



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΔΙΑΔΙΚΤΥΑΚΗ ΑΛΓΟΡΙΘΜΙΚΗ ΓΙΑ ΤΗ ΔΙΑΧΕΙΡΙΣΗ ΕΝΕΡΓΕΙΑΚΩΝ ΔΕΔΟΜΕΝΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Γεώργιος, Π. Καλογεράς

Υπεύθυνος καθηγητής: Λούμος Βασίλειος
Καθηγητής Τομέα Επικοινωνιών, Ηλεκτρονικής και
Συστημάτων Πληροφορικής ΕΜΠ

Επιβλέπων: Νικολόπουλος Βασίλειος
Υποψήφιος Διδάκτορας ΕΜΠ

Αθήνα, Απρίλιος 2010



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΔΙΑΔΙΚΤΥΑΚΗ ΑΛΓΟΡΙΘΜΙΚΗ ΓΙΑ ΤΗ ΔΙΑΧΕΙΡΙΣΗ ΕΝΕΡΓΕΙΑΚΩΝ ΔΕΔΟΜΕΝΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Γεώργιος, Π. Καλογεράς

Υπεύθυνος καθηγητής: Λούμος Βασίλειος

Καθηγητής Τομέα Επικοινωνιών, Ηλεκτρονικής και
Συστημάτων Πληροφορικής ΕΜΠ

Επιβλέπων: Νικολόπουλος Βασίλειος

Υποψήφιος Διδάκτορας ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 30^η Απριλίου 2010.

.....

Λούμος Βασίλειος

Καθηγητής ΕΜΠ

Αθήνα, Απρίλιος 2010

.....
Γεώργιος, Π. Καλογεράς

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

Copyright © Γεώργιος, Π. Καλογεράς , 2010

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.

ΠΕΡΙΛΗΨΗ

Οι παγκόσμιες συνθήκες που έχουν διαμορφωθεί στην αγορά ηλεκτρικής ενέργειας μέσω πληθώρας παραγόντων όπως η απελευθέρωση της αγοράς, η οικονομική κρίση και η κλιματική αλλαγή που έχει οδηγήσει σε στροφή στις εναλλακτικές πηγές ενέργειας, απαιτούν και την εξέλιξη των ενεργειακών συστημάτων σε πιο ευέλικτες μορφές.

Αντικείμενο της διπλωματικής είναι η ανάλυση μαζικών ενεργειακών δεδομένων καταναλωτών και η εξόρυξη σημαντικών πληροφοριών από αυτά. Τα δεδομένα λαμβάνονται από αυτόματους μετρητές και είναι διαθέσιμα προς ανάλυση καθώς αποθηκεύονται σε βάσεις δεδομένων. Η ανάλυσή τους έχει ως στόχο την Ανίχνευση Αιχμών και τη συσχέτισή τους για τον ίδιο ή για διαφορετικούς καταναλωτές. Η ανάλυση αυτή είναι εξαιρετικά σημαντική καθώς οι αιχμές είναι ο κυρίαρχος παράγοντας κινδύνου για το σύστημα στις ώρες αιχμής. Ακόμα συσχετίζονται ποιοτικά οι πελάτες με βάση κάποιους δείκτες ποιοτικής ανάλυσης οι οποίοι καταρτίζονται από εμάς. Οι δείκτες αυτοί βοηθούν στη συσχέτιση ενεργειακών προφίλ και άρα στην κατηγοριοποίηση πελατών με βάση την κατανάλωσή τους με σκοπό την εφαρμογή μιας κοινής τιμολογιακής πολιτικής. Τέλος γίνεται ανάλυση περιοδικότητας, και παρέχονται δυνατότητες εκτίμησης μελλοντικής συμπεριφοράς με βάση την ανάλυση στοιχείων του παρελθόντος σε ενεργειακά συνεπείς καταναλωτές. Η εργασία έγινε στα πλαίσια του ερευνητικού προγράμματος plug-me που έχει ως αντικείμενο την Ανάλυση και Ιεραρχικοποίηση Ενεργειακών Πελατών κάνοντας χρήση του Διαδικτύου και ενός προηγμένου Ενεργειακού Πληροφοριακού Συστήματος. Οι μετρήσεις που έχουμε διαθέσιμες προέρχονται από τα γραφεία του Ιδρύματος Intelen στην Καισαριανή.

Λέξεις κλειδιά:

Ενεργειακό Πληροφοριακό Σύστημα, Ενεργειακό Προφίλ, Ανίχνευση Αιχμών, Διαχείριση Ενέργειας, Διαχείριση Ζήτησης, Εξόρυξη Δεδομένων, Ποιοτική Ανάλυση Προφίλ, Ομαδοποίηση Πελατών, Ανάλυση Περιοδικότητας.

ABSTRACT

The newly formed global conditions in the electrical energy market due to a series of factors such as the market liberation, the economic crisis and the climatic change that has led to an increased interest in alternative “green” energy resources, have led to the need for evolving energy systems to more modern and up to date forms.

The subject of the present Diploma Thesis is the analysis of energy data gathered from consumers and the mining of useful information from it. The data is gathered by automatic meters and stored in databases in order to be readily available for analysis. The analysis performs Peak Detection and correlation of peaks for the same or different consumers. This is extremely important for the energy system as peaks are the main danger factor for the system during peak hours. Moreover a qualitative correlation of clients is performed using a set of quality indices. These indices help us correlate energy profiles and categorize consumers according to their consumption patterns in order to apply common pricing to each group. Finally a periodicity check is performed and the basis for estimating future behavior is set up by analyzing past consumptions of consistent customers. This project is part of the research program “plug-me” ,whose objective is to analyze and rank Energy Customers and Consumptions over Internet using adaptive search machine technology in an web-based intelligent Energy Information System.

Keywords:

Energy Information System, Energy Profile, Peak Detection, Energy Management, Demand Management, Data Mining, Qualitative Profile Analysis, Customer Grouping, Periodicity Analysis.

ΔΟΜΗ ΔΙΠΛΩΜΑΤΙΚΗΣ

Στο Πρώτο Κεφάλαιο επεξηγούνται βασικές έννοιες όπως τι είναι ένα ενεργειακό πληροφοριακό σύστημα και τι είναι η διαχείριση ενέργειας. Ακόμα αναλύεται η ανάγκη για την ανάπτυξη πληροφοριακών συστημάτων νέας γενιάς και επεξηγείται ο όρος ενεργειακό προφίλ.

Στο Δεύτερο Κεφάλαιο γίνεται αναφορά στους έξυπνους μετρητές και εξηγείται το πώς λειτουργεί ένα αυτοματοποιημένο σύστημα μετρήσεων. Αναλύεται ακόμα η προτυποποιημένη διαδικασία θερμοδυναμικής ανάλυσης και ενεργειακής μέτρησης κτιρίων.

Το Κεφάλαιο Τρία επεξηγεί τον όρο διαχείριση ζήτησης ενέργειας και περιγράφει τη δομή και λειτουργία ενός ευφυούς ενεργειακού συστήματος νέας γενιάς.

Στο Τέταρτο Κεφάλαιο αναλύονται στοιχεία μαθηματικών από τη θεωρία ιδιοτιμών και περιγράφεται η διαδικασία που χρησιμοποιείται για την ανάλυση των δεδομένων (Principal Component Analysis).

Στο Πέμπτο Κεφάλαιο περιγράφεται πλήρως ο αλγόριθμος που χρησιμοποιείται και αναλύεται σε βάθος η διαδικασία επιλογής και χρήσης των KPIs.

Στο Έκτο Κεφάλαιο παρουσιάζονται συγκεκριμένα παραδείγματα (case studies) που αποδεικνύουν τη χρησιμότητα της μεθόδου.

Τέλος στο Έβδομο Κεφάλαιο συνοψίζονται τα συμπεράσματα και αναφέρονται μελλοντικές χρήσεις της εργασίας.

Στο Παράρτημα Α παρατίθενται οι συναρτήσεις που υλοποιούν τον αλγόριθμο σε PHP και ο κώδικας MATLAB για την ανάλυση πινάκων.

ΠΕΡΙΕΧΟΜΕΝΑ

1	ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ	10
1.1	ΕΙΣΑΓΩΓΗ	10
1.2	ΔΙΑΧΕΙΡΙΣΗ ΕΝΕΡΓΕΙΑΣ	11
1.3	ΕΝΕΡΓΕΙΑΚΟ ΠΛΗΡΟΦΟΡΙΑΚΟ ΣΥΣΤΗΜΑ	11
1.4	ΕΝΕΡΓΕΙΑΚΟ ΠΡΟΦΙΛ	12
2	AMR SYSTEMS – ΕΝΕΡΓΕΙΑΚΗ ΜΕΤΡΗΣΗ ΚΑΙ ΘΕΡΜΟΔΥΝΑΜΙΚΗ ΑΝΑΛΥΣΗ ΚΤΙΡΙΟΥ	14
2.1	SMART METERS – ΈΞΥΠΝΟΙ ΜΕΤΡΗΤΕΣ	14
2.2	ΣΥΣΤΗΜΑΤΑ AMR	14
2.3	ΕΝΕΡΓΕΙΑΚΗ ΜΕΤΡΗΣΗ ΚΑΙ ΘΕΡΜΟΔΥΝΑΜΙΚΗ ΑΝΑΛΥΣΗ ΚΤΙΡΙΟΥ (ASRAE)	16
3	ΔΙΑΧΕΙΡΙΣΗ ΖΗΤΗΣΗΣ – SMART GRID	21
3.1	ΔΙΑΧΕΙΡΙΣΗ ΖΗΤΗΣΗΣ ΕΝΕΡΓΕΙΑΣ	21
3.2	SMART GRID	23
4	ΜΑΘΗΜΑΤΙΚΗ ΠΕΡΙΓΡΑΦΗ	27
4.1	ΙΔΙΟΤΗΤΕΣ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ ΤΕΤΡΑΓΩΝΙΚΟΥ ΠΙΝΑΚΑ	27
4.2	PRINCIPAL COMPONENTS ANALYSIS (PCA)	28
5	ΑΛΓΟΡΙΘΜΟΣ	32
5.1	ΕΠΙΛΟΓΗ ΠΟΙΟΤΙΚΩΝ ΔΕΙΚΤΩΝ (KPIs)	32
5.2	ΑΝΑΛΥΣΗ ΠΙΝΑΚΩΝ	33
5.3	ΑΛΓΟΡΙΘΜΟΣ	34
5.3.1	Μαθηματική ανάλυση αλγορίθμου	34
5.3.2	Ανάλυση πινάκων στο περιβάλλον MATLAB	43
6	TEST CASES	46
6.1	TEST CASE 1	46
6.2	TEST CASE 2	58
7	ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ ΜΕΛΕΤΗΣ	68
8	ΑΝΑΦΟΡΕΣ	70
9	ΠΑΡΑΡΤΗΜΑ	73
9.1	ΚΑΘΟΡΙΣΜΟΣ ΜΕΤΑΒΛΗΤΩΝ ΕΙΣΟΔΟΥ	73
9.2	ΠΕΡΙΓΡΑΦΗ ΣΥΝΑΡΤΗΣΕΩΝ	73
9.3	ΚΩΔΙΚΑΣ PHP	74
9.4	ΚΩΔΙΚΑΣ MATLAB	93

1

Εισαγωγή

1 Βασικές έννοιες

1.1 Εισαγωγή

Η πρόσφατη απελευθέρωση της ενεργειακής αγοράς στη χώρα μας έχει δημιουργήσει καινούρια δεδομένα για τις εταιρείες παροχής ενέργειας. Οι νέες συνθήκες ανταγωνισμού, τα δεδομένα που προκύπτουν από την παγκόσμια οικονομική κρίση καθώς και η αυξημένη ανάγκη για προστασία του περιβάλλοντος, εν όψει της επερχόμενης κλιματικής αλλαγής, έχουν καταστήσει σαφή την αναγκαιότητα για σωστή διαχείριση ζήτησης ενέργειας. Ο όρος αυτός περιλαμβάνει τη σωστή πρόβλεψη ζήτησης και τη σωστή διαχείριση παραγωγής με βάση τις ιδιαίτερες ανάγκες των πελατών. Εφόσον ο στόχος είναι η μείωση του κόστους παραγωγής από την πλευρά της εταιρείας και η εξοικονόμηση χρημάτων από την πλευρά των καταναλωτών, μια σωστή πολιτική τιμολόγησης μπορεί να είναι αρκετή για να καλύψει τις ανάγκες και των δυο πλευρών.

Το κόστος για την εταιρεία αφορά στο μεγαλύτερο ποσοστό του το κόστος παραγωγής της ενέργειας. Αυτό το κόστος παραγωγής αυξάνεται εκθετικά όταν το επίπεδο ζήτησης ξεπερνά ένα κρίσιμο κατώφλι τις ώρες αιχμής. Όταν ξεπεραστεί το κατώφλι αυτό η εταιρεία προκειμένου να μπορέσει να ικανοποιήσει την αυξημένη ζήτηση καλείται να θέσει σε λειτουργία τις μονάδες εφεδρείας. Αυτές είναι πολυέξοδες μονάδες παραγωγής οι οποίες χρησιμοποιούνται μόνο σε τέτοιες περιπτώσεις ανάγκης. Είναι λογικό λοιπόν η οποιαδήποτε εταιρεία να προσπαθεί να μειώσει τη χρήση των μονάδων αυτών. Αυτό μπορεί να γίνει μόνο εάν υπάρχει ένα σύστημα με το οποίο μπορεί η εταιρεία να είναι σίγουρη ότι ποτέ δε θα ξεπεραστεί το κρίσιμο επίπεδο ζήτησης.

Καθώς όμως η ζήτηση εξαρτάται από τους πελάτες είναι απαραίτητο η εταιρεία να μπορεί να ελέγχει κατά κάποιο τρόπο την ενεργειακή συμπεριφορά τους, ειδικά στις ώρες αιχμής. Αυτό μπορεί να γίνει μετακινώντας τα ενεργειακά προφίλ των πελατών, με ιδιαίτερη έμφαση στις αιχμές τους, προσφέροντάς τους ανταποδοτικά οφέλη με τη μορφή ειδικών τιμολογίων.

Με βάση τις αρχές αυτές και για να γίνει το σύστημα αποδοτικό πρέπει η τιμολόγηση να είναι όσο δυνατόν καλύτερα εξατομικευμένη σε σχέση με την κάθε κατηγορία πελατών (οικιακοί, εταιρίες, βιομηχανίες κτλ) και να είναι ευέλικτη σε αλλαγές στη συμπεριφορά τους. Έτσι χρειάζεται ένα αυτόματο σύστημα κατηγοριοποίησης πελατών το οποίο στηριζόμενο σε πραγματικές μετρήσεις σχηματίζει το χαρακτηριστικό προφίλ του κάθε πελάτη και με βάση αυτό τον κατατάσσει στην ανάλογη κατηγορία. Με τη διαδικασία αυτή ο κάθε πελάτης έχει μια ιδιωτική τιμολογιακή πολιτική, βασισμένη στις δικές του ανάγκες και τρόπο ζωής, στην οποία μπορεί να αντιδρά αναλόγως προκειμένου να εκμεταλλεύεται πλήρως τα οφέλη που του προσφέρονται. Όλα αυτά επιτυγχάνονται με τη δημιουργία ενός ενεργειακού πληροφοριακού συστήματος.

1.2 Διαχείριση ενέργειας

Ο όρος διαχείριση ενέργειας αναφέρεται στη βέλτιστη διαχείριση του ενεργειακού συστήματος καθολικά, από την παραγωγή μέχρι και την κατανάλωση, με τελικό στόχο τη μέγιστη απόδοση.

Το βασικότερο ζήτημα είναι φυσικά η μείωση των απωλειών ενέργειας. Η κάθε εταιρεία παροχής επιθυμεί να διατηρεί την παραγωγή της όσο χαμηλότερα γίνεται χωρίς φυσικά να υπάρχει κίνδυνος κατάρρευσης του συστήματος. Επιθυμούμε δηλαδή τις ώρες αιχμής να έχουμε ελαχιστοποίηση των απαιτήσεων ζήτησης. Έτσι επιτυγχάνεται η μέγιστη απόδοση και η ελάχιστη απώλεια ενέργειας. Για να γίνει αυτό πρέπει να υπάρχει ολοκληρωμένη εποπτεία ολόκληρου του συστήματος ενέργειας από τα αρχικά στάδια παραγωγής-μεταφοράς (συστήματα SCADA, EMS) μέχρι και τις κτιριακές εγκαταστάσεις οι οποίες καταναλώνουν την ενέργεια.

