειστή. Κάτι που αξίζει να σημειωθεί είναι ότι και το μεγαλύτερο μέρος των καταναλώσεων του τοπικού πανεπιστημίου ανήκουν σε αυτόν τον cluster, όχι γιατί το πανεπιστήμιο είναι άδειο, αλλά γιατί η κατανάλωση του δεν κορυφώνεται κάποια στιγμή μέσα στη μέρα, αλλά είναι σχετικά σταθερή, με μία μικρή αύξηση τις ώρες που υπάρχει μεγαλύτερη διδακτική δραστηριότητα. Επειδή δεν υπάρχουν κορυφές μέσα στη μέρα, δεν είναι ξεκάθαρο το πως τα φορτία που ανήκουν σε αυτόν τον cluster θα μπορούσαν να βοηθήσουν σε ένα πρόγραμμα DR. Ωστόσο, επειδή το πανεπιστήμιο έχει πολύ μεγάλη κατανάλωση, ακόμα και μικρές σχετικές μεταβολές στην κατανάλωσή του τις ώρες που αυτό χρειάζεται, θα βοηθούσε πολύ στην επίτευξη των στόχων ενός τέτοιου προγράμματος.

### Cluster 11:



Ο cluster 11 μοιάζει με τον cluster 7 στο ότι δεν έχει μία μεγάλη κορυφή αλλά πολλές μικρότερες κορυφές μέσα στη μέρα. Μία διαφορά είναι ότι τα προφίλ του ο 7 φαίνεται να έχουν μεγαλύτερη κατανάλωση τις πρωινές ώρες, ενώ τα προφίλ του 11 έχουν μεγαλύτερη κατανάλωση αργά το απόγευμα. Με αυτό το δεδομένο, φαίνεται λογικό ότι το μεγαλύτερο μέρος των προφίλ του 7 ανήκουν σε εταιρίες (κατά 50%), ενώ του 11 σε οικιακούς καταναλωτές (κατά 81%). Όπως και ο 7, έτσι και ο 11 θα μπορούσε να χρησιμοποιηθεί για την αντιμετώπιση τόσο του φαινομένου του reverse power flow όσο και της συμφόρησης τις ώρες αιχμής.

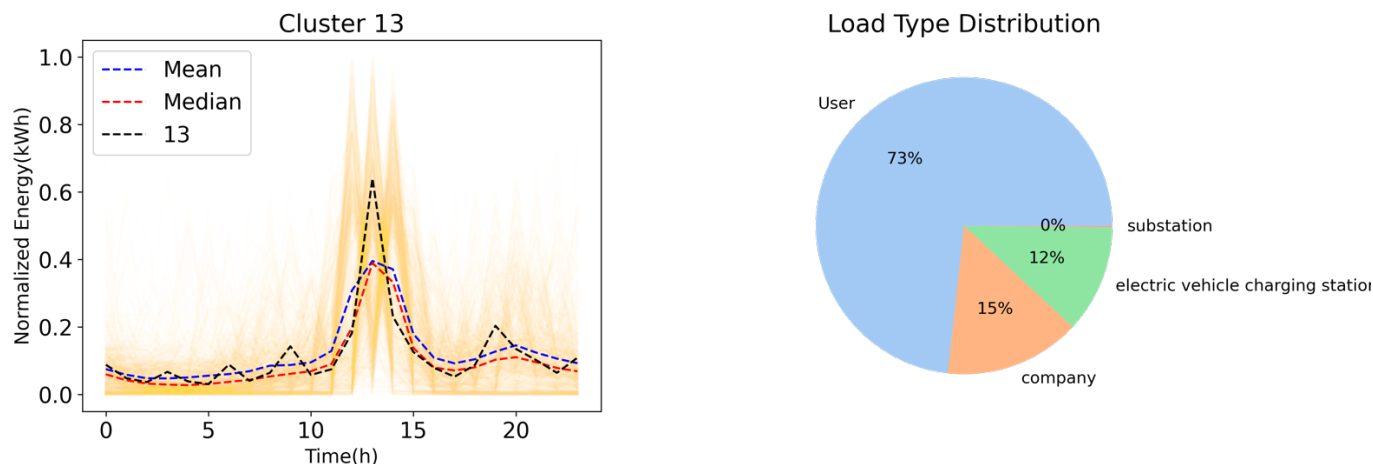
**Cluster 12:**



Περιέχει καταναλώσεις από εταιρίες ή οικιακούς χρήστες που κορυφώνονται νωρίς το πρωί (5:00 με 8:00 το πρωί). Δεδομένου ότι τόσο νωρίς το πρωί συνήθως δεν υπάρχει συμφόρηση στο δίκτυο και οι ώρες που δημιουργείται το πρόβλημα τις έκχυσης φορτίου είναι σχετικά μακριά, δεν υπάρχει λόγος τα φορτία που ανήκουν σε αυτούς τους clusters να ενταχθούν σε πρόγραμμα DR

**Cluster 13:**





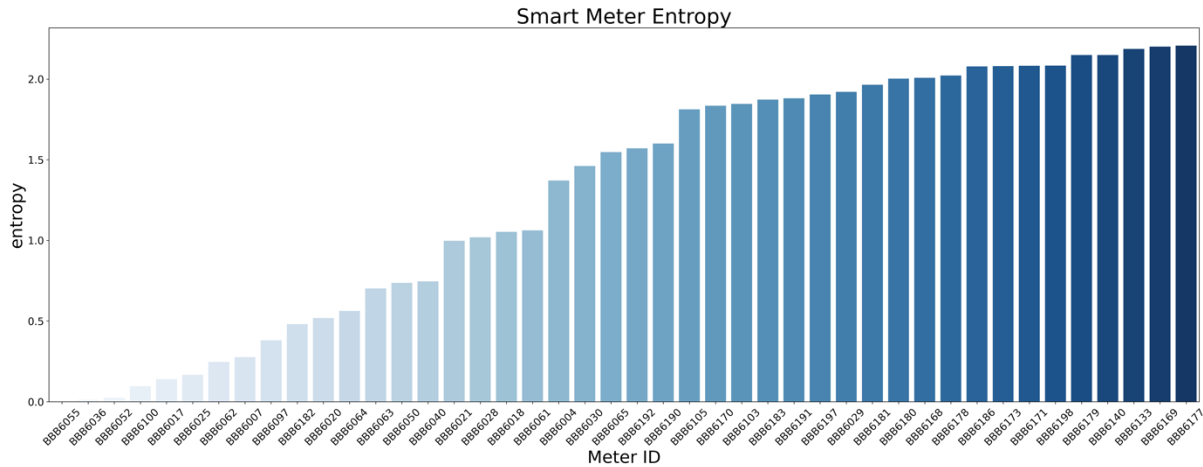
Τα ημερήσια προφίλ που ανήκουν σε αυτόν τον cluster προέρχονται κυρίως από οικιακούς χρήστες (κατά 73%) και χαρακτηρίζονται από μία κορυφή κοντά στο μεσημέρι (από τις 12:00 μέχρι τις 14:00). Μιας και αυτές είναι οι ώρες που κορυφώνεται η αντίστροφη ροή φορτίου, οι καταναλωτές που ανήκουν σε αυτόν τον cluster βοηθούν ήδη στην επίλυση του προβλήματος και δεν θα θέλαμε να αλλάξουν τις συνήθειές τους.

Με βάση τα παραπάνω οι συστάδες 2, 3, 7, 8, 11 θα μπορούσαν να βοηθήσουν στην μείωση του reverse power flow και οι 2, 4, 6, 7, 11 στην μείωση της κατανάλωσης της ώρες αιχμής. Σχετικά με τις χρήσεις των φορτίων, οι συστάδες 2, 4, 5, 6, 11, 13 περιέχουν περισσότερες καταναλώσεις από οικιακούς χρήστες απ' ότι εμπορικούς, ενώ οι συστάδες 0, 3, 8, 9, 12 το αντίθετο. Οι συστάδες 1,7 και 10 είναι πιο ισορροπημένες.