Η σωστή διαχείριση οδηγεί σε μείωση του κόστους του ενεργειακού συστήματος αλλά έχει και άλλα παράπλευρα οφέλη όπως η μείωση των εκπομπών CO_2 , και η βελτίωση ποιότητας ζωής με τη βέλτιστη διαχείριση των ενεργειακών αναγκών των κτιρίων.

1.3 Ενεργειακό πληροφοριακό σύστημα

Με τον όρο ενεργειακό πληροφοριακό σύστημα περιγράφεται ένα πλήρες σύστημα εποπτείας και οργάνωσης της ροής της ενέργειας από τα σημεία παραγωγής της, καθόλη τη διάρκεια της μεταφοράς και διανομής και φυσικά στα πολλαπλά σημεία κατανάλωσης. Ολοκληρωμένα συστήματα τέτοιου τύπου έχουν αναπτυχθεί την τελευταία δεκαετία και χρησιμοποιούνται σήμερα ευρέως για τη σωστή διαχείριση του ενεργειακού δικτύου. Περιλαμβάνουν βάσεις δεδομένων με περιγραφικές πληροφορίες που αφορούν την ενέργεια (οργανισμοί, νομοθεσία, πρότυπα) καθώς και βάσεις που διατηρούν στατιστικές πληροφορίες που αφορούν την παραγωγή, μετατροπή και κατανάλωση της ενέργειας. Τα στατιστικά αυτά δεδομένα μέσω ενός δικτύου επεξεργασίας προσφέρουν ένα απλοποιημένο εργαλείο ενεργειακής ανάλυσης, το οποίο προσομοιώνει το σύστημα ως ροή ενεργειακών προϊόντων μέσα από ενεργειακές διαδικασίες. Έτσι λαμβάνουμε χρήσιμα ποσοτικά στοιχεία για το ενεργειακό μας σύστημα. Ακόμα είναι απαραίτητο και ένα γεωγραφικό σύστημα πληροφοριών που να διαχειρίζεται τις χωρικές πληροφορίες του συστήματος (χάρτες κοιτασμάτων, ενεργειακών δικτύων και εγκαταστάσεων). Ένα παράδειγμα ενός τέτοιου συστήματος είναι το Εθνικό Πληροφοριακό Σύστημα για την Ενέργεια (Ε.Π.Σ.Ε.) το οποίο αναπτύχθηκε στα πλαίσια του Επιχειρησιακού Προγράμματος Ενέργειας του Δεύτερου Κοινοτικού Πλαισίου Στήριξης.

Η σύνδεση όλων των παραπάνω στοιχείων του συστήματος γίνεται μέσω διαδικτύου και ολοκληρώνεται με την προσθήκη μετρητών στα σημεία κατανάλωσης της ενέργειας, οι οποίοι προσφέρουν πληροφορίες πραγματικού χρόνου για την

κατανάλωση. Αυτή η άμεση ενημέρωση προσφέρει τη δυνατότητα σωστής διαχείρισης της κατανάλωσης κτιρίων. Ο χειριστής έχει ανά πάσα στιγμή πλήρη εικόνα του κτιρίου και μπορεί άμεσα να κάνει τις απαραίτητες ενέργειες ώστε να φέρει την απόδοση του κτιρίου σε επιθυμητά επίπεδα, μειώνοντας τα λειτουργικά κόστη. Παράλληλα στο ζήτημα της τιμολόγησης όπως έχει παρουσιαστεί πιο πάνω ένα πληροφοριακό σύστημα πραγματικού χρόνου όπως αυτό που περιγράφεται καταργεί την ανάγκη για λογαριασμούς έναντι καθώς υπάρχει μέγιστη ακρίβεια μέτρησης της καταναλισκόμενης ενέργειας.

1.4 Ενεργειακό προφίλ

Όλα τα δεδομένα που συλλέγονται και επεξεργάζονται από το πληροφοριακό σύστημα μπορούν να βοηθήσουν στη δημιουργία ενεργειακών προφίλ πελατών. Μέσα από το πλήθος μετρήσεων και στατιστικών μπορεί μια διαδικασία εξόρυξης δεδομένων να διαχωρίσει κάποια ιδιαιτέρως σημαντικά στοιχεία ενδεικτικά της συμπεριφοράς ενός ενεργειακού καταναλωτή, όπως περιοδικότητα, ελαστικότητα και αιχμές.

Τα ενεργειακά προφίλ αποτελούν δείγμα της ενεργειακής συμπεριφοράς των πελατών και μπορούν να χρησιμοποιηθούν για να συσταδοποιηθούν πελάτες σε ομάδες κοινής συμπεριφοράς. Ακόμη μπορούν τα προφίλ αυτά να συγκρίνονται με πρότυπα προφίλ κατανάλωσης, με στόχο τη σταδιακή σύγκλιση ενεργοβόρων καταναλωτών σε πιο λογικά επίπεδα κατανάλωσης.

Η δημιουργία προφίλ όπως αναφέρθηκε βασίζεται στη εξόρυξη δεδομένων (data mining). Ο όρος αυτός αναφέρεται στη διαδικασία μελέτης μεγάλου όγκου δεδομένων με στόχο την επιλεκτική εξαγωγή πληροφοριών μετά από ανάλυση. Η ανάλυση αυτή γίνεται με στοχευμένους έξυπνους αλγόριθμους, οι οποίοι συσχετίζουν τα δεδομένα και βρίσκουν ομοιότητες και διαφορές παρουσιάζοντάς τες τελικά σε μορφή εύκολα κατανοητή από το χρήστη.

2

AMR Systems Ενεργειακή μέτρηση και Θερμοδυναμική ανάλυση κτιρίου

2 AMR Systems – Ενεργειακή μέτρηση και θερμοδυναμική ανάλυση κτιρίου

2.1 Smart meters – Έξυπνοι Μετρητές

Με τον όρο έξυπνοι μετρητές περιγράφουμε συσκευές μέτρησης ηλεκτρικής (και όχι μόνο) ενέργειας με μεγαλύτερη ακρίβεια από τους συμβατικούς μετρητές. Επίσης συγκεντρώνονται περισσότερες πληροφορίες για το δίκτυο από μια απλή μέτρηση της ενέργειας που καταναλώνεται. Μερικά στοιχεία που συλλέγονται είναι κάποιοι δείκτες απόδοσης του συστήματος με κυριότερο το συντελεστή ισχύος.

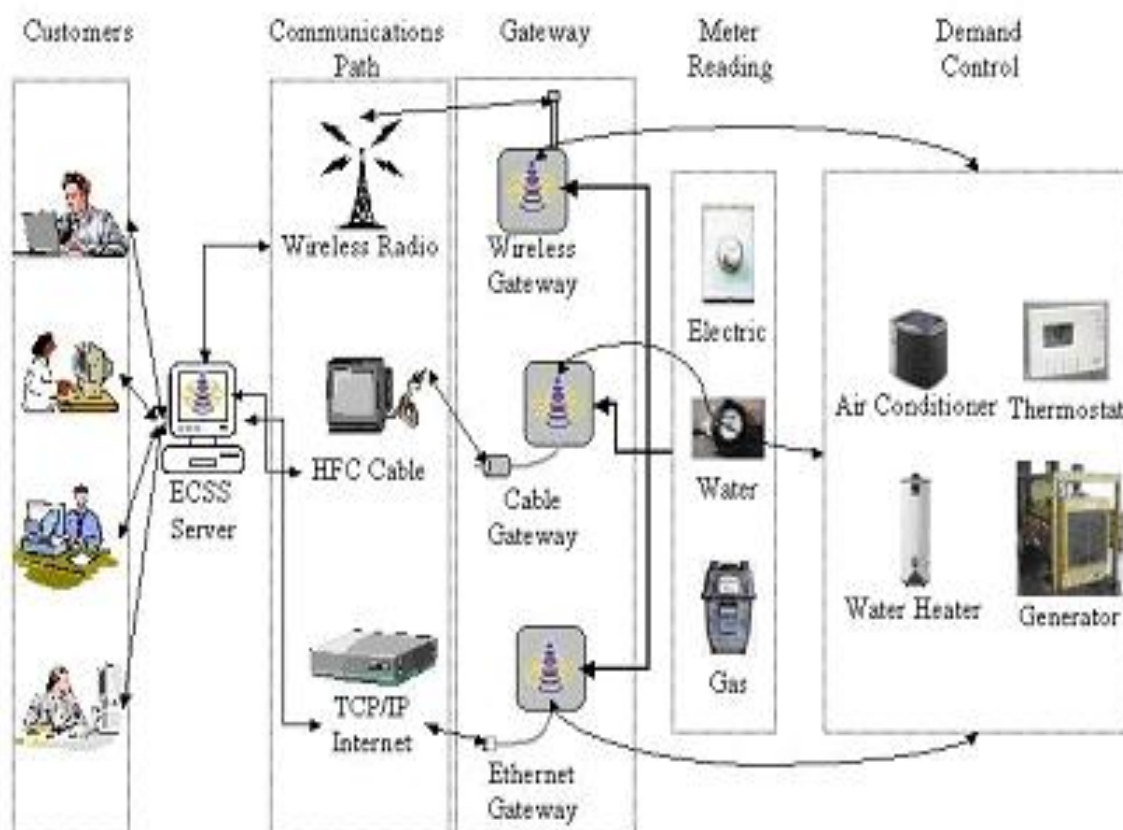
Ένα πλήρες δίκτυο έξυπνων μετρητών εγκατεστημένων σε πλήθος καταναλωτών αποτελεί ένα σημαντικό κομμάτι ενός αυτόματου ευφυούς δικτύου, δίνοντας πληροφορίες πραγματικού χρόνου για την κατανάλωση και παρέχοντας ακριβείς μετρήσεις του ποσού της ενέργειας που καταναλώνεται στο σύνολο του ενεργειακού δικτύου.

Η ύπαρξη του δικτύου μετρητών προϋποθέτει και την ύπαρξη συστήματος επικοινωνίας με τον κεντρικό διαχειριστή. Αυτό το δίκτυο επικοινωνίας είναι ίσως το πιο σημαντικό κομμάτι καθώς οι μετρήσεις από μόνες τους τοπικά δεν προσφέρουν απολύτως τίποτα χωρίς περαιτέρω ανάλυση. Από το πλήθος των τεχνολογιών που χρησιμοποιούνται, όπως Wi-fi και μέσω του δικτύου διανομής ενέργειας, κυρίαρχο ρόλο παίζουν οι εξελιγμένες τεχνολογίες διαδικτύου (TCP/IP).

2.2 Συστήματα AMR

Ο όρος AMR (Automatic Meter Reading) περιγράφει την τεχνολογία της αυτόματης συλλογής δεδομένων κατανάλωσης, κατάστασης και διάγνωσης συστημάτων παροχής ενέργειας (είτε ηλεκτρικής είτε φυσικού αερίου και νερού) και η μεταφορά των δεδομένων αυτών σε βάσεις δεδομένων για αποθήκευση και ανάλυση. Έτσι καταργείται η παλιά τακτική που υποχρέωνε τους παρόχους ενέργειας να επισκέπτονται τακτικά όλους τους πελάτες για να πάρουν μετρήσεις για την κατανάλωση. Ένα άλλο πλεονέκτημα είναι η πραγματικού χρόνου χρέωση των πελατών, κάτι που επιτρέπει σε παρόχους και πελάτες παράλληλα μεγαλύτερο έλεγχο της κατανάλωσης και του κόστους τους.

Υπάρχει πλήθος τακτικών που έχουν χρησιμοποιηθεί για την επικοινωνία των AMR συστημάτων. Κατά καιρούς τα συλλεγόμενα δεδομένα έχουν μεταφερθεί μέσω ραδιοκυμάτων σε δεσμευμένες συχνότητες, είτε με φορητές συσκευές οι οποίες διαβάζουν μετρητές από κάποια απόσταση. Ακόμα κάποιες εταιρείες έχουν αναπτύξει πλήρη αυτόνομα δίκτυα μεταφοράς πληροφοριών με κεραιές και σταθμούς σε διάφορες περιοχές. Παρόλα αυτά τα ανεπτυγμένα συστήματα της τελευταίας δεκαετίας χρησιμοποιούν το διαδίκτυο για να μεταφέρουν τα δεδομένα τους στο κεντρικό σύστημα διαχείρισης όπου θα γίνει η επεξεργασία τους.



Τα συστήματα αυτά προσφέρουν προνόμια σε εταιρίες παραγωγής, παροχής αλλά και σε πελάτες. Οι εταιρίες παραγωγής απολαμβάνουν αυξημένες αποδόσεις, άμεσο εντοπισμό απωλειών και διαρροών καθώς και μειωμένο κόστος για εργατικό δυναμικό εφόσον οι μετρήσεις και οι συνδέσεις – αποσυνδέσεις πελατών γίνονται αυτόματα. Οι εταιρείες παροχής μπορούν να προσφέρουν νέες βελτιωμένες υπηρεσίες βασισμένες στην ανάλυση των δεδομένων ενώ παράλληλα μπορούν να έχουν ευέλικτες πολιτικές τιμολόγησης ανάλογα με τον κάθε πελάτη, κάτι το οποίο ευνοεί και τους πελάτες.

Τα πλεονεκτήματα ενός τέτοιου συστήματος είναι τα εξής:

- Ακριβείς μετρήσεις χωρίς προβλέψεις
- Βελτιωμένη τιμολόγηση
- Σαφείς ενεργειακές κλάσεις πελατών
- Πραγματικό κόστος ενέργειας στον πελάτη
- Βελτιωμένη διαχείριση ζήτησης
- Μειωμένο κόστος μετρήσεων και διορθώσεων λογαριασμών
- Διαφάνεια χρεώσεων

Η λειτουργία ενός AMR είναι η εξής: οι αυτόματοι μετρητές παίρνουν μετρήσεις ενέργειας ανά τακτά χρονικά διαστήματα και σε πραγματικό χρόνο, μέσω της οποιασδήποτε διόδου επικοινωνίας και αν χρησιμοποιείται, αποστέλλει τα δεδομένα κατανάλωσης, μαζί με κάποια άλλα στοιχεία που αφορούν την κατάσταση του συστήματος, στον κεντρικό διακομιστή. Ο διακομιστής δέχεται τις πληροφορίες και τις αποθηκεύει σε βάσεις δεδομένων όπου και φυλάσσονται μέχρι να χρειαστεί να γίνει επεξεργασία τους.

Για τη διαφάνεια των μετρήσεων οι βάσεις αυτές δεδομένων είναι ανοιχτές προς τον ενδιαφερόμενο πελάτη, ο οποίος μπορεί να δει ανά πάσα στιγμή την κατανάλωσή του. Συσσωρεύοντας μετρήσεις για τον κάθε πελάτη σχηματίζεται σιγά σιγά ένα ιστορικό κατανάλωσης, το οποίο μπορεί να χρησιμοποιηθεί για τη δημιουργία του χαρακτηριστικού προφίλ του πελάτη, όπως έχει περιγραφεί πιο πάνω.

Εκτός από το προφίλ του πελάτη μπορούμε με χρήση των μετρήσεων που έχουμε πάρει να δημιουργήσουμε μια σχέση μεταξύ της κατανάλωσης ενός κτιρίου και εξωτερικών παραγόντων όπως η θερμοκρασία. Η σχέση αυτή είναι πολύ σημαντική για την πρόβλεψη της συμπεριφοράς ενός κτιρίου, καθώς έχει αποδειχθεί από μελέτες ότι η θερμοκρασία είναι ένας σημαντικός παράγοντας που επηρεάζει την κατανάλωση ενέργειας σε όλα τα κτίρια. Γιαυτό και καταρτίζονται ειδικά διαγράμματα ενέργειας(KWh) – θερμοκρασίας(C).

2.3 Ενεργειακή μέτρηση και θερμοδυναμική ανάλυση κτιρίου (ASRAE)

Η τακτική της θερμοδυναμικής ανάλυσης και ενεργειακής μέτρησης ενός κτιρίου εφαρμόζεται τα τελευταία χρόνια με στόχο τον προσδιορισμό της απόδοσής του. Η απόδοση του κτιρίου μπορεί να βελτιωθεί με διάφορα μέτρα, κάτι το οποίο είναι και ο απώτερος στόχος. Συγκεκριμένα με ειδικές βελτιώσεις μπορεί να μειωθεί η κατανάλωση κατά 5 – 10%. Αυτή η σημαντική βελτίωση στην απόδοση έχει οδηγήσει σε νομοθετικές ρυθμίσεις οι οποίες προωθούν τη λήψη μέτρων προς την κατεύθυνση αυτή.

Συγκεκριμένα υπάρχει κοινοτική νομοθεσία (EC/2002/91) η οποία θεσπίζει την έκδοση δελτίου ενεργειακής ταυτότητας για κάθε κτίριο, το οποίο είναι ένα πιστοποιητικό που περιγράφει την ενεργειακή απόδοση και συμπεριφορά του κτιρίου. Μετράται η ενεργειακή κατανάλωση και κυρίως η ενέργεια που καταναλώνεται σε θέρμανση και ψύξη του κτιρίου, καθώς έχει αποδειχθεί από μελέτες πως αυτή είναι το μεγαλύτερο κλάσμα της καταναλισκόμενης ενέργειας. Επίσης συνδέεται άμεσα με τις εξωτερικές συνθήκες, επομένως καταρτίζεται μια σχέση κατανάλωσης – θερμοκρασίας σαν μέτρο απόδοσης του κτιρίου. Όλα αυτά περιγράφονται στο δελτίο ενεργειακής ταυτότητας, το οποίο φυσικά και επηρεάζει την τιμή του ακινήτου σε

διάφορες αγοραπωλησίες, και το οποίο θα είναι σε λίγα χρόνια υποχρεωτικό για κάθε νεόδμητο κτίριο.

Η μέθοδος υπολογισμού της ενεργειακής απόδοσης των κτιρίων πρέπει τουλάχιστον να περιλαμβάνει τους ακόλουθους παράγοντες:

- α)Θερμικά χαρακτηριστικά του κτιρίου (κέλυφος και εσωτερικά χωρίσματα, κ.λπ.). Τα χαρακτηριστικά αυτά μπορούν να περιλαμβάνουν και την αεροστεγανότητα.
- β)Εγκατάσταση θέρμανσης και τροφοδοσία θερμού νερού, συμπεριλαμβανομένων των χαρακτηριστικών των μονώσεων τους.
- γ)Εγκατάσταση κλιματισμού.
- δ)Αερισμό.
- ε)Ενσωματωμένη εγκατάσταση φωτισμού (κυρίως στον τομέα που δεν αφορά την κατοικία).
- στ)Θέση και προσανατολισμό των κτιρίων, περιλαμβανομένων των εξωτερικών κλιματικών συνθηκών.
- ζ)Παθητικά ηλιακά συστήματα και ηλιακή προστασία.
- η)Φυσικό αερισμό.
- θ)Εσωτερικές κλιματικές συνθήκες στις οποίες περιλαμβάνονται οι επιδιωκόμενες εσωτερικές κλιματικές συνθήκες.

Η πλήρης μέτρηση της ενεργειακής κατανάλωσης του κτιρίου είναι όπως γίνεται εύκολα αντιληπτό αρκετά δύσκολη και ακριβή διαδικασία. Έτσι έχει καθοριστεί το όριο του 90% της καταναλισκόμενης ενέργειας. Έχει αναπτυχθεί λοιπόν μια ακριβής μεθοδολογία μετρήσεων, σύμφωνα με την οποία μετράται κάθε κατανάλωση ξεκινώντας από τις πιο σημαντικές μέχρι να φτάσουμε στο όριο του 90% προσθέτοντας καταναλώσεις όσο αυτό είναι απαραίτητο. Τα βήματα που ακολουθούνται είναι τα εξής:

Βήμα 1 :Υπολογισμός της συνολικής καταναλισκόμενης ενέργειας.

Για κάθε διαφορετικό καύσιμο υπολογίζουμε ξεχωριστά την κατανάλωση.

Βήμα 2: Εισαγωγή δεδομένων για τις καταναλώσεις.

Στο βήμα αυτό υπολογίζουμε κάθε διαφορετικό τύπο κατανάλωσης. Συνήθως στις πρώτες επαναλήψεις της μεθόδου παίρνουμε 3-4 είδη και στη συνέχεια επεκτείνουμε.

Βήμα 3: Διαχωρισμός των καταναλώσεων.

Κάθε τύπος κατανάλωσης μπορεί να χωριστεί σε επιμέρους μικρότερες καταναλώσεις βάση τοποθεσίας, κυκλώματος, συστήματος κτλ. Έτσι καθίσταται πιο εύκολη η μέτρηση της συνολικής κατανάλωσης και παρέχονται χρήσιμα δεδομένα για το διαχειριστή του κτιρίου.

Βήμα 4 : Μέθοδοι μέτρησης.

Όπως προαναφέραμε το ελάχιστο όριο για την ενέργεια που μετράται είναι το 90%. Υπάρχουν διάφοροι τρόποι μέτρησης που μπορούν να χρησιμοποιηθούν όταν η άμεση μέτρηση καθίσταται δύσκολη ή ακριβή. Τα είδη μετρήσεων είναι τα εξής:

Άμεση μέτρηση :Προσφέρει υψηλή ακρίβεια και αξιοπιστία

Μέτρηση χρόνου: Μετράει τις ώρες λειτουργίας μια συσκευής με σταθερή κατανάλωση ισχύος, συνεπώς μπορεί να υπολογιστεί εύκολα η ενέργεια.

Έμμεση μέτρηση: Η μέτρηση του ζητούμενου μεγέθους προκύπτει από την μέτρηση ενός άλλου μεγέθους. Εδώ χρειάζεται προσοχή στην ακρίβεια.

Διαφορική μέτρηση: Το ζητούμενο μέγεθος προκύπτει από την διαφορά δύο άλλων μετρούμενων μεγεθών. Προσοχή ώστε τα μεγέθη αν είναι περίπου της ίδιας τάξης ώστε να μην υπάρχει μεγάλο σφάλμα.

Προσέγγιση : Στη μέθοδο αυτή ουσιαστικά δεν μετράται αλλά εκτιμάται η κατανάλωση ενός μικρού φορτίου .

Είναι πολύ σημαντική η σωστή επιλογή του μετρητή, της ακρίβειας του αλλά και η κωδικοποίηση του ώστε να μπορούμε τον ταυτοποιήσουμε. Στο σημείο αυτό εξετάζουμε αν είναι **χρήσιμες** οι πληροφορίες που μας δίνει κάθε μέτρηση. Αν όχι **ξαναγυρνάμε στο βήμα 3** και εξετάζουμε αν μπορούμε να διαχωρίσουμε διαφορετικά τις καταναλώσεις.

Βήμα 5: Εκτίμηση της κατανάλωσης βάσης των προτεινόμενων μετρήσεων.

Στο βήμα αυτό υπολογίζουμε την συνολική ενέργεια που μετράμε και την συγκρίνουμε με την αναμενόμενη βάση του σχεδιασμού για να δούμε την αποδοτικότητα του κτιρίου.

Βήμα 6: Έλεγχος ορίου 90%

Στο στάδιο αυτό εξετάζουμε αν η μετρούμενη ενέργεια είναι ικανοποιητική, αν όχι **επιστρέφουμε στο βήμα 3** και **προσθέτουμε** και άλλη μέτρηση.

Βήμα 7: Προγραμματισμός μετρήσεων

Γίνεται προγραμματισμός των μετρητών, καταγραφή της τοποθεσίας τους και ομαδοποίηση τους.

Βήμα 8: Στρατηγική μετρήσεων

Καταστρώνεται ένα διάγραμμα στρατηγικής των μετρήσεων.

Βήμα 9: Προσθήκη λεπτομερειών στα σχέδια και στο αρχείο

Καταγραφή όλων των αποφάσεων και παραδοχών που λήφθηκαν κατά τη διαδικασία της μέτρησης .

Με το τέλος αυτής της διαδικασίας μπορούμε πλέον να δημιουργήσουμε το ενεργειακό προφίλ του κτιρίου και να έχουμε μια σαφή εικόνα της συμπεριφοράς του σε σχέση με τις εξωτερικές καιρικές συνθήκες. Το προφίλ είναι ουσιαστικά ένα διάγραμμα κατανάλωσης (KWh/m^2) με θερμοκρασία (C), βασισμένο στις μετρήσεις που παίρνουμε από το σύστημα AMR όπως έχει περιγραφεί παραπάνω.

Εφόσον έχει δημιουργηθεί το διάγραμμα αυτό μπορεί να χρησιμοποιηθεί σε προβλέψεις κατανάλωσης με βάση τις καιρικές συνθήκες που επικρατούν. Φυσικά και δε μπορούμε να έχουμε πλήρη και ακριβή πρόβλεψη καθώς υπάρχουν και απρόβλεπτοι παράγοντες όπως είναι οι αιχμές. Το φαινόμενο αυτό έχει μελετηθεί και έχει αποδειχθεί ότι η κατανάλωση του κτιρίου χωρίζεται σε δύο συνιστώσες. Η πρώτη είναι η μέση κατανάλωση της ημέρας, η οποία είναι λιγότερο ευμετάβλητη και αλλάζει κυρίως από εποχή σε εποχή. Η δεύτερη συνιστώσα είναι στοχαστική και μας δίνει την απόκλιση από τη μέση τιμή λόγω απότομων αλλαγών σε καιρικές συνθήκες. Το άθροισμα των δύο αυτών συνιστωσών μας δίνει την ολική καμπύλη κατανάλωσης με βάση τη σχέση:

$$L = \mu + r$$

Όπου μ η μέση και r η στοχαστική συνιστώσα.

Η καμπύλη αυτή μπορεί να προσαρμοστεί ώστε να αυξηθεί η απόδοση του κτιρίου με διάφορες ενέργειες όπως βελτίωση μόνωσης, συντήρηση και αντικατάσταση ελαττωματικών καυστήρων και χρήση συσκευών νέας γενιάς με υψηλή ενεργειακή απόδοση.

3

Διαχείριση Ζήτησης Smart Grid

3 Διαχείριση ζήτησης – Smart Grid

3.1 Διαχείριση Ζήτησης Ενέργειας

Η διαχείριση ζήτησης ενέργειας που από εδώ και στο εξής θα αναφέρεται ως διαχείριση ζήτησης (Demand Side Management – DSM) περιλαμβάνει ενέργειες οι οποίες επηρεάζουν την ποσότητα ή τον τρόπο χρήσης της ενέργειας που καταναλώνεται από τους τελικούς χρήστες. Μια τέτοια ενέργεια μπορεί να είναι η μείωση της ζήτησης τις ώρες αιχμής όταν δηλαδή το σύστημα είναι περισσότερο πιεσμένο. Αυτό δεν οδηγεί στη μείωση της συνολικής κατανάλωσης αλλά μειώνει την ανάγκη για μονάδες εφεδρείας, άρα το κόστος παραγωγής της ενέργειας.

Παλαιότερα η διαχείριση ζήτησης ήταν εντελώς ανελαστική και πολλές φορές οι χρεώσεις δεν αντικατόπτριζαν πλήρως την κατανάλωση. Οι εταιρείες πολλές φορές κοστολογούσαν με βάση ιστορικά δεδομένα και προσέφεραν πολύ λίγα αν όχι καθόλου ανταποδοτικά οφέλη στους πελάτες για την μείωση της κατανάλωσής τους. Πλέον η τιμολογιακή πολιτική έχει αλλάξει με στόχο την παροχή ακριβώς αυτών των οφελών στους πελάτες, με στόχο τη μετατόπιση της κατανάλωσής τους σε διαφορετικές στιγμές λιγότερο επικίνδυνες για το σύστημα. Αυτό το εξελιγμένο σύστημα καλείται Διαχείριση Ζήτησης.

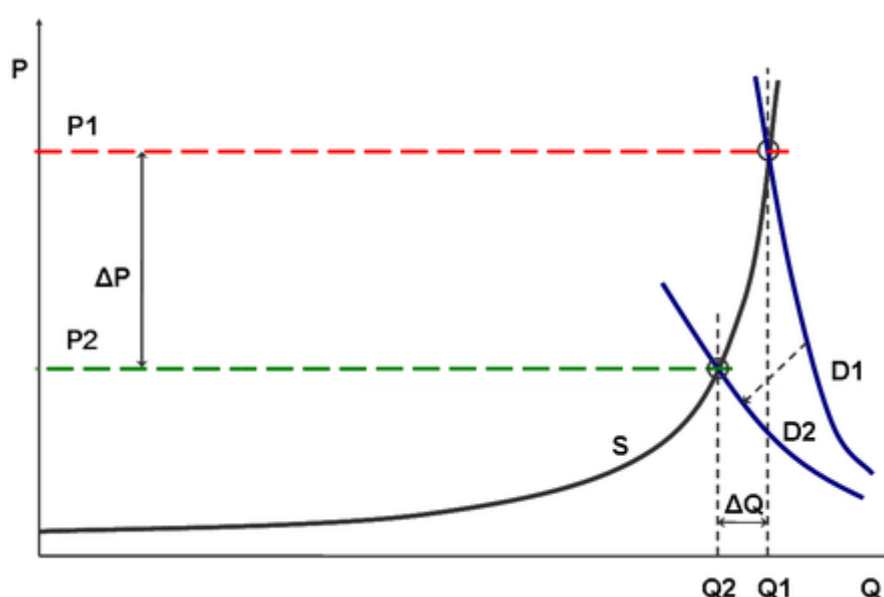
Η διαχείριση αυτή γίνεται δυναμικά αντιδρώντας σε κρίσιμες τιμές συνολικής κατανάλωσης ή σε αλλαγή στις τρέχουσες τιμές ενέργειας. Αυτό γίνεται μέσω μηνυμάτων από την εταιρεία παροχής είτε για την αποσύνδεση ενεργοβόρων συσκευών, είτε για την εκκίνηση μικρών μονάδων παραγωγής με σκοπό την αποφόρτιση του συστήματος. Αυτή η διαδικασία μπορεί να αφορά τόσο βιομηχανικές μονάδες όσο και οικιακούς χρήστες.

Σε γενικές γραμμές αναφερόμενοι στη διαχείριση ζήτησης εννοούμε μηχανισμούς μείωσης της ζήτησης σε ώρες αιχμής. Αυτό όπως είναι κατανοητό μειώνει τα ανώτατα επίπεδα απαιτούμενης ενέργειας, επομένως το κόστος παραγωγής καθώς χρειάζονται λιγότερες παραγωγικές μονάδες. Παρόλα αυτά επειδή ένα άλλο σημαντικό στοιχείο για το ενεργειακό δίκτυο είναι και η συνολική απόδοση του συστήματος ο ίδιος μηχανισμός μπορεί να απαιτήσει και αύξηση της κατανάλωσης από ορισμένους καταναλωτές σε στιγμές που υπάρχει υψηλή παραγωγή και χαμηλή ζήτηση.

Όπως είναι φυσικά εύκολα αντιληπτό οι πελάτες δεν είναι πρόθυμοι να παραδώσουν πλήρη έλεγχο στον πάροχο ηλεκτρικής ενέργειας. Γιαυτό το λόγο κυρίως εφαρμόζεται η τακτική των ανταποδοτικών οφελών με τη μορφή οικονομικών εκπτώσεων. Ακόμη πριμοδοτείται η επένδυση σε μικρές μονάδες παραγωγής και κυρίως σε μονάδες ανανεώσιμων πηγών ενέργειας, οι οποίες όχι μόνο καλύπτουν μέρος της ζητούμενης κατανάλωσης αλλά και επιστρέφουν ενέργεια στο σύστημα της

ώρες αδράνειας του τοπικού φορτίου. Όλα αυτά γίνονται δυνατά και άκρως αποδοτικά με τη δημιουργία ενός ευφυούς δικτύου όπως περιγράφεται παρακάτω.