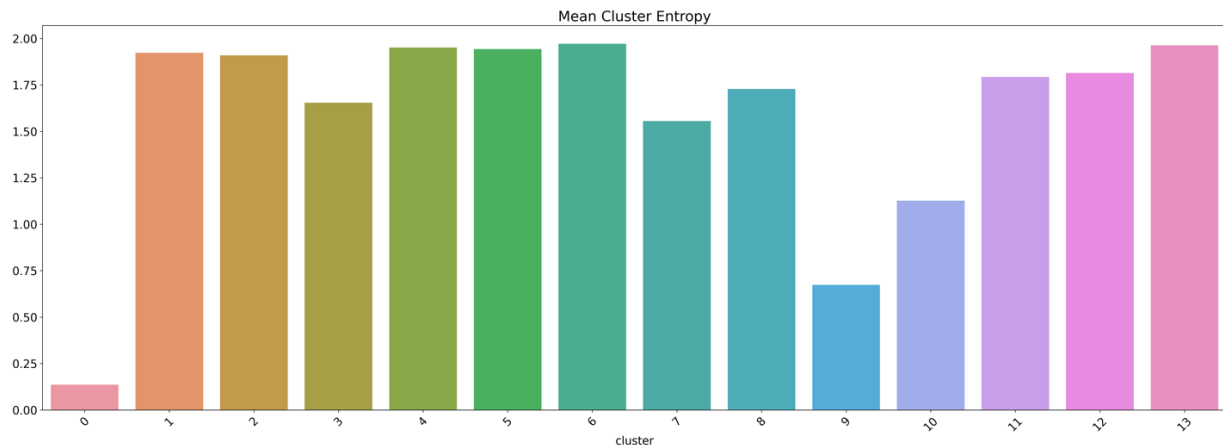
### 7.3 Ανάλυση εντροπίας

Όπως είδαμε στην ενότητα 0 ένα ακόμα χαρακτηριστικό που επηρεάζει τη διαδικασία επιλογής καταναλωτών για ένα πρόγραμμα απόκρισης ζήτησης είναι η εντροπία ή μεταβλητότητα στις καταναλωτικές τους συνήθειες.

Στο διάγραμμα της εικόνας Εικόνα 12 οι μετρητές τους συνόλου δεδομένων είναι τοποθετημένοι σε αύξουσα σειρά ανάλογα με την εντροπία. Στο διάγραμμα της Εικόνα 13 παρουσιάζεται η μέση εντροπία για κάθε μία τις 14 συστάδες. Η μέση εντροπία κάθε συστάδας ορίζεται ως ο σταθμισμένος μέσος όρος των εντροπιών των καταναλωτών που ανήκουν σε αυτή, με το βάρος κάθε καταναλωτή να είναι ίσο με το ποσοστό των καταναλώσεων της συστάδας που του αντιστοιχούν.



**Εικόνα 12 - Εντροπία για κάθε μετρητή στο σύνολο δεδομένων. Οι μετρητές έχουν ταξινομηθεί ανάλογα με την εντροπία (χαμηλή εντροπία αριστερά, υψηλή δεξιά)**



**Εικόνα 13 - Μέση εντροπία για κάθε συστάδα**

Παρατηρούμε ότι:

1. Οι συστάδες με τη μικρότερη εντροπία είναι αυτές που περιέχουν προφίλ από εγκαταστάσεις παραγωγής ενέργειας ή μικτά φορτία (συστάδες 0 και 9). Αυτό είναι αναμενόμενο γιατί σε αυτά τα προφίλ περιέχεται και παραγωγή και κατανάλωση (στη συστάδα 0 μόνο παραγωγή) και η παραγωγή έχει πολύ μικρή μεταβλητότητα (το σχήμα της καμπύλης παραγωγής είναι σχεδόν ίδιο κάθε μέρα: μία κοιλάδα τις ώρες που έχει ήλιο).

2. Γενικά επιβεβαιώνεται η διαπίστωση της ενότητας 0 ότι οι συστάδες που αποτελούνται περισσότερο από καταναλώσεις εταιρειών έχουν μικρότερη μέση εντροπία από τις συστάδες που αποτελούνται περισσότερο από καταναλώσεις οικιακών χρηστών. Ενδεικτικά, από τις 5 συστάδες με τη μεγαλύτερη μέση εντροπία (1,4,5,6,13) και οι 5 έχουν περισσότερες καταναλώσεις οικιακών χρηστών απ' ότι εταιριών. Για τις 5 συστάδες με τη μικρότερη μέση εντροπία (0,3,7,9,10) συμβαίνει το αντίθετο.

## 8. Συμπεράσματα – Προτάσεις

Σκοπός του κεφαλαίου αυτού είναι η χρήση των αποτελεσμάτων του κεφαλαίου 6 σε συνδυασμό με τους μεθοδολογικούς κανόνες της ενότητας 6.8 ώστε να εξάγουμε συμπεράσματα για το ποιες συστάδες είναι πιο ιδανικές για απόκριση ζήτησης και τι είδους προγράμματα θα ταίριαζαν σε κάθε συστάδα.

Ωστόσο ένα πρόβλημα που προέκυψε είναι ότι αυτές οι συστάδες δεν ήταν συστάδες καταναλωτών αλλά καταναλώσεων, δηλαδή ένας καταναλωτής δεν ανήκε σε μία μόνο συστάδα αλλά πολλές. Αυτό θα δυσκόλευε την εξαγωγή συγκεκριμένων προτάσεων για απόκριση ζήτησης, γιατί θα υπήρχαν καταναλωτές που ανήκουν σε συστάδες με διαφορετικές ανάγκες και προτεινόμενα προγράμματα. Στη συνέχεια της εργασίας ακολουθήθηκαν δύο διαφορετικές προσεγγίσεις για την διαχείριση αυτής της ιδιαιτερότητας:

1. στην ενότητα 8.1 θεωρήσαμε ότι κάθε καταναλωτής ανήκει στη συστάδα που περιέχει τις περισσότερες καταναλώσεις του. Έτσι, κάθε καταναλωτής πλέον ανήκε σε μία μόνο συστάδα και με βάση τα χαρακτηριστικά της συστάδας από το κεφάλαιο 7 (σχήμα, χρήσεις φορτίων, μέση εντροπία) μπορέσαμε να προτείνουμε τη δημιουργία προγραμμάτων με συγκεκριμένη δομή και σκοπό για κάθε συστάδα και τελικά για κάθε καταναλωτή
2. στην ενότητα 8.2 προτείναμε τη δημιουργία ενός ολοκληρωμένου συστήματος DR που λειτουργεί καθημερινά και παράγει προτάσεις για κάθε καταναλωτή (πχ. τιμολογιακή πολιτική τύπου RTP), δημιουργώντας πρώτα μία πρόγνωση για την επόμενη μέρα και χρησιμοποιώντας αυτή την πρόγνωση για να τοποθετήσει τον καταναλωτή σε μία από τις συστάδες

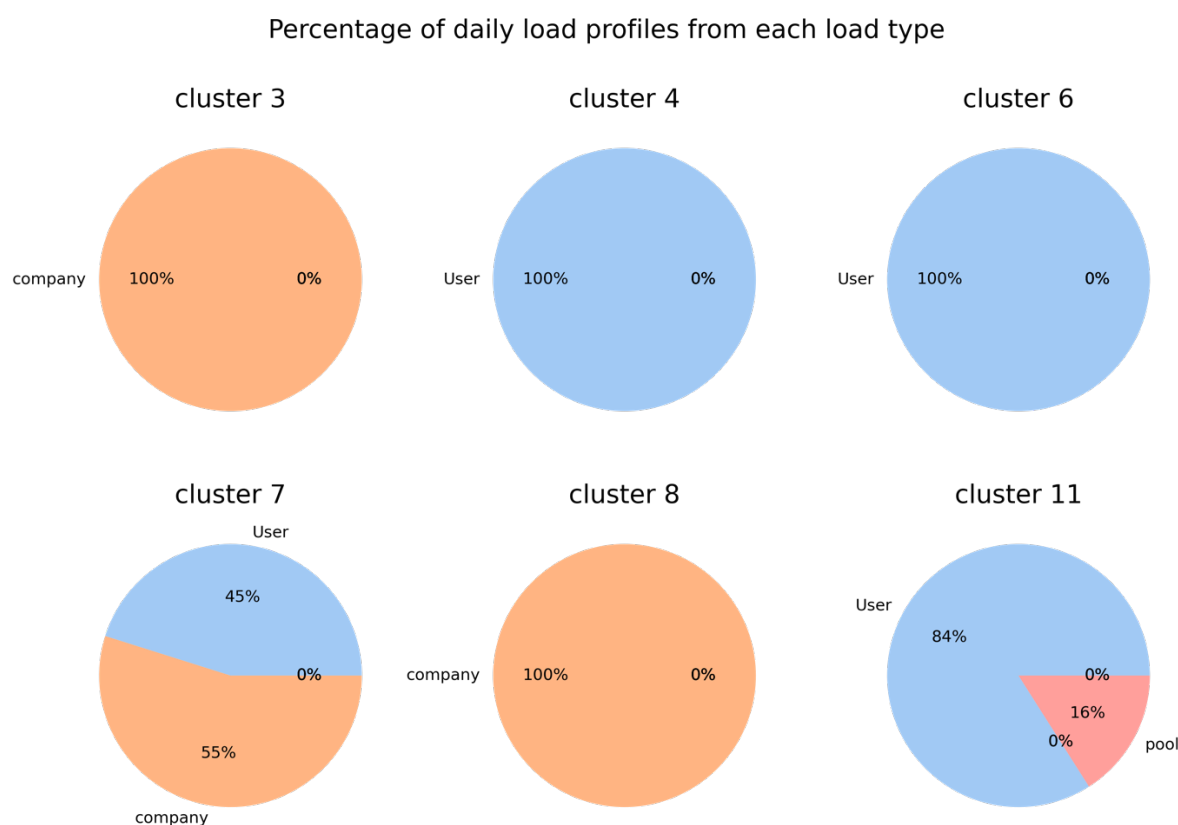
### 8.1 Προτάσεις ανά συστάδα

#### 8.1.1 Χρήσεις φορτίων και μέση εντροπία ανά συστάδα

Σε αυτή την ενότητα καταλήξαμε σε προτάσεις για κάθε μία από τις συστάδες που δείξαμε (ενότητα 7.2) ότι μπορούν να βοηθήσουν στην επίλυση των προβλημάτων της αντίστροφης ροής φορτίου και της συμφόρησης τις ώρες αιχμής, δηλαδή τις συστάδες 2,3,4,6,7,8 και 11.

Όπως αναφέραμε στην εισαγωγή του κεφαλαίου 6, για να παράγουμε αυτές τις προτάσεις θεωρήσαμε ότι κάθε φορτίο ανήκει στη συστάδα που περιέχει τις περισσότερες καταναλώσεις του. Αυτό σημαίνει ότι τα ποσοστά των καταναλώσεων των συστάδων που προέρχονταν από κάθε χρήση και η μέση εντροπία κάθε συστάδας διέφεραν από αυτά που είδαμε στις ενότητες 7.2 και 7.3. Συγκεκριμένα:

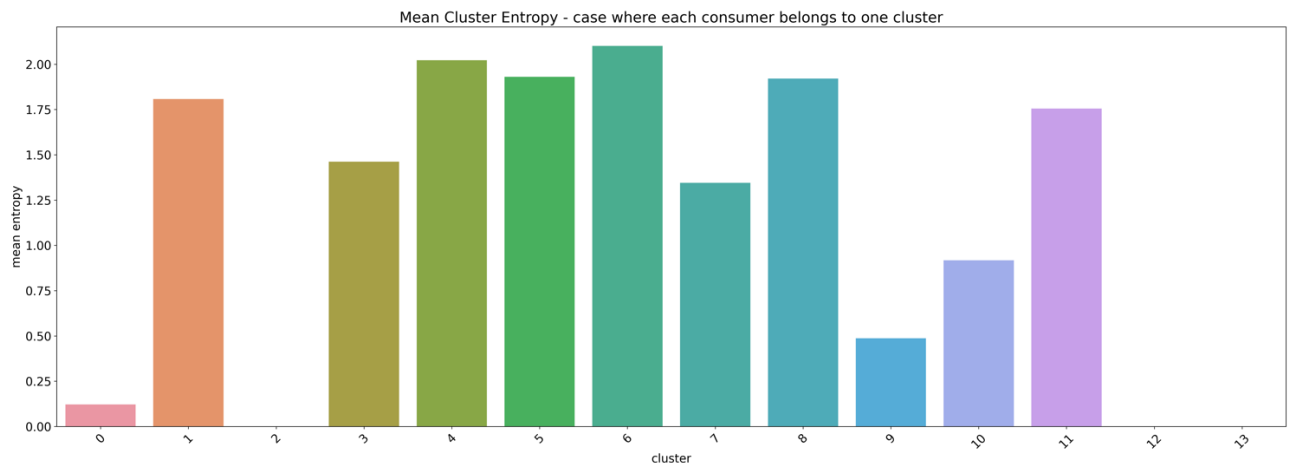
- **Χρήσεις των φορτίων που ανήκουν σε κάθε συστάδα:** στο διάγραμμα της εικόνας Εικόνα 14 έχουμε τα ποσοστά των δειγμάτων κάθε συστάδας που αντιστοιχούν σε κάθε χρήση – τύπο φορτίου. Παρατηρούμε ότι πλέον οι συστάδες 3 και 8 αποτελούνται αμιγώς από καταναλώσεις εμπορικών φορτίων και οι συστάδες 4 και 6 από καταναλώσεις οικιακών. Η συστάδα 7 έχει καταναλώσεις τόσο από οικιακούς χρήστες όσο και εταιρείες, ενώ η συστάδα 11 έχει καταναλώσεις από οικιακά φορτία και από τη δημόσια πισίνα.



**Εικόνα 14 – Ποσοστά προφίλ κάθε συστάδας που αντιστοιχούν στις διαφορετικές χρήσεις (περίπτωση που κάθε φορτίου ανήκει σε μία μόνο συστάδα)**

- **Μέση εντροπία συστάδας:** στο διάγραμμα της εικόνας Εικόνα 15 βλέπουμε τη μέση εντροπία για κάθε μία από τις 14 συστάδες. Αυτή οι τιμή προέκυψε με τον ίδιο τρόπο που προέκυψε στην ενότητα 7.3 (δηλαδή ως ο σταθμισμένος μέσος όρος των εντροπιών των καταναλωτών που ανήκουν στη συστάδα), μόνο που εδώ λήφθηκαν υπόψη μόνο οι εντροπίες των καταναλωτών που ανήκουν αποκλειστικά στη συστάδα. Οι τιμές της εντροπίας κάθε καταναλωτή είναι αυτές που υπολογίστηκαν στην ενότητα

7.3 με βάση τον τύπο της ενότητας **Error! Reference source not found.**, παρόλο που για την εξαγωγή των συμπερασμάτων της παρούσας ενότητας θεωρήσαμε ότι κάθε καταναλωτής ανήκει σε μία μόνο συστάδα. Οι συστάδες 2, 12 και 13 δεν έχουν τιμή εντροπίας στο σχήμα γιατί μετά από την αλλαγή στους κανόνες συμμετοχής των καταναλωτών στις συστάδες κανένας καταναλωτής δεν ανήκε σε αυτές.