Τα τεράστια οφέλη στην τιμολόγηση από τη μείωση της ζήτησης φαίνονται σαφώς στο επόμενο διάγραμμα. Όπως βλέπουμε η ανελαστική ζήτηση D1 οδηγεί σε υψηλό φόρτο δικτύου και πολύ υψηλή τιμή. Με την εφαρμογή της διαχείρισης ζήτησης σε ένα ευφύες δίκτυο η ζήτηση γίνεται πολύ πιο ελαστική (D2). Αυτό οδηγεί σε μειωμένη ζήτηση για ενέργεια, κάτι το οποίο μειώνει εκθετικά την τιμή της. Έτσι έχουμε μειωμένο κόστος για την εταιρεία παραγωγής αλλά και μειωμένα τιμολόγια για τους καταναλωτές.



Σύμφωνα με εκτιμήσεις μια μείωση της ζήτησης κατά μόλις 5% τις ώρες αιχμής μπορεί να οδηγήσει σε μείωση έως και 50% της τιμής της ηλεκτρικής ενέργειας. Αξίζει λοιπόν η οποιαδήποτε βραχυπρόθεσμη δυσκολία και μείωση του ελέγχου που ασκεί ο κάθε πελάτης πάνω στην κατανάλωσή του με στόχο τα μακροπρόθεσμα οικονομικά οφέλη που θα απολαύσει.

3.2 Smart Grid

Το ηλεκτρικό δίκτυο μεταφέρει ενέργεια από τα σημεία παραγωγής στους καταναλωτές και το δίκτυο παροχής λειτουργεί μέσω δύο βασικών συστημάτων, του συστήματος μετάδοσης και του συστήματος διανομής. Το σύστημα μεταφοράς μεταφέρει ενέργεια από τις μονάδες παραγωγής στους υποσταθμούς διανομής, ενώ το σύστημα διανομής μεταφέρει την ενέργεια από τους υποσταθμούς στους καταναλωτές. Το δίκτυο ενσωματώνει ακόμη στις μέρες μας πολλά τοπικά δίκτυα που χρησιμοποιούν κατανεμημένες πηγές ενέργειας για να ικανοποιούν τις ανάγκες τοπικών φορτίων.

Ο στόχος του ηλεκτρικού δικτύου είναι φυσικά η σύνδεση όλων των πελατών, ανεξαρτήτως τόπου και χρόνου, με τις πηγές ενέργειας οι οποίες παρέχουν άφθονη, καθαρή, φθηνή και αξιόπιστη ηλεκτρική ενέργεια. Για να γίνει αυτό με το βέλτιστο τρόπο χρειάζεται η δημιουργία ενός εξελιγμένου δικτύου με μεγάλες δυνατότητες και χρήση της σύγχρονης τεχνολογίας για τον έλεγχο και τη διαχείριση της ενεργειακής ροής.

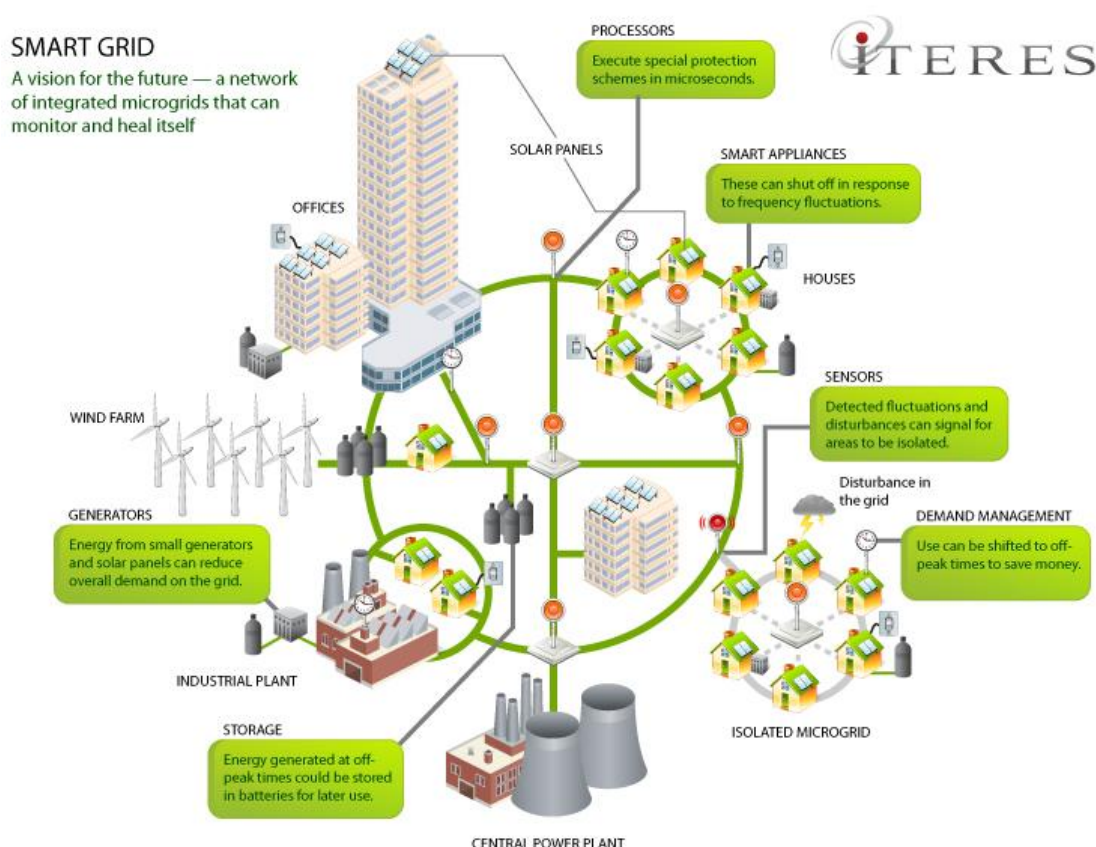
Για το λόγο αυτό έχει προβλεφθεί και οργανωθεί η δημιουργία ενός *έξυπνου δικτύου* το οποίο στη βέλτιστη λειτουργία του πρέπει να ικανοποιεί τα παρακάτω χαρακτηριστικά:

- Δυνατότητα αυτόματης επαναφοράς από ενεργειακές διαταραχές
- Άμεση συμμετοχή των καταναλωτών στη διαχείριση ζήτησης
- Αυξημένη αντοχή σε βλάβες υποδομών και επικοινωνιών
- Αυξημένη ποιότητα ενέργειας για χρήση στον 21^ο αιώνα
- Πλήρες σύστημα παραγωγής και αποθήκευσης
- Δυνατότητα συνεχούς εισαγωγής νέων υπηρεσιών
- Βελτιστοποίηση περιουσιακών στοιχείων και αποδοτική λειτουργία

Το έξυπνο αυτό δίκτυο χρησιμοποιεί μοντέρνα ψηφιακή τεχνολογία για τον ευφυή έλεγχο της ροής ενέργειας. Ακόμα χρησιμοποιεί τεχνολογία υπεραγωγών για μείωση απωλειών και ενσωματώνει όλες τις σύγχρονες εναλλακτικές πηγές ενέργειας όπως ηλιακή και αιολική. Σε συνδυασμό με τα συστήματα έξυπνων μετρητών που έχουν περιγραφεί πιο πάνω το δίκτυο εποπτεύει αυτόματα την ενεργειακή κατανάλωση. Έτσι έχει πλήρη συνολική εικόνα την οποία αναλύει και επιλέγει τη βέλτιστη κίνηση προς την οικονομία του συστήματος.

Μέσα στις βασικές ενέργειες που μπορεί να πραγματοποιήσει το ευφυές δίκτυο είναι η αμφίδρομη ψηφιακή επικοινωνία με τους τελικούς χρήστες. Με βασικό στόχο τη μείωση του κινδύνου στις ώρες αιχμής και την όσο το δυνατόν πιο ομοιόμορφη κατανάλωση στη διάρκεια της ημέρας, μπορεί να αυξομειώνει τις χρεώσεις και να ειδοποιεί τους καταναλωτές για αυτές, προφέροντάς τους κίνητρα και αντικίνητρα για την αλλαγή της κατανάλωσής τους. Τις ώρες αιχμής οι χρεώσεις αυξάνονται ενώ το

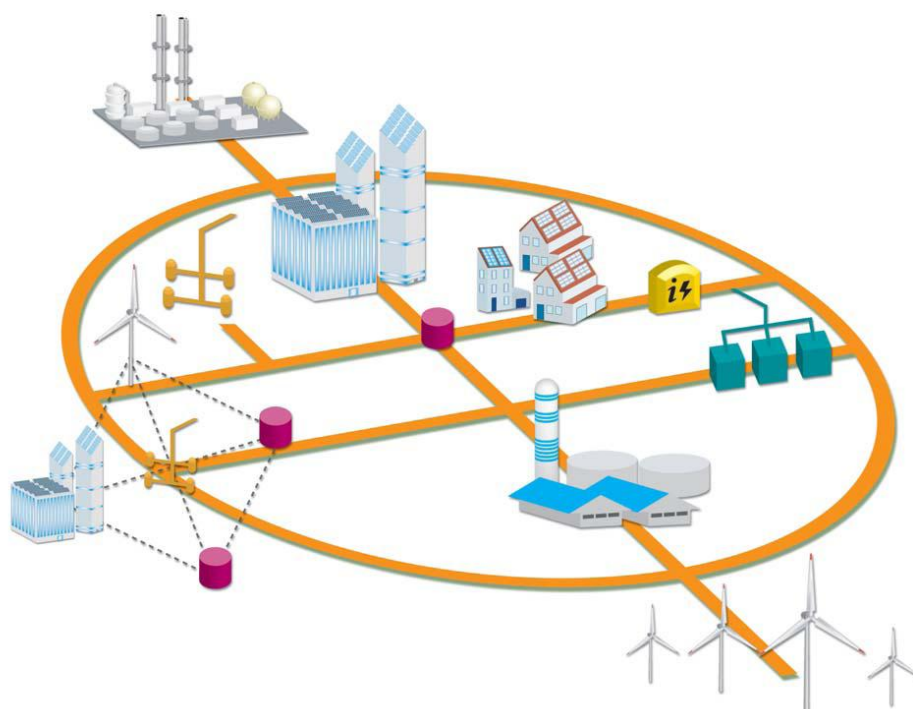
αντίστροφο συμβαίνει τις ώρες χαμηλής κατανάλωσης. Σε ένα πλήρως αυτοματοποιημένο δίκτυο του μέλλοντος η κεντρική διαχείριση θα μπορούσε ακόμη και να ελέγχει συσκευές καταναλωτών, ενεργοποιώντας ή απενεργοποιώντας τις με βάση μια λίστα προτεραιότητας την οποία έχει δώσει ο χρήστης και το ύψος της κατανάλωσής τους. Φυσικά πρέπει να προσφερθούν στον πελάτη και τα κατάλληλα ανταλλάγματα με τη μορφή συνήθως μειωμένων χρεώσεων.



Ένα τέτοιο δίκτυο γίνεται όλο και πιο απαραίτητο όσο περισσότεροι καταναλωτές συμμετέχουν ενεργά στην παραγωγή της ενέργειας με ανανεώσιμες οικιακές πηγές. Με αυτή την τακτική ο πελάτης καλύπτει όλες ή μέρος των αναγκών του και όταν υπάρχει περίσσεια παραγόμενης ενέργειας την προσφέρει πίσω στο δίκτυο επί πληρωμή. Αυτό φυσικά δημιουργεί πολλές διακυμάνσεις στο επίπεδο της διαθέσιμης στο σύστημα ενέργειας.

Η διακύμανση αυξάνεται ακόμα περισσότερο αναλογικά με το ποσοστό των ανανεώσιμων πηγών που συμμετέχουν στο σύστημα. Αυτό συμβαίνει γιατί οι πηγές αυτές είναι εγγενώς εξαρτημένες από φυσικούς παράγοντες και τις καιρικές

συνθήκες. Το ευφυές δίκτυο όμως είναι σε θέση να ανταπεξέρχεται στις μεταβολές αυτές, αυξομειώνοντας κατάλληλα τις δικές του παραγωγικές μονάδες.



Η πολυπλοκότητα του ευφυούς δικτύου απαιτεί συνεχή τεχνολογική εξέλιξη σε όλους τους τομείς. Στον τομέα της επικοινωνίας των συσκευών με τα κέντρα διαχείρισης απαιτείται ανάπτυξη πλήρως αυτοματοποιημένων υποσταθμών και συστημάτων διανομής, δίκτυα ασύρματης επικοινωνίας και οπτικών ινών για την απρόσκοπτη, ασφαλή και αξιόπιστη μετάδοση δεδομένων καθώς και εξελιγμένα συστήματα ελέγχου και διαχείρισης δεδομένων (supervisory control and data acquisition, SCADA). Στον τομέα των μετρήσεων εκτός από τους έξυπνους μετρητές που έχουν αναφερθεί πιο πάνω αναπτύσσονται συνεχώς νέα όργανα μετρήσεων όπως θερμικοί μετρητές πραγματικού χρόνου (RTTR) και μονάδες μέτρησης φάσης (PMU) οι οποίες λειτουργούν σε υψηλές ταχύτητες και ελέγχουν την ποιότητα της ηλεκτρικής ενέργειας, αντιδρώντας πολλές φορές προς τη βελτίωσή της. Ακόμα χρησιμοποιούνται και όλες οι πρόσφατες εξελίξεις στην τεχνολογία υλικών και κυρίως στους υπεραγωγούς καθώς και στα συστήματα λήψης αποφάσεων.

4

Μαθηματική περιγραφή

4 Μαθηματική περιγραφή

Η ανάλυση των δεδομένων βασίζεται σε μεθόδους γραμμικής άλγεβρας πάνω σε πίνακες. Πιο συγκεκριμένα χρησιμοποιείται κατά κύριο λόγο η μέθοδος Principal Component Analysis (PCA) και η θεωρία γύρω από ιδιοτιμές και ιδιοδιανύσματα.

4.1 Ιδιοτιμές και ιδιοδιανύσματα τετραγωνικού πίνακα

Οι ιδιοτιμές και τα ιδιοδιανύσματα αποτελούν ιδιότητες ενός πίνακα και δίνουν σημαντικές πληροφορίες για τον πίνακα αυτό. Σε γενικές γραμμές η επίδραση ενός πίνακα πάνω σε ένα διάνυσμα είναι η αλλαγή τόσο του μήκους όσο και της κατεύθυνσης του διανύσματος αυτού. Παρόλα αυτά υπάρχουν ορισμένα διανύσματα για κάθε πίνακα στα οποία η επίδραση του πίνακα αφορά μόνο το μήκος τους και όχι την κατεύθυνση, η οποία παραμένει ανεπηρέαστη. Αυτά τα διανύσματα αποτελούν τα ιδιοδιανύσματα του πίνακα. Ο παράγοντας κατά τον οποίο μεταβάλλεται το μήκος των διανυσμάτων είναι η αντίστοιχη ιδιοτιμή. Κάθε ιδιοτιμή μαζί με τα αντίστοιχα ιδιοδιανύσματά της, αν σε αυτά προσθέσουμε και το μηδενικό διάνυσμα αποτελούν έναν ιδιοχώρο του πίνακα.

Έστω A τετραγωνικός πίνακας μεγέθους $n \times n$. Μια ιδιοτιμή και ένα ιδιοδιάνυσμα του πίνακα A είναι αντίστοιχα ένας συντελεστής λ και ένα μη μηδενικό διάνυσμα x τέτοια ώστε

$$Ax = \lambda x.$$

Από τη σχέση αυτή προκύπτει ένας απλός τρόπος υπολογισμού των ιδιοτιμών και των ιδιοδιανυσμάτων του πίνακα A , μέσω της εξίσωσης

$$(A - \lambda I)x = 0, \quad x \neq 0.$$

Αυτό υπονοεί ότι ο πίνακας $A - \lambda I$ είναι μοναδιαίος και άρα

$$\det(A - \lambda I) = 0.$$

Αυτός ο ορισμός της ιδιοτιμής λ , ο οποίος δεν περιλαμβάνει άμεσα το αντίστοιχο ιδιοδιάνυσμα είναι η χαρακτηριστική εξίσωση ή το χαρακτηριστικό πολυώνυμο του A . Ο βαθμός του πολυωνύμου είναι και η τάξη του πίνακα. Είναι επομένως σαφές ότι ένας $n \times n$ πίνακας έχει n ιδιοτιμές, αν σε αυτές μετρηθούν φυσικά και οι πολλαπλές.