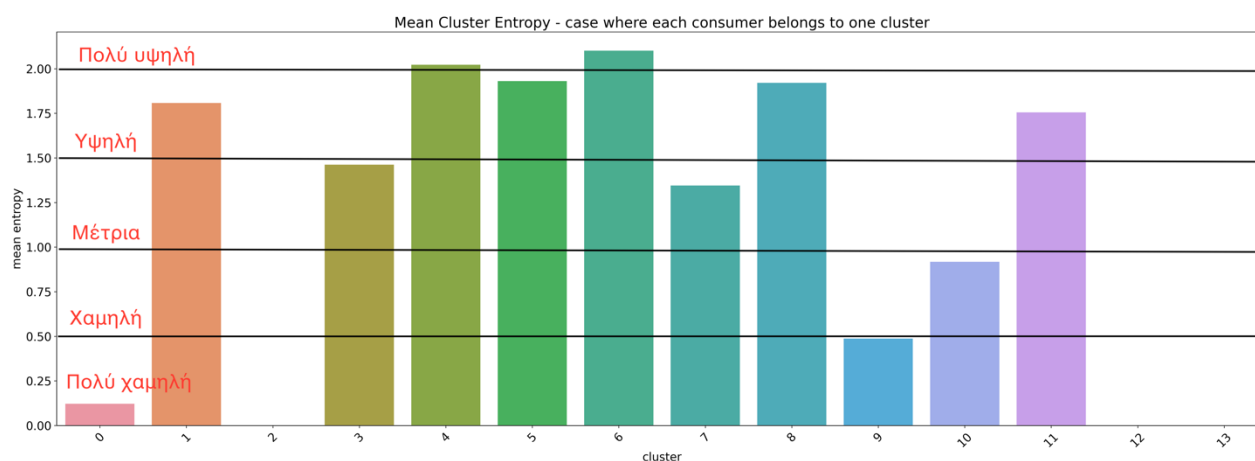


**Εικόνα 15 – Μέση εντροπία κάθε συστάδας (περίπτωση που κάθε φορτίο ανήκει σε μία συστάδα)**

### 8.1.2 Κατηγορικός χαρακτηρισμός της εντροπίας

Στην ενότητα 4.2.2 αναφέρθηκαν κάποιες προτάσεις που υπάρχουν στη βιβλιογραφία για τους τύπους προγραμμάτων που ταιριάζουν σε καταναλωτές χαμηλής εντροπίας και υψηλής εντροπίας. Ωστόσο, στις εργασίες που μελετήθηκαν δεν υπήρχε κάποια αντικειμενική στάθμη για να αποφασίσουμε ποιες τιμές εντροπίας θεωρούνται υψηλές και ποιες χαμηλές. Αυτό μπορούσε να γίνει μόνο συγκρίνοντας τις εντροπίες των συστάδων μεταξύ τους.

Συγκεκριμένα, αποφασίσαμε να χωρίσουμε το πεδίο τιμών ανά 0.5 και καταλήξαμε σε 5 «περιοχές» στις οποίες δώσαμε ένα χαρακτηρισμό: πολύ χαμηλή, χαμηλή, μέτρια, υψηλή και πολύ υψηλή εντροπία. Αυτός ο διαχωρισμός φαίνεται στο διάγραμμα της εικόνας Εικόνα 16.



Εικόνα 16 – κατηγορικός χαρακτηρισμός της εντροπίας

### 8.1.3 Τελικές προτάσεις

Στον πίνακα Πίνακας 3 οργανώσαμε τα αποτελέσματα του κεφαλαίου 7 και τα παραπάνω συμπεράσματα για κάθε συστάδα που μας ενδιέφερε. Στη συνέχεια, χρησιμοποιώντας τους γενικούς κανόνες της ενότητας 6.8, καταλήξαμε σε ένα προτεινόμενο πρόγραμμα απόκρισης ζήτησης για κάθε συστάδα.

**Πίνακας 3 – προτεινόμενα προγράμματα για κάθε συστάδα και χαρακτηριστικά που βοήθησαν στην επιλογή**

Συστάδα	Συνεισφέρει στην αντίστροφη ροή φορτίου	Επιβαρύνει τις ώρες αιχμής	Εντροπία	Χρήση	Προτεινόμενο πρόγραμμα
3	✓		Μέτρια	Εταιρείες	1
4		✓	Πολύ Υψηλή	Κατοικίες	1*,2,3
6		✓	Πολύ Υψηλή	Κατοικίες	1*,2,3
7	✓	✓	Μέτρια	Μικτή	1
8	✓		Υψηλή	Εταιρείες	1*,3
11	✓	✓	Υψηλή	Κατοικίες + Δημόσια Πισίνα	1*,2,3

**Πρόγραμμα 1 (TOU):** Το πιο απλό πρόγραμμα προκειμένου να έχουμε μείωση της αντίστροφης ροής φορτίου και της κατανάλωσης τις ώρες αιχμής και αυτό που χρειάζεται τις λιγότερες επενδύσεις τόσο για τους συμμετέχοντες όσο και για την ASM είναι ένα πρόγραμμα Time Of Use. Αφού επιθυμούμε να κινητοποιήσουμε τους καταναλωτές να καταναλώνουν περισσότερο από τις 11 π.μ. μέχρι τις 2 μ.μ. (ενότητα ...) πρέπει να θέσουμε την τιμή χαμηλά εκείνες τις ώρες και συγκριτικά πιο ψηλά νωρίτερα (8 π.μ. με 11 π.μ.) και αργότερα (2 μ.μ. με 5 μ.μ.). Το αντίθετο πρέπει να γίνει για τις ώρες αιχμής (5 μ.μ. με 7 μ.μ.), δηλαδή η τιμή πρέπει να είναι υψηλή εκείνες τις ώρες και συγκριτικά πιο χαμηλή νωρίτερα (2 μ.μ. με 5 μ.μ.) και αργότερα (8 μ.μ. με 12 μ.μ.). Τελικά μπορούμε να προτείνουμε ένα πρόγραμμα με την εξής τιμολόγηση:

Περίοδος	Τιμή
12 π.μ. με 7 π.μ.	ΧΑΜΗΛΗ
7 π.μ με 11 π.μ.	ΜΕΣΑΙΑ
11 π.μ. με 2 μ.μ.	ΧΑΜΗΛΗ
2 μ.μ. με 5 μ.μ.	ΜΕΣΑΙΑ
5 μ.μ. με 7 μ.μ	ΥΨΗΛΗ
8 μ.μ. με 12 μ.μ	ΜΕΣΑΙΑ

Στην ενότητα 4.2 αναφέραμε ότι τα προγράμματα Time Of Use ταιριάζουν περισσότερο σε καταναλωτές με χαμηλή εντροπία και σε εμπορικά φορτία. Γι' αυτό και το πρόγραμμα προτείνεται για τις συστάδες 1 και 7. Ωστόσο, επειδή είναι απλό στην εφαρμογή και χρειάζεται μικρότερες επενδύσεις σε εξοπλισμό και εργατοώρες για να λειτουργήσει, θα μπορούσε να εφαρμοστεί και στις συστάδες με οικιακούς καταναλωτές και υψηλότερη εντροπία.

### Πρόγραμμα 2 (CPP)

Όπως αναφέραμε στην ενότητα 4.1.2, τα γεγονότα απόκρισης ζήτησης στα προγράμματα CPP έχουν περιστασιακό χαρακτήρα αλλά κατά τη διάρκεια των γεγονότων αποδίδουν πιο σημαντικά αποτελέσματα (μεγαλύτερη αλλαγή στη συμπεριφορά του καταναλωτή) από τα προγράμματα TOU. Ένα πρόγραμμα CPP για το δικό μας use case θα μπορούσε να ενεργοποιείται 20-30 φορές το χρόνο όταν προβλέπεται ότι τα προβλήματα στο δίκτυο θα είναι έντονα (πολύ μεγάλη αντίστροφη ροή φορτίου ή πολύ μεγάλο peak τις απογευματινές ώρες). Η ASM θα προειδοποιεί τους καταναλωτές για τα DR γεγονότα μία μέρα πριν, ανακοινώνοντας την τιμή της ενέργειας για την περίοδο του γεγονότος και προτείνοντας τι μπορούν να κάνουν για να βοηθήσουν το δίκτυο και την τσέπη τους (πχ. να θερμάνουν το νερό ή το χώρο νωρίτερα από το γεγονός αν απαιτείται μείωση της συνολικής κατανάλωσης). Ένα άλλο πλεονέκτημα ενός τέτοιου προγράμματος σε σχέση με το TOU είναι ότι μπορεί να προσαρμοστεί καλύτερα



στις ιδιαιτερότητες της κάθε μέρας και να εφαρμοστεί και σε έκτακτα γεγονότα, ενώ το ΤΟΥ λειτουργεί με συγκεκριμένη τιμολογιακή δομή κάθε μέρα. Βέβαια, όπως αναφέρεται στην ενότητα 4.2.2, τα προγράμματα CPP ταιριάζουν καλύτερα σε καταναλωτές με υψηλή εντροπία και οικιακούς καταναλωτές. Γι' αυτό και προτείνονται για τις συστάδες 4,6 και 11.