Έστω $\lambda_1, \lambda_2, \dots, \lambda_n$ οι ιδιοτιμές του πίνακα A , $x_1, x_2, \dots, x_n$ το σύνολο των αντίστοιχων ιδιοδιανυσμάτων και έστω Λ ο διαγώνιος $n \times n$ πίνακας που περιλαμβάνει στη διαγώνιό του τα στοιχεία λ_j και X ο $n \times n$ πίνακας του οποίου η j -οστή στήλη είναι το ιδιοδιάνυσμα x_j . Τότε

$$AX = X\Lambda.$$

Προσοχή πρέπει να δοθεί στο γεγονός ότι στο δεύτερο μέλος πρέπει το Λ να τοποθετηθεί δεξιά έτσι ώστε κάθε στήλη του X να πολλαπλασιαστεί με τη σωστή ιδιοτιμή. Αν τώρα όλα τα ιδιοδιανύσματα είναι γραμμικώς ανεξάρτητα (κάτι το οποίο φυσικά δεν ισχύει για όλους τους πίνακες) τότε υπάρχει ο αντίστροφος X^{-1} και επομένως

$$A = X \Lambda X^{-1}.$$

Αυτό καλείται *eigenvalue decomposition* και είναι μια πολύ σημαντική ιδιότητα καθώς μας επιτρέπει να ελέγχουμε τις ιδιότητες του A απλά ελέγχοντας το διαγώνιο πίνακα Λ . Ένα απλό παράδειγμα της χρησιμότητας της μεθόδου αυτής είναι η ύψωση σε δύναμη, μια επίπονη διαδικασία η οποία όμως εδώ απλοποιείται σημαντικά καθώς ισχύει

$$A^p = X \Lambda^p X^{-1}.$$

Αν τα ιδιοδιανύσματα δεν είναι γραμμικώς ανεξάρτητα τότε δεν υπάρχει ο διαγώνιος πίνακας Λ και φυσικά δεν ισχύει η παραπάνω σχέση καθώς και όλα τα πλεονεκτήματα που μας προσφέρει σε υπολογισμούς.

Οι ιδιοτιμές παίζουν πολύ σημαντικό ρόλο σε περιπτώσεις όπου ένας πίνακας αποτελεί ένα μετασχηματισμό από ένα διανυσματικό χώρο στον ίδιο το χώρο. Το πιο σημαντικό παράδειγμα ενός τέτοιου μετασχηματισμού είναι τα συστήματα συνήθων γραμμικών διαφορικών εξισώσεων.

4.2 Principal Components Analysis (PCA)

Η principal component analysis χρησιμοποιείται ευρέως στη στατιστική και στη διαχείριση μεγάλου όγκου δεδομένων. Η βασική ιδιότητα της μεθόδου είναι η μείωση του χώρου των πραγματικών δεδομένων σε ένα μικρότερο υποχώρο στον οποίο συμμετέχουν τα πιο σημαντικά στοιχεία. Έτσι ορίζεται ως ορθογωνικός γραμμικός μετασχηματισμός, ο οποίος μετασχηματίζει τα δεδομένα σε ένα νέο σύστημα συντεταγμένων έτσι ώστε η μεγαλύτερη διακύμανση οποιασδήποτε προβολής των δεδομένων στο σύστημα αυτό να περιέχεται στο πρώτο στοιχείο, το οποίο καλείται και πρώτο βασικό συστατικό (principal component).

Βασικό τμήμα της μεθόδου είναι η αφαίρεση του μέσου της κατανομής από τα δεδομένα. Αυτό οδηγεί στην εύρεση μιας βάσης η οποία κατά την προσέγγιση των δεδομένων ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα.

Αν έχουμε λοιπόν έναν πίνακα δεδομένων X^T με μηδενικό εμπειρικό μέσο τότε η μέθοδος PCA δίνεται από τη σχέση

$$Y^T = X^T W = V \Sigma$$

όπου $W\Sigma V^T$ είναι η αποσύνθεση σε μοναδιαίες τιμές (Singular Value Decomposition – SVD) του πίνακα X .

Αν έχουμε ένα σύνολο σημείων σε έναν Ευκλείδειο χώρο (τα οποία αναπαριστούν τα δεδομένα μας), το πρώτο principal component αντιστοιχεί σε μια γραμμή η οποία περνά ανάμεσα στα σημεία με στόχο να ελαχιστοποιήσει το συνολικό τετραγωνικό σφάλμα με τα σημεία αυτά. Το κάθε επόμενο συστατικό κάνει ακριβώς το ίδιο με τη διαφορά ότι έχει αφαιρεθεί από τα σημεία η συσχέτισή τους με το προηγούμενο συστατικό. Έτσι τα επόμενα συστατικά έχουν ολοένα και μικρότερη σημασία σε σχέση με τα προηγούμενα. Ένας σαφής δείκτης της σημασίας τους είναι η αντίστοιχη ιδιοτιμή του καθενός, αφού το μέγεθός της αντικατοπτρίζει και τη συμμετοχή του κάθε διανύσματος στη λύση.

Έχοντας κατανοήσει τη σημασία των principal components ο υπολογισμός τους γίνεται με βάση τη σχέση:

$$w_1 = \arg_{||w||=1} \max \text{Var}\{w^T x\} = \arg_{||w||=1} \max E\{(w^T x)^2\}$$

Πρόκειται για επαναληπτική μέθοδο στην οποία το k -οστό συστατικό μπορεί να βρεθεί αφαιρώντας τα πρώτα $k-1$ συστατικά από το x :

$$\hat{x}_{k-1} = x - \sum_{i=1}^{k-1} w_i w_i^T x$$

Και χρησιμοποιώντας το νέο αυτό σύνολο σαν τα αρχικά δεδομένα έχουμε το νέο PC ως εξής:

$$w_k = \arg_{||w||=1} \max E\{(w^T \hat{x}_{k-1})^2\}$$

Επομένως η εύρεση των βασικών συστατικών είναι ισοδύναμη με την εύρεση της αποσύνθεσης σε μοναδιαίες τιμές (Singular Value Decomposition) του πίνακα δεδομένων X ,

$$X = W\Sigma V^T$$

Και στη συνέχεια την εύρεση του μειωμένου πίνακα δεδομένων Y προβάλλοντας τον πίνακα X στο μειωμένο χώρο που δημιουργείται από τα L πρώτα μοναδιαία διανύσματα W_L :

$$Y = W_L^T X = \Sigma_L V_L^T$$

Ακολουθώντας αυτή τη μέθοδο ουσιαστικά ο πίνακας W των μοναδιαίων διανυσμάτων του X είναι πανομοιότυπος με τον πίνακα W των ιδιοδιανυσμάτων του πίνακα των συνδιακυμάνσεων $C = XX^T$ καθώς ισχύει:

$$XX^T = W\Sigma\Sigma^T W^T$$

Αυτή η τελευταία παρατήρηση μας δίνει μια ευκολότερη μέθοδο από αυτή με τους επαναληπτικούς τύπους που περιγράφηκε παραπάνω, η οποία για μεγάλους πίνακες δεδομένων μπορεί να παρουσιάζει μεγάλη υπολογιστική πολυπλοκότητα. Η νέα αυτή μέθοδος η οποία βασίζεται στη συνδιακύμανση περιγράφεται αναλυτικά παρακάτω και είναι αυτή η οποία χρησιμοποιείται στους αλγόριθμους.

Έστω X ο πίνακας των δεδομένων μεγέθους $M \times N$. Τότε ο εμπειρικός μέσος της κάθε διάστασης $m=1, \dots, M$ είναι

$$U[m] = \frac{1}{N} \sum_{n=1}^N X[m, n]$$

Στη συνέχεια σχηματίζουμε τον πίνακα B ο οποίος ορίζεται ως

$$B = X - uh$$

και περιλαμβάνει τα δεδομένα μετά την αφαίρεση του εμπειρικού μέσου. Στη σχέση αυτή h είναι ένα μοναδιαίο διάνυσμα $h[n] = 1$ για $n=1 \dots N$.

Από τον πίνακα B μπορούμε τώρα να υπολογίσουμε τον πίνακα συνδιακύμανσης (covariance matrix) ως

$$C = \frac{1}{N} \sum B B^*.$$

Αυτό ουσιαστικά είναι ισοδύναμο με μια πράξη πινάκων $C = B B^T$.

Έχοντας τώρα τον πίνακα αυτό αναζητούμε τον πίνακα V ο οποίος περιλαμβάνει τα ιδιοδιανύσματα του πίνακα C και ουσιαστικά διαγωνοποιεί τον C έτσι ώστε να ισχύει η σχέση

$$V^{-1} C V = D$$

όπου D ο διαγώνιος πίνακας των ιδιοτιμών του C . Έχοντας φτάσει στο σημείο αυτό ο πίνακας V των ιδιοδιανυσμάτων ουσιαστικά αποτελεί τον πίνακα των ζητούμενων βασικών στοιχείων (principal components). Τα στοιχεία αυτά δίνονται κατά σειρά προτεραιότητας η οποία ορίζεται με βάση τις ιδιοτιμές. Το στοιχείο με τη μεγαλύτερη ιδιοτιμή είναι και το πρώτο pc το οποίο έχει και τη μεγαλύτερη σημασία. Ακολουθούν τα υπόλοιπα στοιχεία με φθίνουσα σειρά.

Με βάση όλα αυτά μπορούμε πλέον να επιλέξουμε τα στοιχεία με τη μεγαλύτερη σημασία, ο αριθμός των οποίων καθορίζεται ανάλογα με το πρόβλημα και χρησιμοποιώντας τα στοιχεία αυτά να περιγράψουμε το πρόβλημα σε μια νέα βάση, με σαφώς μειωμένο αριθμό μεταβλητών και δεδομένων.

5

Αλγόριθμος

5 Αλγόριθμος

5.1 Επιλογή ποιοτικών δεικτών (KPIs)

Η εξατομίκευση των χρεώσεων και υπηρεσιών που παρέχονται από εταιρείες παροχής ενέργειας στους πελάτες τους κρίνεται ιδιαίτερα σημαντική μετά την απελευθέρωση της ενεργειακής αγοράς. Για να γίνει αυτό με τον πιο αποδοτικό τρόπο πρέπει να υπάρξει ένα σύστημα συσταδοποίησης των πελατών σε ομάδες με παρόμοια ενεργειακά προφίλ. Η συσταδοποίηση αυτή μπορεί να γίνει με βάση κάποιους δείκτες οι οποίοι περιγράφουν το κάθε προφίλ ποιοτικά. Συγκεκριμένα χρησιμοποιούμε τέσσερις δείκτες οι οποίοι μπορούν να περιγράψουν ικανοποιητικά ένα προφίλ με τη μέγιστη αφαιρετικότητα και παράλληλα με την ελάχιστη δυνατή απώλεια πληροφορίας. Οι δείκτες αυτοί είναι:

1. P_{max}/P_{avg} (Το πηλίκο της μέγιστης αιχμής προς τη μέση τιμή των μετρήσεων μας δίνει μια εικόνα του ύψους της αιχμής ως προς την υπόλοιπη γραφική παράσταση)
2. Αριθμός αιχμών (Μας δίνεται μια σαφής εικόνα του αριθμού των αιχμών που εμφανίζει το προφίλ στο συγκεκριμένο διάστημα)
3. P_{rang}/P_{avg} (Το πηλίκο της μέσης τιμής των αιχμών προς τη μέση τιμή των μετρήσεων μας δίνει πληροφορία για τη διακύμανση της γραφικής παράστασης, όπως και για τη διαφορετικότητα του ύψους των διαφόρων αιχμών του διαστήματος αυτού)
4. $(P_{end}-P_{start})/P_{avg}$ (Ο δείκτης αυτός μας δίνει ένα σταθμισμένο ρυθμό ανόδου ή καθόδου της γραφικής παράστασης)

Οι δείκτες αυτοί υπολογίζονται ανά τρίωρο. Προκειμένου να έχουμε μια πλήρη εικόνα της ημέρας με μικρή απώλεια πληροφορίας επιλέγουμε η ανάλυση να γίνει σε 6 τρίωρα ξεκινώντας από τις 06:00 το πρωί και τελειώνοντας τα μεσάνυχτα (00:00 το βράδυ) κρίνοντας ότι τις υπόλοιπες ώρες υπάρχει μικρή κατανάλωση για το ενεργειακό σύστημα.

Έτσι για το κάθε τρίωρο σχηματίζεται ένα διάνυσμα $f_n = [KPI_1, KPI_2, KPI_3, KPI_4]$, $n=1 \dots 6$ το οποίο περιγράφει ποιοτικά τη συμπεριφορά του πελάτη για το συγκεκριμένο τρίωρο. Μια ολόκληρη ημέρα μπορεί να δοθεί σαν $d = [f_1, f_2, f_3, f_4, f_5, f_6]$, περιγράφοντας και τα 6 τρίωρα μαζί σε ένα σύνθετο διάνυσμα.

Με δεδομένα τα παραπάνω καταρτίζονται πίνακες για τον κάθε πελάτη. Ο κάθε πίνακας αφορά ένα δείκτη και καλύπτει διάστημα ενός έτους. Έτσι για τον κάθε πελάτη έχουμε 4 πίνακες μεγέθους 365×6 , που περιγράφουν πλήρως ποιοτικά το προφίλ του. Επιπροσθέτως σχηματίζουμε έναν επιπλέον πίνακα ο οποίος έχει μέγεθος 365×4 και περιλαμβάνει σε κάθε θέση την τιμή του τρίωρου στο οποίο εμφανίζεται η

μέγιστη τιμή του δείκτη της συγκεκριμένης στήλης. Έτσι έχουμε μια πλήρη ποιοτική ανάλυση για τον κάθε πελάτη.

5.2 Ανάλυση πινάκων

Η χρήση και ανάλυση των πινάκων αυτών γίνεται στο περιβάλλον Matlab, χρησιμοποιώντας στοιχεία από γραμμική άλγεβρα και συγκεκριμένα τη μέθοδο Principal Component Analysis και τις ιδιοτιμές-ιδιοδιανύσματα τετραγωνικών πινάκων.

Έστω $A1$ ο πίνακας του δείκτη P_{max}/P_{avg} για τον πελάτη 1 και $A2$ ο ίδιος πίνακας για τον πελάτη 2. Και οι 2 πίνακες έχουν μέγεθος 365×6 . Κάνοντας τη στοιχειώδη πράξη $A1 * A1'$, δηλαδή πολλαπλασιάζοντας τον πρώτο πίνακα με τον αντίστροφό του έχουμε ένα τετραγωνικό πίνακα 365×365 . Σύμφωνα με τη μέθοδο PCA τα ιδιοδιανύσματα του πίνακα αυτού αποτελούν τα principal components τα οποία θα αναλύσουμε. Το πιο σημαντικό pc είναι το πρώτο, το οποίο καθορίζεται από τη μεγαλύτερη ιδιοτιμή και δίνει τη μεγαλύτερη συσχέτιση των πινάκων. Έτσι υπολογίζοντας τις ιδιοτιμές του πίνακα γινομένου, η μεγαλύτερη από αυτές μας δίνει το ιδιοδιάνυσμα του πρώτου pc, ενώ οι μεγαλύτερες κατά απόλυτο τιμές του PC μας δίνουν τη γραμμή (και άρα την ημέρα) στην οποία ο πίνακας παρουσιάζει τη μεγαλύτερη αυτοσυσχέτιση. Κάνοντας το ίδιο και για τον άλλο πελάτη και συγκρίνοντας τα PCs τους συσχετίζουμε τους δύο αυτούς πελάτες.

Αν τώρα χρησιμοποιήσουμε τον πέμπτο πίνακα ο οποίος περιλαμβάνει και τους 4 δείκτες μαζί με τον ίδιο τρόπο έχουμε και πάλι σαν αποτέλεσμα την ημέρα στην οποία τα προφίλ των 2 πελατών παρουσιάζουν τη μεγαλύτερη ποιοτική συσχέτιση μεταξύ τους. Παράλληλα μελετώντας τον πίνακα ανά στήλες μπορούμε να βρούμε τη διακύμανση του κάθε δείκτη, μια πληροφορία που μας βοηθάει στο να καταλάβουμε τη συνέπεια ή μη του πελάτη στη χρήση της ηλεκτρικής ενέργειας, παράγοντας σημαντικός για την τιμολόγησή του.

Ακόμα πιο ενδιαφέρον είναι ίσως το γεγονός ότι μετρώντας τη διακύμανση αυτή μπορούμε να κάνουμε πρόβλεψη για τη συμπεριφορά του πελάτη στο μέλλον, πράγμα εξαιρετικά σημαντικό για μια εταιρεία παροχής. Η μέθοδος αυτή λοιπόν προσφέρεται για τη δημιουργία μιας εκτιμήτριας για τον κάθε πελάτη, αρκεί φυσικά να υπάρχει κάποιος δεδομένος βαθμός συνέπειας από την πλευρά του.

5.3 Αλγόριθμος

5.3.1 Μαθηματική ανάλυση αλγορίθμου

Τα στοιχεία προς επεξεργασία προέρχονται από μετρήσεις σε διάφορα κτήρια με βάση τα πρότυπα της ASRAE. Τα πρότυπα αυτά καθορίζουν μια πλήρη διαδικασία για την ενεργειακή μέτρηση και λειτουργική θερμοδυναμική ανάλυση του κτηρίου, μετά την οποία μπορεί το κτήριο να καταταχθεί σε κάποια ενεργειακή κλάση. Οι κλάσεις καθορίζουν την ενεργειακή λειτουργικότητα του κτηρίου με βάση κάποια διεθνώς αναγνωρισμένα κριτήρια.

Οι μετρήσεις γίνονται με τη χρήση συστημάτων AMR (Automatic Meter Reading). Ένα τέτοιο σύστημα περιλαμβάνει μετρητές ενέργειας οι οποίοι είναι εγκατεστημένοι στο κάθε κτήριο και οι οποίοι μετρούν την ηλεκτρική ισχύ που καταναλώνεται. Ακόμα υπάρχει ένα κεντρικό σύστημα διαχείρισης το οποίο συγκεντρώνει και αποθηκεύει τα δεδομένα. Ενδιάμεσο στάδιο είναι η αποστολή των στοιχείων από τους μετρητές στο σύστημα διαχείρισης με ειδικές συσκευές επικοινωνίας.

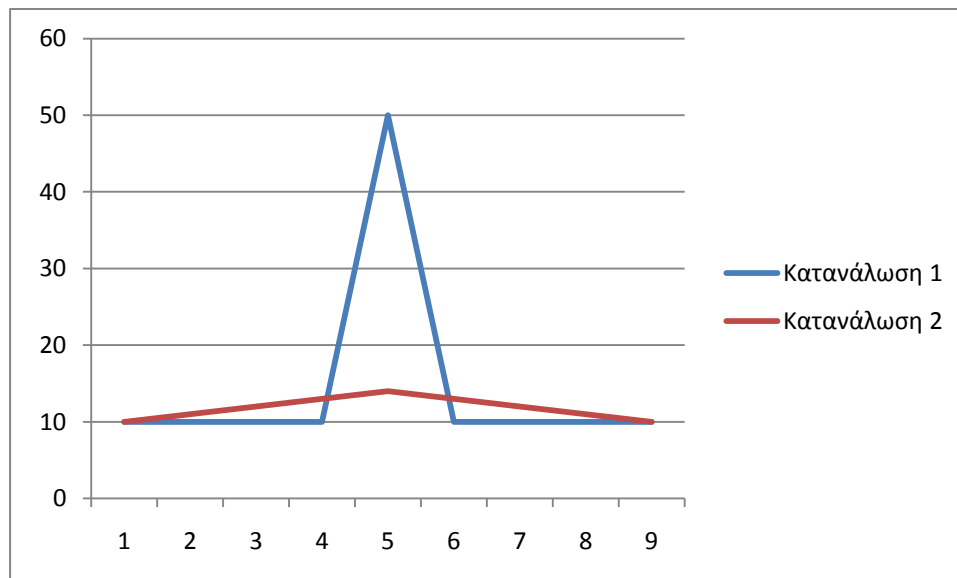
Η αποθήκευση των στοιχείων γίνεται σε βάση δεδομένων και τα στοιχεία που μας ενδιαφέρουν είναι η κατανάλωση σε KWh και η χρονική στιγμή στην οποία έγινε η μέτρηση αυτή ως UNIX timestamp.

Το πρώτο και βασικό σημείο του αλγορίθμου είναι η ανίχνευση αιχμών στο διάγραμμα που προκύπτει από τα στοιχεία της βάσης δεδομένων. Τα δεδομένα αφορούν μετρήσεις ανά χρονικό διάστημα 15 λεπτών. Έτσι για να έχουμε το διάγραμμα μιας ημέρας χρειαζόμαστε $86400/(15*60)=96$ μετρήσεις. Σχηματίζεται λοιπόν ένα διάγραμμα με τιμές οριζόντιου άξονα 0-95, οι οποίες αφορούν την τιμή του δεκαπεντάλεπτου στο οποίο έχει ληφθεί η συγκεκριμένη μέτρηση (time offset) και κατακόρυφο άξονα τιμές ηλεκτρικής ισχύος (KWh). Όπως είναι προφανές το διάγραμμα αυτό θα περιέχει διακυμάνσεις στην ισχύ, ανάλογα με την ώρα της ημέρας και την χρήση ηλεκτρικής ενέργειας που συνέβη στο κτήριο. Σε συγκεκριμένες λοιπόν χρονικές στιγμές κατά τις οποίες μπαίνουν σε λειτουργία συσκευές του κτηρίου εμφανίζεται μια απότομη αύξηση της καταναλισκόμενης ισχύος, η οποία οδηγεί σε κάποια κορυφή στο διάγραμμα. Στόχος μας είναι η υπό προϋποθέσεις αναγνώριση μιας τέτοιας κορυφής ως αιχμή ή όχι. Για να δεχθούμε μια κορυφή ως αιχμή θα πρέπει αυτή να πληροί κάποιες προϋποθέσεις. Αυτές είναι η γωνία ανόδου και καθόδου να ξεπερνούν μια συγκεκριμένη τιμή, και η ποσοστιαία άνοδος της κορυφής να είναι αρκετά μεγάλη.

Η πρώτη προϋπόθεση είναι απαραίτητη για να αποκόπτε κορυφές οι οποίες σχηματίζονται από μια διαρκώς αυξανόμενη μεν, με πολύ μικρό ρυθμό δε κατανάλωση. Σε τέτοιες περιπτώσεις δεν προκύπτει κάποιος κίνδυνος για το δίκτυο επομένως δεν υπάρχει και ενδιαφέρον. Ο στόχος μας είναι η ανίχνευση κορυφών οι οποίες αν υπερτεθούν (από πολλούς καταναλωτές την ίδια χρονική στιγμή) να

οδηγήσουν το σύστημά μας στα όρια της ευσταθούς λειτουργίας του, οπότε θα χρειαστεί και ειδική αντιμετώπιση από την εταιρεία παροχής ενέργειας. Είναι απολύτως σαφές λοιπόν ότι μια ελαφρώς αυξανόμενη κατανάλωση δεν παρουσιάζει τέτοιους κινδύνους και άρα μπορεί να αφεθεί χωρίς επεξεργασία. Τα ακόλουθα σχήματα δίνουν μια εποπτική εικόνα μιας αιχμής και μιας απλής κορυφής η οποία δε μας ενδιαφέρει, ενώ δίνονται και οι τιμές κατανάλωσης σε πίνακα.

<u>Κατανάλωση 1</u>	<u>Κατανάλωση 2</u>
10	10
10	11
10	12
10	13
50	14
10	13
10	12
10	11
10	10



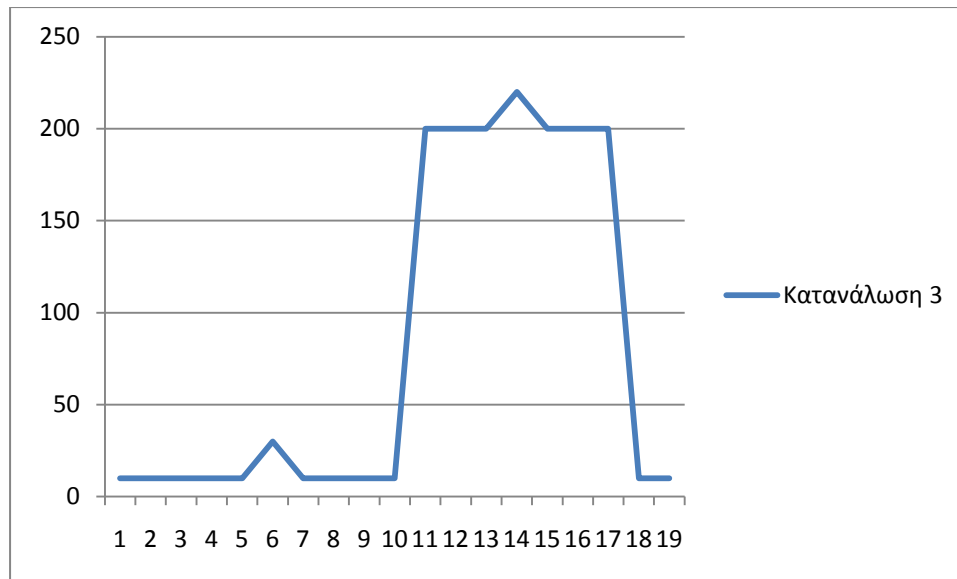
Με τον τρόπο αυτό ελέγχουμε ταυτόχρονα τόσο την ύπαρξη κορυφής όσο και το ύψος της αιχμής, καθώς γνωρίζοντας εκ των προτέρων τη διαβάθμιση του οριζόντιου άξονα αν ελέγξουμε τη γωνία ανόδου και καθόδου άμεσα ξέρουμε και το ύψος της κορυφής. Έτσι ουσιαστικά η γωνία (η οποία εισάγεται ως μεταβλητή

deg_threshold) μας δίνει το κατώφλι για την αποδοχή μιας κορυφής ως αιχμή, με βάση το απόλυτο ύψος της κορυφής.

Η δεύτερη προϋπόθεση έχει ως στόχο να αποκλείσει κορυφές μικρού ενδιαφέροντος οι οποίες παρουσιάζουν μικρή διακύμανση από τις γειτονικές χρονικά τιμές ενέργειας. Αυτό πρακτικά σημαίνει ότι δύο κορυφές ίδιου ύψους αλλά με διαφορετικές τιμές βάσης θα τυχουν διαφορετικής αντιμετώπισης.

Είναι κατανοητό ότι όταν έχουμε χαμηλή κατανάλωση ενέργειας μια απότομη άνοδος, έστω και αν είναι σχετικά μικρή πρέπει να θεωρηθεί ως αιχμή. Για παράδειγμα μια άνοδος από την τιμή 10 στην τιμή 30 είναι σαφώς μια αιχμή που μας ενδιαφέρει και πρέπει να ανιχνευθεί. Η απόλυτη τιμή της ανόδου αυτής καθορίζεται όπως περιγράφηκε παραπάνω από την τιμή της γωνίας. Είναι λιγότερο σαφές όμως το τι θα πρέπει να γίνει όταν έχουμε μια ίδια άνοδο αλλά σε μεγαλύτερη κατανάλωση. Για παράδειγμα μια αύξηση κατανάλωσης από 200 σε 220 παρουσιάζει ίδια απόλυτη αύξηση της τάξεως των 20 μονάδων. Παρόλα αυτά μπροστά στην ήδη υπάρχουσα κατανάλωση η κορυφή αυτή είναι τόσο μικρή που θα μπορούσε να προκληθεί ακόμη και από λάθος μέτρηση, ή από κάποια μικρή διακύμανση λόγω αστοχίας κάποιας συσκευής, παράγοντες οι οποίοι δε θα θέλαμε να συμπεριληφθούν στην ανάλυσή μας. Για το λόγο αυτό εισάγουμε μια νέα μεταβλητή με όνομα val_threshold η οποία ορίζει με βάση τις συγκεκριμένες ανάγκες του κάθε συστήματος την κατώτατη ποσοστιαία διαφορά για την οποία θα πρέπει να ανιχνευτεί η αιχμή. Η τιμή αυτή είναι ουσιαστικά το πηλίκο της τιμής της κορυφής προς την τιμή βάσης. Στο παραπάνω παράδειγμα, το οποίο παρουσιάζεται και στο επόμενο σχήμα, στην πρώτη περίπτωση έχουμε μια τιμή $A=30/10=3$ ενώ στη δεύτερη έχουμε $B=220/200=1,1$. Είναι προφανής η τεράστια διαφορά καθώς στην περίπτωση A έχουμε αύξηση 300% ενώ στην περίπτωση B μόλις 10%.

<u>Κατανάλωση 3</u>	
10	200
10	200
10	200
10	220
10	200
30	200
10	200
10	10
10	10
10	



Φυσικά ίσως υπάρξουν και περιπτώσεις στις οποίες είναι απαραίτητη και η επεξεργασία της κορυφής B ως αιχμή. Ακριβώς για το λόγο αυτό η τιμή κατωφλίου εισάγεται ως μεταβλητή καθοριζόμενη από το χρήστη, ούτως ώστε να υπάρχει πλήρης έλεγχος.

Έχοντας αναλύσει τα θεωρητικά στοιχεία και τη χρησιμότητα των διαφόρων μεταβλητών που χρησιμοποιούνται, πρέπει πλέον να δούμε τον ακριβή τρόπο με τον οποίο γίνεται η ανίχνευση μιας αιχμής από ένα διάγραμμα κατανάλωσης. Έστω λοιπόν ένας πίνακας που περιλαμβάνει τιμές κατανάλωσης ενός συγκεκριμένου κτηρίου για μια συγκεκριμένη χρονική περίοδο. Έστω ακόμη ότι έχουμε ήδη καθορίσει με βάση τις ανάγκες της συγκεκριμένης περίπτωσης τις δύο βασικές μεταβλητές κατωφλίου γωνίας και ποσοστιαίας ανόδου με τιμές α και β αντίστοιχα. Προφανώς για να υπολογίσουμε γωνία ανόδου και γωνία καθόδου πρέπει να έχουμε σε κάθε φάση υπολογισμού 3 τιμές κατανάλωσης. Έτσι κάθε φορά παίρνουμε από τον πίνακα 3 διαδοχικές τιμές προς επεξεργασία, έστω left, middle και right. Από βασική τριγωνομετρία η κλίση μιας ευθείας από ένα σημείο $A(x_1, y_1)$ σε ένα σημείο $B(x_2, y_2)$ δίνεται από τη σχέση $\lambda = \frac{y_2 - y_1}{x_2 - x_1}$, η οποία αποτελεί και την εφαπτομένη της ευθείας. Στη δική μας περίπτωση και έστω για τη ευθεία ανόδου βρίσκουμε την κλίση μέσω της σχέσης $\lambda = \frac{\text{middle} - \text{left}}{15}$. Ο παρονομαστής έχει την τιμή 15 καθώς αυτή είναι η τιμή μεταξύ 2 διαδοχικών μετρήσεων κατανάλωσης (θεωρούμε ότι το διάγραμμα είναι σε άξονες ενέργειας-λεπτών). Έχοντας πλέον την τιμή της εφαπτομένης μπορούμε εύκολα να βρούμε τη γωνία με την τριγωνομετρική σχέση $\varphi_1 = \arctan(\lambda)$. Ομοίως για την γωνία καθόδου έχουμε τη σχέση $\varphi_2 = \arctan(\frac{\text{right} - \text{middle}}{15})$.

Επομένως για να έχουμε αιχμή θα πρέπει η τιμή της γωνίας φ_1 να είναι μεγαλύτερη από το κατώφλι μας `deg_threshold` και ταυτόχρονα η τιμή της φ_2 να

είναι μικρότερη από το (-deg_threshold) καθώς εφόσον μιλάμε για κάθοδο θα πρέπει να είναι αρνητική.

Εφόσον μια κορυφή περάσει από αυτό τον πρώτο έλεγχο μένει να δούμε πλέον αν ικανοποιεί και τη δεύτερη συνθήκη κατωφλίου. Αυτή δίνεται από τη σχέση $\frac{middle}{left} > val_threshold$ και εφόσον ικανοποιηθεί και η σχέση αυτή η κορυφή αναγνωρίζεται ως αιχμή και αποθηκεύεται σε πίνακα της μορφής [timestamp, KWh], όπου timestamp η χρονική στιγμή στην οποία συνέβη η αιχμή και KWh η τιμή της αιχμής.