**Πρόγραμμα 3 (RTP):** Ένα πρόγραμμα RTP θα μπορούσε να προσαρμόζει την τιμή του ρεύματος ανάλογα με τις ανάγκες του δικτύου πιο αποτελεσματικά από τα προγράμματα ΤΟΥ και CPP. Ωστόσο, όπως αναφέρεται στην ενότητα 4.1.3, προκειμένου οι καταναλωτές να αντιδρούν στις τιμές μεταβάλλοντας την κατανάλωσή τους θα χρειάζονται εξοπλισμό που να τους επιτρέπει χρονοπρογραμματισμό ή αυτόματο χειρισμό των διάφορων φορτίων, όπως έξυπνες συσκευές και ελεγκτές. Εφόσον δεν υπάρχουν τέτοιες συσκευές, θα ήταν μάλλον προτιμότερο να εφαρμοστεί ένα από τα προγράμματα 1 και 2. Εάν όμως υπάρχει η διάθεση για επένδυση σε τέτοιο εξοπλισμό, ένα πρόγραμμα RTP θα μπορούσε να αποδώσει καλά, ιδιαίτερα για τις συστάδες με μεγάλη εντροπία (4,6,8 και 11). Προφανώς, ο μηχανισμός που θα παράγει τις τιμές θα πρέπει να είναι έτσι σχεδιασμένος ώστε να δίνει χαμηλές τιμές όταν προβλέπεται ότι θα υπάρχει αντίστροφη ροή φορτίου και υψηλές όταν προβλέπεται ότι θα υπάρχει συμφόρηση στο τοπικό δίκτυο ή ότι η τιμή του ρεύματος στο γενικότερο δίκτυο (το ιταλικό) θα είναι υψηλή.

## 8.2 Ολοκληρωμένο σύστημα προγνώσεων, συσταδοποίησης και DR

### 8.2.1 Περιγραφή και παράδειγμα

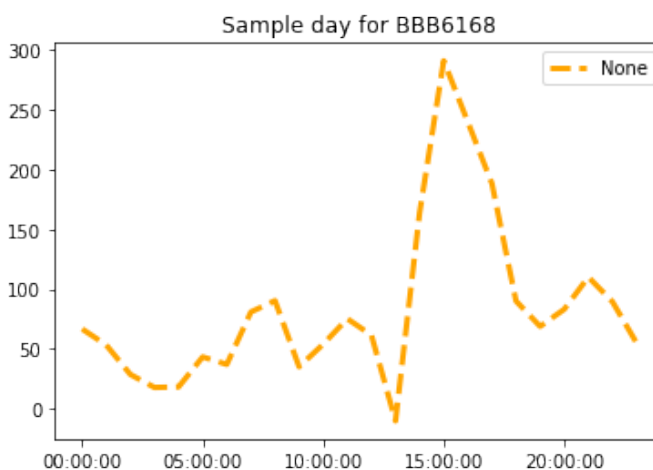
Σε αυτή την ενότητα προτείνουμε ένα σύστημα το οποίο θα χρησιμοποιεί τα αποτελέσματα της συσταδοποίησης του κεφαλαίου 7 για να δημιουργεί προτάσεις απόκρισης ζήτησης για τα φορτία του εικονικού δικτύου.

Το προτεινόμενο σύστημα αποτελείται τρία μέρη:

1. μία υπηρεσία προγνώσεων (forecasting service), με προγνωστικά μοντέλα που θα παράγουν προγνώσεις για την κατανάλωση κάθε φορτίου στο πρόγραμμα
2. μία υπηρεσία συσταδοποίησης (clustering service), με ένα μοντέλο που θα κατατάσσει κάθε πρόγνωση σε μία από τις συστάδες της ενότητας 7.2
3. ένα σύστημα προτάσεων (recommendation – pricing engine) που, με βάση την συστάδα στην οποία ανήκει κάθε πρόγνωση και τους στόχους του προγράμματος απόκρισης ζήτησης, θα παράγει προτάσεις για το πως οι υπεύθυνοι του προγράμματος μπορούν να επηρεάσουν τους καταναλωτές ώστε η κατανάλωσή τους να μην είναι αυτή που προβλέφθηκε από το πρώτο κομμάτι του συστήματος, αλλά πιο ιδανική για την επίτευξη των στόχων του προγράμματος.

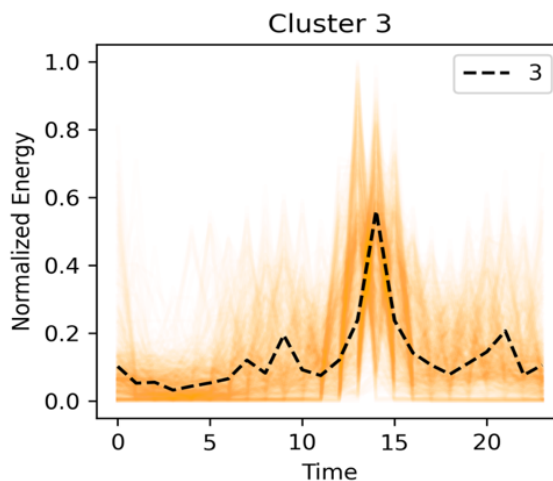
Ας δούμε ένα παράδειγμα λειτουργίας του παραπάνω συστήματος για έναν συγκεκριμένο καταναλωτή:

1. Ένα προγνωστικό μοντέλο παράγει μία πρόγνωση για την κατανάλωση που θα έχει το φορτίο με έξυπνο μετρητή BBB6168 την επόμενη μέρα. Ας πούμε ότι είναι αυτή στην Εικόνα 17. Το προγνωστικό μοντέλο επίσης προσδιορίζει, αθροίζοντας όλες της προγνώσεις κατανάλωσης και παραγωγής ενέργειας, ότι τις ώρες 11:00-13:00 η παραγωγή στο δίκτυο θα είναι μεγαλύτερη από την κατανάλωση και θα υπάρχει αντίστροφη ροή φορτίου από το δίκτυο της ASM προς τα δίκτυα με τα οποία είναι συνδεδεμένο.



**Εικόνα 17 - Πρόγνωση ημερήσιας κατανάλωσης για μετρητή BBB6168**

2. Το μοντέλο της υπηρεσίας συσταδοποίησης λαμβάνει την πρόγνωση για τον καταναλωτή με έξυπνο μετρητή BBB6168 και την κατατάσσει σε μία από τις συστάδες της ενότητας **Error! Reference source not found.**, με τον ίδιο τρόπο που θα το έκανε και ο αλγόριθμος συσταδοποίησης. Σε αυτό το παράδειγμα θα την κατέτασσε στην συστάδα 3, όπως φαίνεται στην Εικόνα 18.