Με βάση όλα τα παραπάνω μπορούμε από οποιαδήποτε κατανάλωση να ανιχνεύσουμε τις επιθυμητές αιχμές. Δημιουργούμε έτσι ημερήσιους πίνακες αιχμών οι οποίοι είναι και τα βασικά εργαλεία για την περαιτέρω ανάλυση του συστήματος.

Ένα άλλο σημείο που είναι άξιο αναφοράς είναι το πώς ανακτούμε τα στοιχεία από τη βάση δεδομένων. Η βασική προσέγγιση είναι φυσικά η ανάκτηση όλων των στοιχείων ανά ημέρα και η επεξεργασία όλων τους ανά τριάδες. Αυτό σημαίνει ότι ψάχνουμε για αιχμές μέσα σε ένα παράθυρο 45 λεπτών, το ελάχιστο που μπορούμε να έχουμε. Κάποιες όμως φορές είναι επιθυμητό να ανιχνεύουμε αιχμές σε μεγαλύτερο παράθυρο, για παράδειγμα δύο ή τριών ωρών ώστε να μπορούμε να προβλέψουμε στο μέλλον την κατανάλωση. Γι αυτό το λόγο εισάγεται μια ελεγχόμενη από το χρήστη μεταβλητή data_interval, η οποία καθορίζει το πόσες μετρήσεις θα ανακτήσουμε από τη βάση. Η default τιμή 1 σημαίνει ότι παίρνουμε όλες τις μετρήσεις της ημέρας και εργαζόμαστε όπως παραπάνω. Αν ανεβάσουμε την τιμή της για παράδειγμα στην τιμή 2, αυτό σημαίνει ότι πλέον ανακτούμε από τη βάση όχι όλες τις μετρήσεις αλλά μία παρά μία, παραλείποντας την ενδιάμεση τιμή. Έτσι ουσιαστικά στα προς επεξεργασία στοιχεία δεν έχουμε τιμές ανά 15 λεπτά αλλά ανά 30, και παίρνοντας μια τριάδα στοιχείων όπως πάντα ανιχνεύουμε αιχμές σε ένα παράθυρο 90 λεπτών. Σαφώς περαιτέρω αύξηση της τιμής data_interval αυξάνει το εν λόγω παράθυρο.

Όλη η παραπάνω επεξεργασία παραμένει εν πολλοίς ίδια εκτός από το σημείο στο οποίο υπολογίζουμε τις γωνίες ανόδου και καθόδου. Εκεί όπως γίνεται εύκολα αντιληπτό το διάστημα στον οριζόντιο άξονα δεν είναι πλέον 15 αλλά εξαρτάται από την τιμή data_interval. Έτσι ο τύπος για τη γωνία ανόδου γίνεται $\phi1 = \arctan\left(\frac{middle-left}{data_interval*15}\right)$ για να μας δώσει το σωστό αποτέλεσμα και αναλόγως προσαρμόζεται και η σχέση για τη γωνία καθόδου.

Έχοντας όλες τις αιχμές σε πίνακες μπορούμε να συνεχίσουμε την ανάλυση με τη χρονική συσχέτισή τους. Αυτό σημαίνει την ανίχνευση των αιχμών που συμβαίνουν σε κοντινές χρονικές στιγμές διαφορετικών ημερών. Το ενδιαφέρον που παρουσιάζεται στο σημείο αυτό είναι προφανές, καθώς οι αιχμές αποτελούν βασικό σημείο της ενεργειακής συμπεριφοράς ενός πελάτη και καθορίζουν σε μεγάλο βαθμό το ενεργειακό του προφίλ. Αν μπορούμε επομένως να συσχετίσουμε τις ημερήσιες

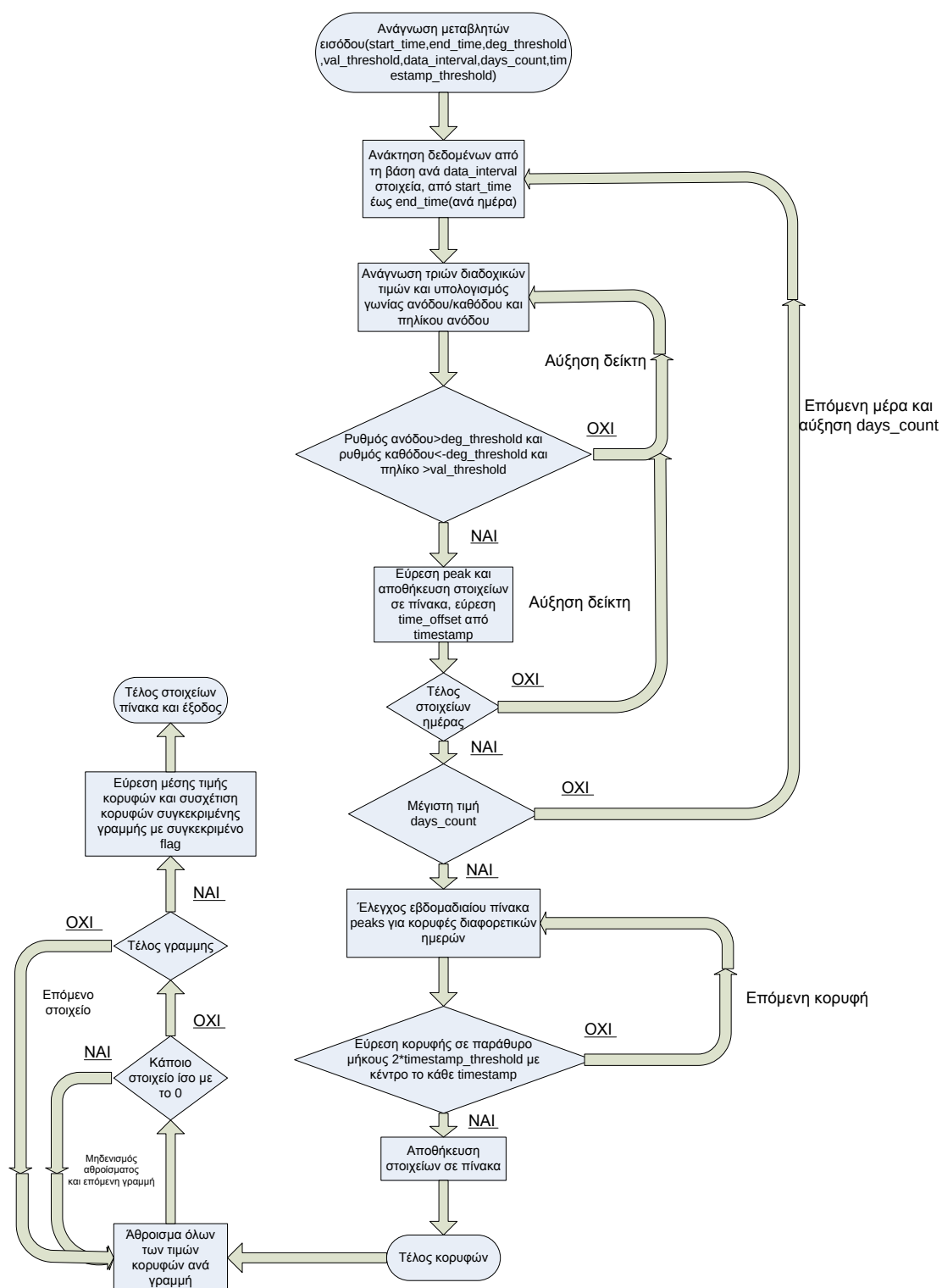
αιχμές για μια ολόκληρη εβδομάδα μπορούμε αυτομάτως να δούμε το κατά πόσο ο συγκεκριμένος πελάτης παρουσιάζει συνέπεια ή όχι στην κατανάλωσή του.

Για να γίνει η συσχέτιση πρέπει αρχικά να ορίσουμε την έννοια του `timestamp_offset`. Όπως έχουμε ήδη πει οι αιχμές αποτελούν ζεύγη (χρονικής στιγμής, ενέργειας). Με δεδομένο ότι η ανάλυση κάθε ημέρας ξεκινάει από την ίδια χρονική στιγμή (για παράδειγμα πάντα τα μεσάνυχτα), μπορούμε να ορίσουμε ένα μέτρο της απόστασης μιας χρονικής στιγμής από την αρχή των μετρήσεων. Αυτό είναι το `offset` και δίνει ανεξαρτήτως του για ποια ημέρα μιλάμε το σημείο της ημέρας στο οποίο συνέβη η αιχμή. Ο υπολογισμός του είναι απλός αν σκεφτούμε ότι μεταξύ διαδοχικών μετρήσεων μεσολαβούν 15 λεπτά ή αλλιώς $15 \cdot 60 = 900$ δευτερόλεπτα. Γνωρίζοντας επομένως την αρχική στιγμή της ημέρας (`start_time`) και το τρέχον χρονικό σημείο της αιχμής (`timestamp`) το `offset` δίνεται από τη σχέση $\text{offset} = (\text{timestamp} - \text{start_time}) / 900$ και είναι ένας ακέραιος από 1 έως και 96.

Βρίσκοντας τα `offsets` όλων των αιχμών και συγκρίνοντας τα μπορούμε να βρούμε συσχετισμένες αιχμές. Στο σημείο αυτό εισάγουμε και ένα παράγοντα ευαισθησίας ως προς τη σύγκριση των `offsets`. Δημιουργούμε ένα παράθυρο γύρω από κάθε χρονική στιγμή αιχμής, εύρους $2 \cdot \text{timestamp_threshold}$, όπου `timestamp_threshold` είναι μια μεταβλητή ορισμένη από το χρήστη. Έτσι δεχόμαστε μικρές αποκλίσεις στη χρονική στιγμή στην οποία καταγράφηκε η αιχμή, οι οποίες μπορεί να οφείλονται σε σφάλματα εξοπλισμού (όπως λάθος στην καταγραφή χρονικής στιγμής ή καθυστέρηση αποθήκευσης στοιχείων στη βάση δεδομένων κτλ). Πρακτικά αυτό σημαίνει ότι δύο αιχμές σε χρονικές στιγμές `time1` και `time2` έχουν ίδιο `offset` αν επικαλύπτονται σε οποιοδήποτε σημείο τα παράθυρα (`time1 - timestamp_threshold`, `time1 + timestamp_threshold`) και (`time2 - timestamp_threshold`, `time2 + timestamp_threshold`).

Με βάση όλα αυτά ανιχνεύουμε για μια ολόκληρη εβδομάδα τις κοινές αιχμές. Ιδιαίτερο ενδιαφέρον παρουσιάζουν περιπτώσεις όπου έχουμε επαναλαμβανόμενες αιχμές για αρκετές ημέρες της εβδομάδας. Ένας καταναλωτής με συσχετισμένες αιχμές για 7 ημέρες είναι εξαιρετικά συνεπής και άρα προβλέψιμος ως προς την κατανάλωσή του. Επειδή μάλιστα συνήθως η ενεργειακή συμπεριφορά αλλάζει τα σαββατοκύριακα μπορούμε να χωρίσουμε την εβδομάδα σε $5 + 2$ ημέρες και να ελέγχουμε για συνέπεια τις 5 πρώτες ημέρες οι οποίες αντιστοιχούν στις εργάσιμες. Σαν έξοδος του αλγορίθμου αυτού επιστρέφεται το `offset` των αιχμών, οι τιμές τους καθώς και η μέση τιμή αυτών για να έχουμε μια εποπτική εικόνα του ενεργειακού προφίλ του πελάτη.

Το διάγραμμα ροής δεδομένων του κομματιού του αλγορίθμου που περιγράφηκε πιο πάνω δίνεται στο σχήμα που ακολουθεί:



Έχοντας πλέον τελειώσει με την εύρεση και τη χρονική συσχέτιση των αιχμών, το δεύτερο κομμάτι της ανάλυσης περιλαμβάνει την ποιοτική ανάλυση του προφίλ ενός πελάτη. Για το σκοπό αυτό και όπως έχει ήδη αναφερθεί πιο πάνω έχουμε ορίσει τους τέσσερις δείκτες ποιοτικής ανάλυσης. Το πρώτο βήμα για τον υπολογισμό τους είναι ο σωστός χωρισμός της ημέρας σε τρίωρα. Εφόσον όλη η προηγούμενη ανάλυση έχει γίνει με βάση τα offsets θα τα χρησιμοποιήσουμε και εδώ. Υποθέτοντας και πάλι ότι η κάθε ημέρα ξεκινά από τα μεσάνυχτα και ότι η ανάλυσή μας αρχίζει από τις 06:00 το πρωί, η πρώτη μέτρηση που πρέπει να πάρουμε θα έχει offset 24. Αυτό γιατί από τη στιγμή 00:00 μέχρι τη στιγμή 06:00 μεσολαβούν 6 ώρες, δηλαδή 24 δεκαπεντάλεπτα. Κάθε τρίωρο περιλαμβάνει 12 μετρήσεις, επομένως η ανάλυσή μας γίνεται για τα εξής διαστήματα [24-36), [36-48), [48-60), [60-72), [72-84) και τέλος [84-96). Έτσι καλύπτουμε το μεγαλύτερο εύρος της ημέρας μέσα στο οποίο έχουμε τη μεγαλύτερη κατανάλωση.

Για το κάθε ένα από τα διαστήματα αυτά πρέπει πρώτα να υπολογίσουμε το μέσο όρο της κατανάλωσης. Αυτό γίνεται απλά αθροίζοντας όλες τις τιμές ενέργειας του διαστήματος και διαιρώντας με τον αριθμό τους, το 12 ($P_{avg} = \frac{\sum_{i=1}^{12} P_i}{12}$), όπου P_i η κατανάλωση στο offset i .

Στη συνέχεια για το πρώτο KPI ελέγχουμε τον πίνακα των αιχμών για το ίδιο διάστημα offset και αν υπάρχει αιχμή μέσα στο διάστημα αυτό τότε υπολογίζουμε τον δείκτη με βάση τη σχέση $KPI_1 = \frac{P_{max}}{P_{avg}}$, όπου P_{max} η τιμή της μέγιστης αιχμής.

Για το δεύτερο KPI γίνεται ένας απλός έλεγχος του πίνακα των αιχμών στο συγκεκριμένο διάστημα και αύξηση ενός μετρητή που δίνει τον αριθμό των αιχμών κάθε φορά που βρίσκουμε αιχμή εντός του διαστήματος.