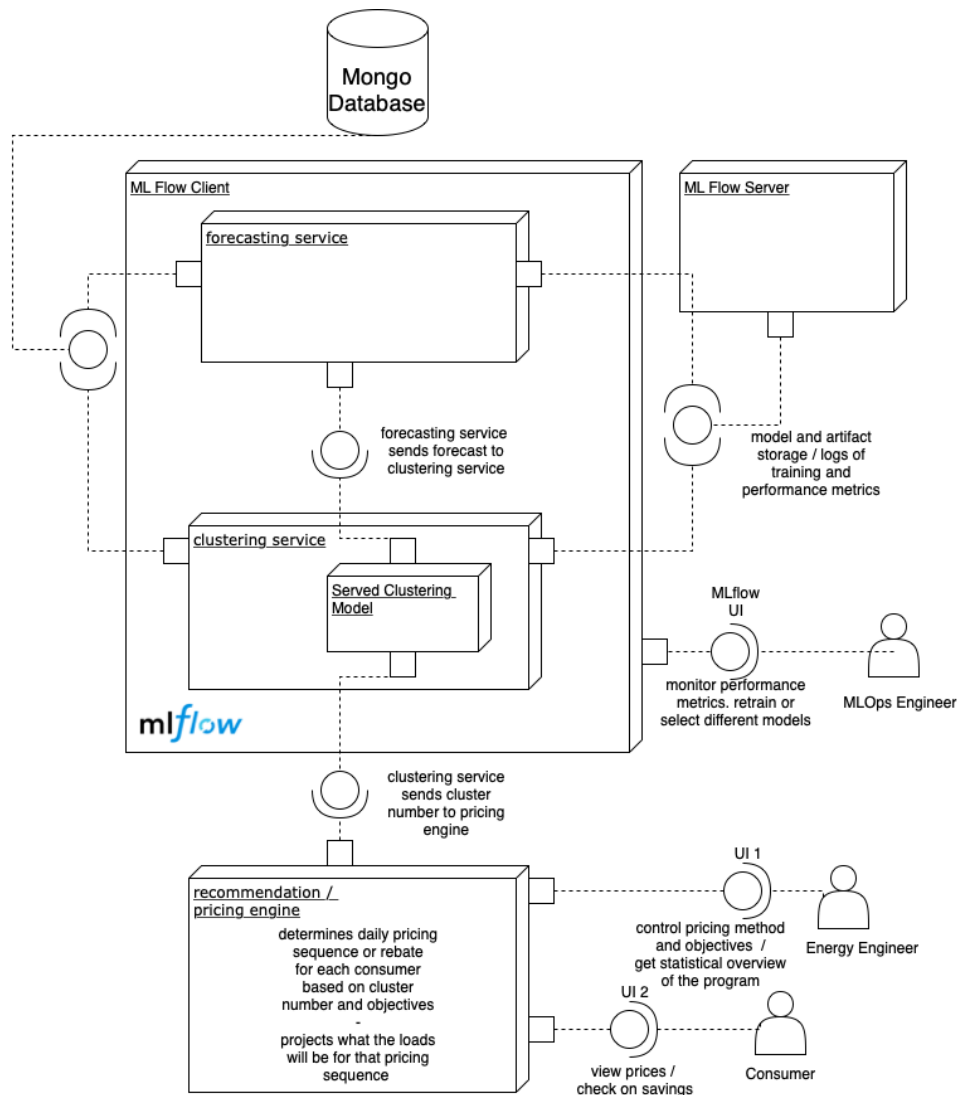


**Εικόνα 18 - Συστάδα στην οποία αντιστοιχεί η πρόβλεψη για τον μετρητή BBB6168**

3. Η συστάδα 3 έχει κορυφή στις 13:00-15:00 το μεσημερι. Θα ήταν συμφέρον για το πρόγραμμα ο καταναλωτής με έξυπνο μετρητή BBB6168 να «μετακυλίσει» αυτή την κορυφή στην κατανάλωση περίπου 2 ώρες νωρίτερα, ώστε να είναι πιο κοντά στην περίοδο 11:00-13:00 και η διαφορά παραγωγής και κατανάλωσης, και άρα η αντίστροφη ροή φορτίου, να είναι μικρότερη. Για να γίνει αυτό θα πρέπει η ASM να προσφέρει στον καταναλωτή ένα χρηματικό κίνητρο. Το χρηματικό κίνητρο θα εξαρτάται από τη δομή του προγράμματος απόκρισης ζήτησης. Μπορεί να έχει τη μορφή μία τιμολογιακής πολιτικής με διαφορετικά επίπεδα τιμής μέσα στη μέρα (price based πρόγραμμα) ή να προτείνει μία μείωση στον τελικό λογαριασμό του ρεύματος για κάθε επιπλέον kWh (δηλαδή κάθε kWh πάνω από την πρόβλεψη) που θα κάψει ο καταναλωτής τις ώρες της αντίστροφης ροής φορτίου (incentive based πρόγραμμα). Σε κάθε περίπτωση, το τελευταίο κομμάτι του συστήματος παράγει αυτό το χρηματικό κίνητρο αυτόματα ώστε η ASM να το στείλει στον καταναλωτή.

### 8.2.2 Διάγραμμα του συστήματος

Στην Εικόνα 19 παρουσιάζεται ένα UML διάγραμμα του συστήματος που περιγράφεται παραπάνω. Απεικονίζονται τα 3 βασικά κομμάτια του συστήματος (forecasting service, clustering service, recommendation-pricing engine) και οι βασικές οντότητες που αλληλοεπιδρούν με το σύστημα (μηχανικός ML Ops, μηχανικός ενέργειας και τελικός καταναλωτής). Επίσης συμπεριλαμβάνονται μία βάση MongoDB που περιέχει τις μετρήσεις κατανάλωσης και παραγωγής ενέργειας για τους καταναλωτές που συμμετέχουν στο πρόγραμμα και ένας server που υποστηρίζει τη λειτουργία του MLflow, αποθηκεύοντας τα μοντέλα forecasting και clustering, μετρικές σχετικές με την εκπαίδευση και την επίδοσή τους καθώς και άλλα απαραίτητα στοιχεία (πχ. γραφήματα).



## Εικόνα 19 - UML διάγραμμα του συστήματος απόκρισης ζήτησης

### 8.2.3 Recommendation Engine

Όπως αναφέρεται στο παράδειγμα της ενότητας 8.2.1, το πως θα λειτουργεί το τελευταίο κομμάτι, αυτό που παράγει τα κίνητρα / τιμές που θα κινητοποιήσουν τον καταναλωτή να αλλάξει τη συμπεριφορά του, εξαρτάται από τη δομή του προγράμματος απόκρισης ζήτησης.

Μία ιδέα θα ήταν το πρόγραμμα να ακολουθεί δομή Real Time Pricing, με συγκεκριμένες ζώνες τιμής μέσα στη μέρα και τις τιμές να ανακοινώνονται στους καταναλωτές το προηγούμενο βράδυ. Σε αυτή την περίπτωση, η ASM μπορεί να ζητήσει από ειδικούς να παράγουν μία διαφορετική τιμολογιακή πολιτική για κάθε συστάδα (άρα αν έχουμε 14 συστάδες 14 διαφορετικές τιμολογιακές πολιτικές) και το μόνο που θα χρειάζεται για την εξαγωγή των τιμών του ρεύματος της επόμενης μέρας θα είναι η συστάδα στην οποία ανήκει η πρόγνωση κάθε καταναλωτή.

Οι τιμές θα μπορούσαν να εξαχθούν και με μοντελοποίηση των καταναλωτών με τη χρήση συναρτήσεων ζήτησης (demand functions) και τη διενέργεια προσομοίωσης. Σε μία τέτοια προσομοίωση, είσοδοι θα ήταν οι προγνώσεις της επόμενης ημέρας και οι τιμές του ρεύματος για κάθε καταναλωτή και έξοδοι οι τελικές καταναλώσεις. Από τις τελικές καταναλώσεις μπορούν να μετρηθούν η αντίστροφη ροή φορτίου, η κατανάλωση τις ώρες αιχμής, καθώς και η εξοικονόμηση χρημάτων για κάθε καταναλωτή.

### 8.2.4 Οντότητες που αλληλοεπιδρούν με το σύστημα

**Data Scientist / ML Ops Engineer:** είναι υπεύθυνος για τη συντήρηση και βελτίωση των πρώτων δύο κομματιών του συστήματος, της υπηρεσίας προγνώσεων και της υπηρεσίας συσταδοποίησης. Αυτό περιλαμβάνει την επιτήρηση, αξιολόγηση και επανεκπαίδευση των μοντέλων μηχανικής μάθησης που χρησιμοποιούνται από το σύστημα.

**Μηχανικός Ενέργειας (Energy Engineer):** είναι υπεύθυνος για την επίβλεψη του συστήματος προτάσεων και έχει τη δυνατότητα να μεταβάλλει τον τρόπο λειτουργίας του ανάλογα με τις ανάγκες κάθε περιόδου σε ελαστικότητα και απόκριση ζήτησης. Επιπλέον, μπορεί να εξάγει διάφορα σενάρια αποδοχής των μέτρων / πολιτικών που προτείνει το σύστημα προτάσεων και να ελέγξει το πως κάθε σενάριο επηρεάζει συνολικά το δίκτυο.

**Τελικός χρήστης (End User):** αλληλοεπιδρά με το σύστημα προτάσεων μέσω αποκλειστικής διεπαφής, μέσω της οποίας μπορεί να λάβει τις τιμές / προτάσεις για την επομένη μέρα καθώς και στατιστικά για τα χρήματα που έχει εξοικονομήσει και θα εξοικονομήσει από τη συμμετοχή του στο πρόγραμμα.

## 9. Μελλοντική Εργασία

Το πειραματικό κομμάτι αυτής της εργασίας κατέληξε με την πρόταση κάποιων προγραμμάτων DR για συγκεκριμένες ομάδες καταναλωτών (ενότητα 8.1) και ενός συστήματος που θα χρησιμοποιεί μοντέλα μηχανικής μάθησης και προσομοιώσεις για την παραγωγή προτάσεων DR και τιμολογιακών πολιτικών για κάθε καταναλωτή στο δίκτυο της ASM (ενότητα 8.2).

Κάτι που λείπει είναι ο έλεγχος ότι οι παραπάνω προτάσεις θα οδηγούσαν πράγματι στην επίτευξη των στόχων της ASM (μείωση αντίστροφης ροής φορτίου, μείωση κατανάλωσης τις ώρες αιχμής) και ότι θα ήταν συμφέρουσες και για τους καταναλωτές, πράγμα κρίσιμο για να δεχτούν να συμμετέχουν στην απόκριση ζήτησης. Αυτό είναι ένα απαραίτητο βήμα για να προκριθούν οι επενδύσεις, σε εξοπλισμό και εργατοώρες, που θα χρειάζονταν για την έναρξη ενός προγράμματος.

Μία ιδέα είναι ο έλεγχος των προτάσεων με χρήση προσομοιώσεων, που θα έπαιρναν ως είσοδο πχ. τις τιμές του ρεύματος για κάθε καταναλωτή (που θα προέκυπταν από το σύστημα της ενότητας 8.2 για κάθε καταναλωτή ξεχωριστά ή από τα προτεινόμενα προγράμματα της ενότητας 8.1 για ομάδες καταναλωτών) και θα επέστρεφαν στην έξοδο την αλλαγμένη κατανάλωση κάθε καταναλωτή και τη συνολική κατανάλωση στο δίκτυο.

Μία τέτοια προσομοίωση απαιτεί κάποια μοντελοποίηση της αντίδρασης των καταναλωτών σε διαφορετικά σήματα τιμής. Πολλά άρθρα σε αυτόν τον τομέα χρησιμοποιούν την έννοια της ελαστικότητας της ζήτησης (elasticity of demand) [50]–[52], κατά την οποία η τιμή και η ζήτηση συνδέονται μεταξύ τους με έναν συντελεστή, τον συντελεστή ελαστικότητας. Αυτός ο συντελεστής μπορεί να είναι είτε σταθερός είτε να αλλάζει από καταναλωτή σε καταναλωτή και από ώρα σε ώρα. Η ανάλυση της συμπεριφοράς των καταναλωτών και ο διαχωρισμός σε συστάδες θα μπορούσε να βοηθήσει στον προσδιορισμό αυτών των συντελεστών, πχ. κάθε συστάδα να έχει τους δικούς της συντελεστές ελαστικότητας ανάλογα με το σχήμα του προφίλ της και το είδος των φορτίων (οικιακά, εμπορικά κλπ) που ανήκουν σε αυτή.

Με βάση την παραπάνω μοντελοποίηση των καταναλωτών, μπορούν να δοκιμαστούν και διάφοροι αλγόριθμοι προσδιορισμού της βέλτιστης τιμολογιακής πολιτικής για κάθε καταναλωτή. Μία ιδέα θα ήταν η μοντελοποίηση και επίλυση του προβλήματος με χρήση ενισχυτικής μάθησης (reinforcement learning – RL). Σε αυτή την περίπτωση πράκτορας (agent) θα ήταν ο aggregator και θα προσπαθούσε μέσω της επιλογής διαφορετικών τιμών για κάθε ώρα της ημέρας να βελτιστοποιήσει μία συνάρτηση που θα συνδύαζε τα κόστη της ASM από την αντίστροφη ροή φορτίου και από τη συμφόρηση τις ώρες αιχμής με το κόστος αγοράς του ρεύματος από τους καταναλωτές και την δυσαρέσκεια των καταναλωτών που οφείλεται στην αλλαγή των καταναλωτικών τους συνηθειών. Παρόμοιες εργασίες για αναφορά υπάρχουν στη βιβλιογραφία [53]–[56].



## 10. Κατάλογος Βιβλιογραφίας

- [1] “Demand Response,” *US Department of Energy*. <https://www.energy.gov/oe/demand-response> (accessed Feb. 28, 2023).
- [2] J. S. Vardakas, N. Zorba, and C. v. Verikoukis, “A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms,” *IEEE Communications Surveys and Tutorials*, vol. 17, no. 1, pp. 152–178, Jan. 2015, doi: 10.1109/COMST.2014.2341586.
- [3] X. Yan, Y. Ozturk, Z. Hu, and Y. Song, “A review on price-driven residential demand response,” *Renewable and Sustainable Energy Reviews*, vol. 96. Elsevier Ltd, pp. 411–419, Nov. 01, 2018. doi: 10.1016/j.rser.2018.08.003.
- [4] S. Shao, T. Zhang, M. Pipattanasomporn, and S. Rahman, “Impact of TOU Rates on Distribution Load Shapes in a Smart Grid with PHEV Penetration,” in *IEEE/PES Transmission and Distribution Conference*, 2010.
- [5] S. Gottwalt, W. Ketter, C. Block, J. Collins, and C. Weinhardt, “Demand side management- A simulation of household behavior under variable prices,” *Energy Policy*, vol. 39, no. 12, pp. 8163–8174, Dec. 2011, doi: 10.1016/j.enpol.2011.10.016.
- [6] M. L. Nicolson, M. J. Fell, and G. M. Huebner, “Consumer demand for time of use electricity tariffs: A systematized review of the empirical evidence,” *Renewable and Sustainable Energy Reviews*, vol. 97. Elsevier Ltd, pp. 276–289, Dec. 01, 2018. doi: 10.1016/j.rser.2018.08.040.
- [7] G. R. Newsham and B. G. Bowker, “The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review,” *Energy Policy*, vol. 38, no. 7, pp. 3289–3296, Jul. 2010, doi: 10.1016/j.enpol.2010.01.027.
- [8] M. B. Rosenzweig, H. Fraser, J. Falk, and S. P. Voll, “Market power and demand responsiveness: Letting customers protect themselves,” *Electricity Journal*, vol. 16, no. 4, pp. 11–23, May 2003, doi: 10.1016/S1040-6190(03)00048-4.
- [9] K. Herter, “Residential implementation of critical-peak pricing of electricity,” *Energy Policy*, vol. 35, no. 4, pp. 2121–2130, Apr. 2007, doi: 10.1016/j.enpol.2006.06.019.
- [10] A. Faruqui and S. Sergici, “Household response to dynamic pricing of electricity: A survey of 15 experiments,” *Journal of Regulatory Economics*, vol. 38, no. 2. pp. 193–225, 2010. doi: 10.1007/s11149-010-9127-y.
- [11] H. Allcott, “Real time pricing and electricity markets,” Feb. 2009.



- [12] D. Hammerstrom, R. Ambrosio, and T. Carlon, “Pacific Northwest GridWise Testbed demonstration projects. Part I,” 2007.
- [13] H. Allcott, “Rethinking real-time electricity pricing,” *Resour Energy Econ*, vol. 33, no. 4, pp. 820–842, Nov. 2011, doi: 10.1016/j.reseneeco.2011.06.003.
- [14] L. Lutzenhiser *et al.*, “Behavioral Assumptions Underlying California Residential Sector Energy Efficiency Programs,” 2009.
- [15] F. McLoughlin, A. Duffy, and M. Conlon, “A clustering approach to domestic electricity load profile characterisation using smart metering data,” *Appl Energy*, vol. 141, pp. 190–199, Mar. 2015, doi: 10.1016/j.apenergy.2014.12.039.
- [16] J. Kwac, J. Flora, and R. Rajagopal, “Household energy consumption segmentation using hourly data,” *IEEE Trans Smart Grid*, vol. 5, no. 1, pp. 420–430, Jan. 2014, doi: 10.1109/TSG.2013.2278477.
- [17] B. A. Smith, Wong, and Rajagopal, “A Simple Way to Use Interval Data to Segment Residential Customers for Energy Efficiency and Demand Response Program Targeting Use of Segmentation Schemes within the Utility Industry: A Review,” 2012. [Online]. Available: [http://my.epri.com/portal/server.pt?Abstract\\_id=00000000001016386.CLASSIFYmanualsandreportsareavailableonlineatwww.epri.com](http://my.epri.com/portal/server.pt?Abstract_id=00000000001016386.CLASSIFYmanualsandreportsareavailableonlineatwww.epri.com).
- [18] J. Kwac, J. Flora, and R. Rajagopal, “Household energy consumption segmentation using hourly data,” *IEEE Trans Smart Grid*, vol. 5, no. 1, pp. 420–430, Jan. 2014, doi: 10.1109/TSG.2013.2278477.
- [19] D. Zhou, M. Balandat, and C. Tomlin, “Residential Demand Response Targeting Using Machine Learning with Observational Data,” Jul. 2016, [Online]. Available: <http://arxiv.org/abs/1607.00595>
- [20] C. Si, S. Xu, C. Wan, D. Chen, W. Cui, and J. Zhao, “Electric Load Clustering in Smart Grid: Methodologies, Applications, and Future Trends,” *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 2, pp. 237–252, Mar. 2021, doi: 10.35833/MPCE.2020.000472.
- [21] “Types of Electrical Loads.” <https://powersystemsinternational.com/types-of-electrical-loads/> (accessed Feb. 25, 2023).
- [22] United States Environmental Protection Agency (EPA), “Electricity Customers.” <https://www.epa.gov/energy/electricity-customers> (accessed Feb. 27, 2023).
- [23] U.S. Energy Information Administration, “Homes show greatest seasonal variation in electricity use,” <https://www.eia.gov/todayinenergy/detail.php?id=10211>.

- [24] S. Kiliccote, “Findings from Seven Years of Field Performance Data for Automated Demand Response in Commercial Buildings,” 2010.
- [25] C. Álvarez Bel, M. A. Ortega, G. E. Escrivá, and A. Gabaldón Marín, “Technical and economical tools to assess customer demand response in the commercial sector,” *Energy Convers Manag*, vol. 50, no. 10, pp. 2605–2612, Oct. 2009, doi: 10.1016/j.enconman.2009.06.002.
- [26] I. Khan, “Household factors and electrical peak demand: a review for further assessment,” *Advances in Building Energy Research*, vol. 15, no. 4. Taylor and Francis Ltd., pp. 409–441, 2021. doi: 10.1080/17512549.2019.1575770.
- [27] US Department of Energy, “Benefits of Demand Response in Electricity Markets and Recommendations for Achieving them, report to the United States Congress, Feb. 2006.”.
- [28] J. Kwac, J. Flora, and R. Rajagopal, “Lifestyle Segmentation Based on Energy Consumption Data,” *IEEE Trans Smart Grid*, vol. 9, no. 4, pp. 2409–2418, Jul. 2018, doi: 10.1109/TSG.2016.2611600.
- [29] H.-A. Cao, C. Beckel, and T. Staake, “Are domestic load profiles stable over time? An attempt to identify target households for demand side management campaigns,” *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, 2013.
- [30] N. Bokde, M. W. Beck, F. Martínez Álvarez, and K. Kulat, “A novel imputation methodology for time series based on pattern sequence forecasting,” *Pattern Recognit Lett*, vol. 116, pp. 88–96, Dec. 2018, doi: 10.1016/j.patrec.2018.09.020.
- [31] J. Peppanen, X. Zhang, S. Grijalva, and M. J. Reno, “Handling bad or missing smart meter data through advanced data imputation,” in *2016 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2016*, Dec. 2016. doi: 10.1109/ISGT.2016.7781213.
- [32] I. Assent, “Clustering high dimensional data,” *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 2, no. 4, pp. 340–350, Jul. 2012, doi: 10.1002/widm.1062.
- [33] A. Hinneburg and D. A. Keim, “Optimal Grid-Clustering : Towards Breaking the Curse of Dimensionality in High-Dimensional Clustering,” 1999. [Online]. Available: <http://www.ub.uni-konstanz.de/kops/volltexte/2008/7041/>
- [34] R. Mehra, N. Bhatt, F. Kazi, and N. M. Singh, “Analysis of PCA Based Compression and Denoising of Smart Grid Data under Normal and Fault Conditions.”
- [35] M. Cui, J. Wang, and M. Yue, “Machine Learning-Based Anomaly Detection for Load Forecasting Under Cyberattacks,” *IEEE Trans Smart Grid*, vol. 10, no. 5, pp. 5724–5734, Jan. 2019, doi: 10.1109/tsg.2018.2890809.

- [36] R. Li, F. Li, and N. D. Smith, "Multi-Resolution Load Profile Clustering for Smart Metering Data," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4473–4482, Nov. 2016, doi: 10.1109/TPWRS.2016.2536781.
- [37] S. Ryu, H. Choi, H. Lee, and H. Kim, "Convolutional Autoencoder Based Feature Extraction and Clustering for Customer Load Analysis," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1048–1060, Mar. 2020, doi: 10.1109/TPWRS.2019.2936293.
- [38] R. Bellman and R. Kalaba, "On adaptive control processes," *IRE Transaction on Automatic Control*, vol. 4, no. 2, pp. 1–9, 1959.
- [39] S. Salvador and P. Chan, "FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space."
- [40] H. Sakoe and S. Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition," *IEEE Trans Acoust*, no. 1, p. 43, 1978.
- [41] D. Xu and Y. Tian, "A Comprehensive Survey of Clustering Algorithms," *Annals of Data Science*, vol. 2, no. 2, pp. 165–193, Jun. 2015, doi: 10.1007/s40745-015-0040-1.
- [42] S. K. Bhatia, "Adaptive K-Means Clustering," in *FLAIRS Conference*, 2004. [Online]. Available: [www.aaai.org](http://www.aaai.org)
- [43] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," in *Published in Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining*, 1996.
- [44] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science (1979)*, vol. 344, no. 6191, pp. 1492–1496, 2014, doi: 10.1126/science.1242072.
- [45] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," 1987.
- [46] D. L. Davies and D. W. Bouldin, "A Cluster Separation Measure," *IEEE Trans Pattern Anal Mach Intell*, vol. 1, no. 2, pp. 224–227, 1979, doi: <https://doi.org/10.1109/2FTPAMI.1979.4766909>.
- [47] S. Lin, F. Li, E. Tian, Y. Fu, and D. Li, "Clustering load profiles for demand response applications," *IEEE Trans Smart Grid*, vol. 10, no. 2, pp. 1599–1607, Mar. 2019, doi: 10.1109/TSG.2017.2773573.
- [48] M. Darsun, M. A. Alam, A. S. Masud, and M. S. Islam, "Seasonal and diurnal variation of electrical energy consumption in residential sector," in *2016 International Conference on Electrical, Electronics, Engineering Trends, Communication, Optimization and Sciences (EEECOS)*, 2016, pp. 1–6.

- [49] Andrew Kirmse, “Topographic Prominence,” <https://www.andrewkirmse.com/prominence>.  
<https://www.andrewkirmse.com/prominence> (accessed Feb. 22, 2023).
- [50] R. Lu, S. H. Hong, and X. Zhang, “A Dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach,” *Appl Energy*, vol. 220, pp. 220–230, Jun. 2018, doi: 10.1016/j.apenergy.2018.03.072.
- [51] M. G. Lijesen, “The real-time price elasticity of electricity,” *Energy Econ*, vol. 29, no. 2, pp. 249–258, Mar. 2007, doi: 10.1016/j.eneco.2006.08.008.
- [52] M. Miller and A. Alberini, “Sensitivity of price elasticity of demand to aggregation, unobserved heterogeneity, price trends, and price endogeneity: Evidence from U.S. Data,” *Energy Policy*, vol. 97, pp. 235–249, Oct. 2016, doi: 10.1016/j.enpol.2016.07.031.
- [53] R. Lu, S. H. Hong, X. Zhang, X. Ye, and W. S. Song, “A Perspective on Reinforcement Learning in Price-Based Demand Response for Smart Grid,” in *Proceedings - 2017 International Conference on Computational Science and Computational Intelligence, CSCI 2017*, Dec. 2018, pp. 1822–1823. doi: 10.1109/CSCI.2017.327.
- [54] R. Lu, S. H. Hong, and X. Zhang, “A Dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach,” *Appl Energy*, vol. 220, pp. 220–230, Jun. 2018, doi: 10.1016/j.apenergy.2018.03.072.
- [55] B. G. Kim, Y. Zhang, M. Van Der Schaar, and J. W. Lee, “Dynamic pricing and energy consumption scheduling with reinforcement learning,” *IEEE Trans Smart Grid*, vol. 7, no. 5, pp. 2187–2198, Sep. 2016, doi: 10.1109/TSG.2015.2495145.
- [56] A. Ghasemkhani and L. Yang, “Reinforcement learning based pricing for demand response,” in *2018 IEEE International Conference on Communications Workshops, ICC Workshops 2018 - Proceedings*, Jul. 2018, pp. 1–6. doi: 10.1109/ICCW.2018.8403783.

--- Τέλος εγγράφου ---