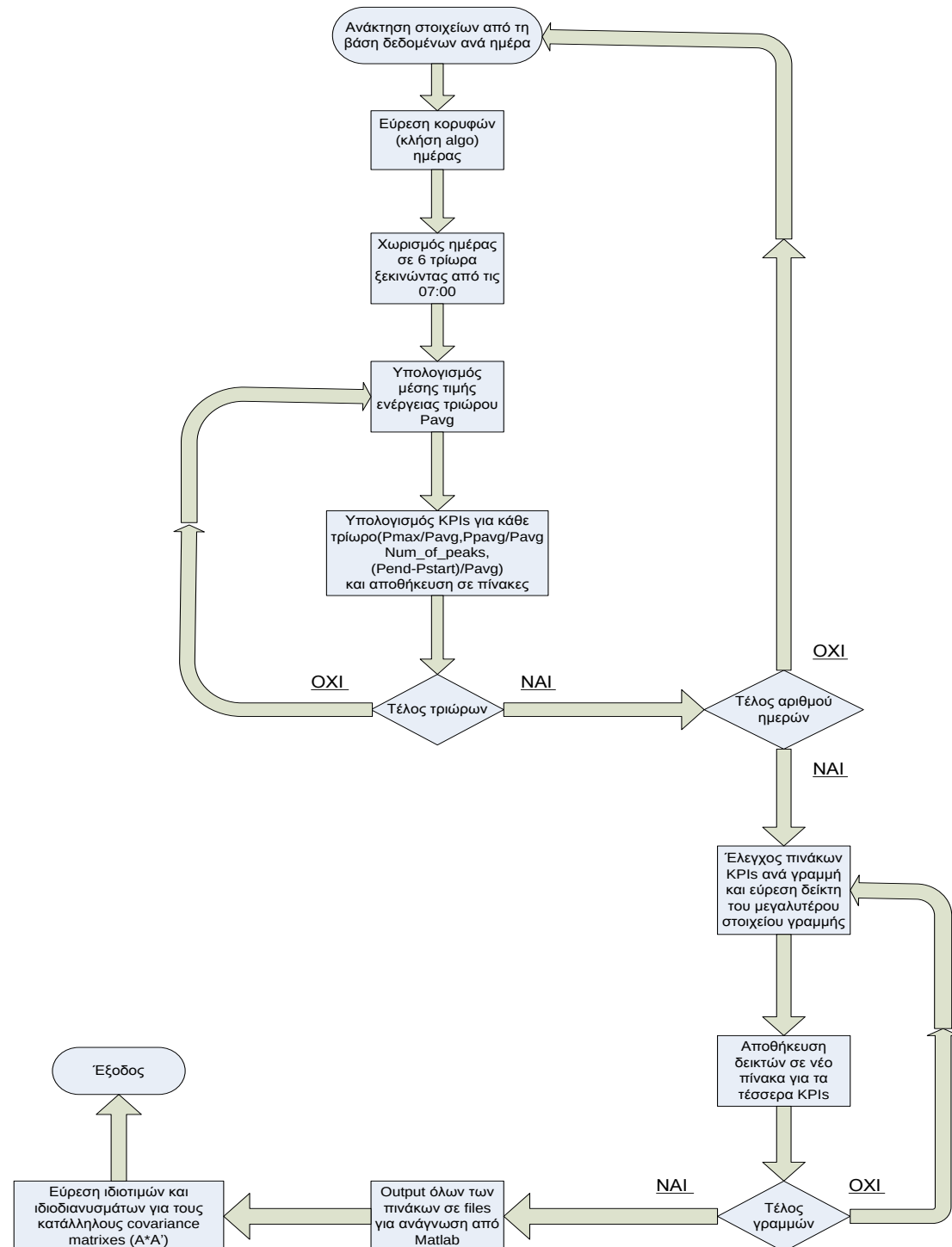
Για το τρίτο KPI χρειαζόμαστε επιπλέον τη μέση τιμή των αιχμών του διαστήματος. Αυτή υπολογίζεται με τον τύπο $P_{avg} = \frac{\sum P_{peak}}{count}$, όπου ο αριθμητής αποτελεί το άθροισμα των τιμών των κορυφών του διαστήματος και ο παρονομαστής ο αριθμός τους (ουσιαστικά είναι το KPI_2). Στη συνέχεια απλά υπολογίζουμε τον τρίτο δείκτη $KPI_3 = \frac{P_{avg}}{P_{avg}}$.

Τέλος μας μένει να υπολογίσουμε τον τέταρτο και τελευταίο δείκτη ο οποίος αποτελεί το ρυθμό ανόδου ή καθόδου της γραφικής παράστασης σύμφωνα με τη σχέση $KPI_4 = \frac{P_{end} - P_{start}}{P_{avg}}$, όπου P_{end} , P_{start} είναι η τελική και αρχική μέτρηση του διαστήματος αντίστοιχα.

Με τη διαδικασία αυτή έχουμε καταρτίσει τους πίνακες των ποιοτικών δεικτών. Μας απομένει η συγχώνευση όλων των δεικτών σε ένα πίνακα. Για να το κάνουμε αυτό διατρέχουμε την κάθε γραμμή του κάθε πίνακα KPI και καταγράφουμε το δείκτη του τριώρου στο οποίο έχουμε τη μεγαλύτερη τιμή. Έτσι έχουμε τελικά ένα

πίνακα ο οποίος σε κάθε γραμμή περιλαμβάνει τέσσερις τιμές, από 1 έως 6 η καθεμία, οι οποίες δείχνουν που ανιχνεύθηκε η μέγιστη τιμή του κάθε KPI.

Ακολουθεί το διάγραμμα ροής δεδομένων:



5.3.2 Ανάλυση πινάκων στο περιβάλλον MATLAB

Έστω οι πίνακες έξοδος των αλγορίθμων όπως έχουν περιγραφεί προηγουμένως. Επιλέγουμε να ασχοληθούμε με τους πίνακες που αφορούν το KPI_1 δηλαδή το $\frac{P_{max}}{P_{avg}}$. Η ανάλυση που ακολουθεί γίνεται ενδεικτικά για τους πίνακες αυτούς αλλά είναι ακριβώς ίδια για οποιαδήποτε άλλη κατηγορία πινάκων.

Έστω λοιπόν A ο 365X6 πίνακας που αναφέρεται στις ετήσιες μετρήσεις του χρήστη A και B ο αντίστοιχος πίνακας του χρήστη B. Στόχος μας είναι η ανάλυση των πινάκων αυτών χρησιμοποιώντας ιδιότητες των ιδιοτιμών και τη μέθοδο PCA. Αρχικά πρέπει να υπολογίσουμε τις ιδιοτιμές και τα ιδιοδιανύσματα των πινάκων αυτών, με τη χρήση του περιβάλλοντος MATLAB.

Σύμφωνα με τη θεωρία τα principal components είναι τα ιδιοδιανύσματα του πίνακα συνδιασποράς, δηλαδή του πίνακα $A \cdot A'$ στην περίπτωση του πίνακα A. Υπολογίζουμε λοιπόν για το λόγο αυτό τους πίνακες $A \cdot A'$ και $B \cdot B'$ οι οποίοι είναι πλέον τετραγωνικοί μεγέθους 365X365. Οι πίνακες αυτοί από θεωρία γραμμικής άλγεβρας θα έχουν πραγματικές ιδιοτιμές.

Στη συνέχεια υπολογίζουμε τις ιδιοτιμές και τα ιδιοδιανύσματα των πινάκων αυτών. Έτσι προκύπτουν δύο πίνακες ιδιοτιμών με 365 ιδιοτιμές ο καθένας και δύο πίνακες 6X365 των οποίων οι στήλες αποτελούν τα 365 ιδιοδιανύσματα για τον κάθε πίνακα. Φυσικά ανάλογα με τις τιμές των αρχικών πινάκων θα υπάρχουν μικρές και μεγάλες πραγματικές πάντα ιδιοτιμές. Από τη θεωρία της μεθόδου PCA οι μεγαλύτερες ιδιοτιμές είναι αυτές οι οποίες αντιστοιχούν στα ιδιοδιανύσματα με τη μεγαλύτερη σημασία. Η συμμετοχή των ιδιοδιανυσμάτων στη διαδικασία είναι φθίνουσα όσο μειώνεται η τιμή της αντίστοιχης ιδιοτιμής.

Αυτό που θέλουμε ουσιαστικά είναι ένας μηχανισμός ο οποίος να αναγνωρίζει ποιες διαστάσεις του προβλήματος είναι οι πιο σημαντικές και άρα πρέπει να επιλεγούν. Με άλλα λόγια στη δική μας περίπτωση πρέπει να βρούμε ποιες γραμμές των πινάκων έχουν τη μεγαλύτερη συσχέτιση μεταξύ τους, δηλαδή σε ποιες ημέρες οι δύο καταναλωτές παρουσιάζουν μεγάλες ομοιότητες στην ενεργειακή τους συμπεριφορά. Αυτό μπορεί να γίνει επιλέγοντας το σημαντικότερο ιδιοδιάνυσμα από κάθε πελάτη και προσθέτοντάς τα. Προκύπτει ένα νέο διάνυσμα με 365 τιμές. Συγκρίνοντας τις τιμές αυτές κατά απόλυτη τιμή μπορούμε να πούμε ότι οι μεγαλύτερες τιμές αντιστοιχούν σε σημαντικές για το πρόβλημα διαστάσεις οι οποίες πρέπει να ληφθούν υπόψη. Μάλιστα μπορούμε να ορίσουμε ότι σημαντική συνεισφορά στο πρόβλημα έχουν οι διαστάσεις των οποίων η τιμή είναι μεγαλύτερη από ένα συγκεκριμένο κατώφλι, για παράδειγμα απαιτούμε η τιμή να είναι μεγαλύτερη από 0,6. Έτσι βρίσκουμε ποιες είναι οι ημέρες στις οποίες έχουμε τη μεγαλύτερη συσχέτιση προφίλ των δύο πελατών. Συνήθως αυτό είναι αρκετό για σωστές παρατηρήσεις. Σε περίπτωση που θέλουμε μεγαλύτερη ακρίβεια μπορούμε να επιλέξουμε περισσότερα ιδιοδιανύσματα, αν και εφόσον οι ιδιοτιμές τους έχουν μη αμελητέες σε σχέση με τη μεγαλύτερη τιμές.

Με τον τρόπο αυτό μπορούμε άμεσα να βρούμε ομοιότητες των προφίλ των καταναλωτών σε σχέση με το εκάστοτε στοιχείο το οποίο βρίσκεται υπό μελέτη. Ελέγχοντας όλα τα στοιχεία του κάθε πελάτη και κάνοντας τη σύγκριση με τους υπόλοιπους μπορούμε να κατηγοριοποιήσουμε πελάτες με βάση την ενεργειακή συμπεριφορά τους. Για παράδειγμα αν 2 πελάτες έχουν υψηλή συσχέτιση σε όλα τα στοιχεία τους σε κοινές ημέρες (δηλαδή έχουν αντίστοιχες τιμές στα ιδιοδιανύσματά τους μεγαλύτερες από 0,6), τότε μπορούμε να πούμε ότι οι δύο αυτοί πελάτες μπορούν να ανήκουν στην ίδια κατηγορία.

Ακόμα μπορούμε να αναγνωρίσουμε κοινές συμπεριφορές (pattern matching) με βάση τα ιδιοδιανύσματα (κοινές τιμές σε 5+2 ημέρες κτλ). Αυτό γίνεται με βάση το γεγονός ότι οι συνιστώσες του ιδιοδιανύσματος εκφράζουν το μέτρο της συνεισφοράς της συγκεκριμένης διάστασης στο γενικό πρόβλημα. Επομένως για ένα προφίλ με κάποια ενεργειακή συνέπεια, και επομένως με επαναλαμβανόμενες τιμές, το ιδιοδιάνυσμα προς μελέτη θα εμφανίζει και αυτό με τη σειρά του ανάλογες ιδιότητες όπως επαναληπτικότητα. Για παράδειγμα ένας καταναλωτής με συνεπή ενεργειακή συμπεριφορά τις πέντε καθημερινές της εβδομάδας, και μοναδικές αποκλίσεις τα σαββατοκύριακα θα παρουσιάζει στο ιδιοδιάνυσμά του αυτή ακριβώς τη συμπεριφορά, με πεντάδες τιμών που θα επαναλαμβάνονται, με κάποιο λογικό ποσοστό σφάλματος, και δύο τιμές που θα παρουσιάζουν απόκλιση. Κοιτώντας επομένως μόνο το πρώτο PC μπορούμε να αναγνωρίσουμε ιδιότητες του ενεργειακού προφίλ. Έτσι μπορούμε και να κάνουμε κατηγοριοποίηση των πελατών με βάση αυτό το γεγονός.

6

Test cases

6 Test cases

6.1 Test case 1

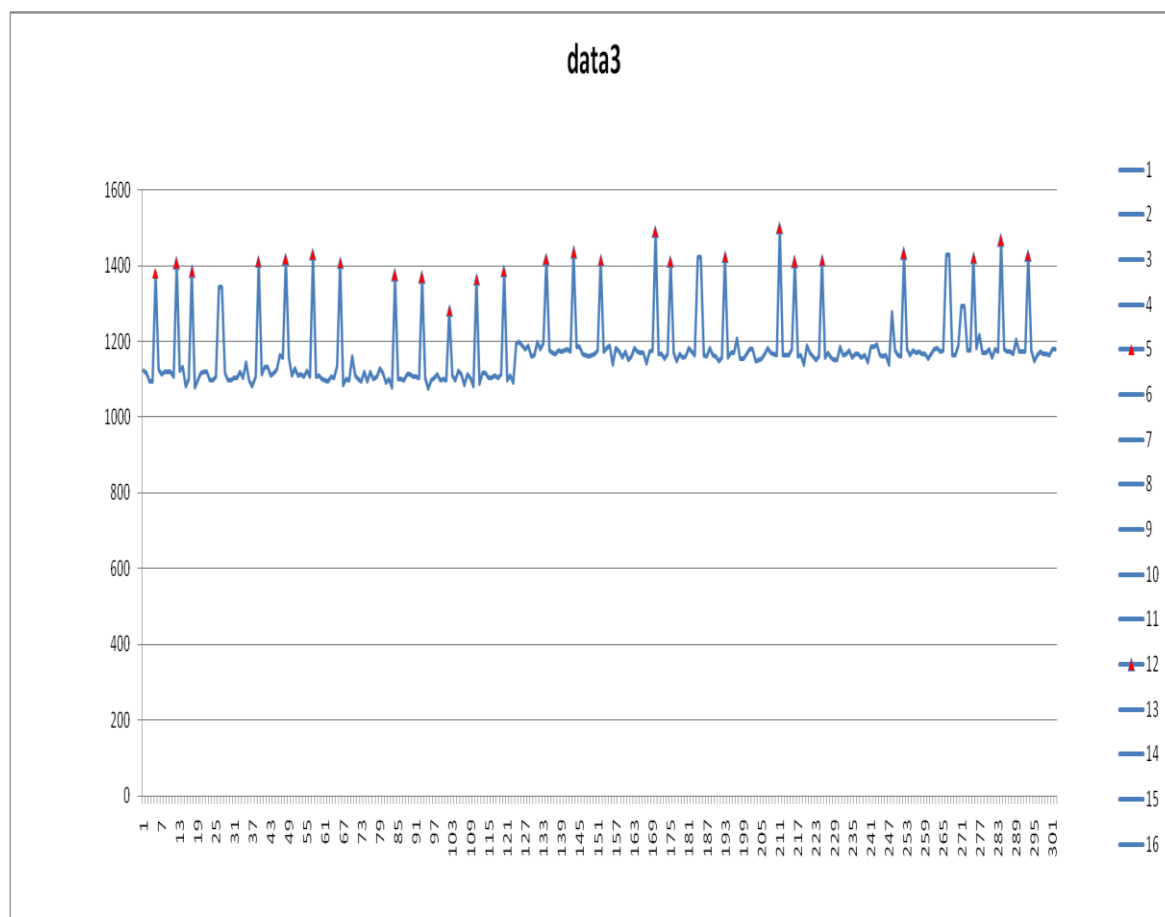
Οι μετρήσεις του δοκιμαστικού dataset ξεκινούν από 12:19:29 στις 4/5/2009 και σταματούν 00:01:59 στις 5/6/2009 (με κάποια κενά σε ώρες ή και ολόκληρες ημέρες). Έχουν ληφθεί μετρήσεις ανά χρονικό διάστημα 6 δευτερολέπτων. Επειδή το χρονικό αυτό διάστημα είναι πολύ μικρότερο από αυτό που έχουμε υποθέσει στην όλη ανάλυση θα πρέπει να κάνουμε μικρές τροποποιήσεις στον κώδικά μας. Το πολύ μικρό διάστημα οδηγεί σε τεράστιο αριθμό μετρήσεων ανά ημέρα. Η ανάλυση σε τόσο μεγάλο αριθμό μετρήσεων είναι εξαιρετικά χρονοβόρα. Έτσι δε θα εργασθούμε με ολόκληρες ημέρες αλλά με συγκεκριμένα διαστήματα.

Στην ανάλυσή μας είχαμε μετρήσεις ανά δεκαπεντάλεπτο και έτσι μια ημέρα είχε καλυφθεί πλήρως με 96 μετρήσεις. Θα ξεκινήσουμε με ανάλυση αιχμών του δείγματος προς έλεγχο του κώδικα. Χρειαζόμαστε εδώ τουλάχιστον 100 μετρήσεις για να έχουμε καλό δείγμα. Περισσότερες από 300 μετρήσεις κρίνονται υπερβολικές για την ανάλυση επομένως θα προσομοιώσουμε μια ημέρα με 300 μετρήσεις ανά 6 δευτερόλεπτα. Αυτό δημιουργεί ένα χρονικό παράθυρο της τάξεως των 30 λεπτών ($300 \cdot 6/60$). Αυτές οι 300 μετρήσεις θα χρησιμοποιηθούν στην πρωτεύουσα ανάλυση για την ανίχνευση αιχμών. Επιλέγουμε μια καθημερινή και σε ώρες που να υπάρχει φόρτο εργασίας σε γραφεία και υπηρεσίες μιας και οι μετρήσεις μας έχουν παρθεί από γραφείο. Μπορούμε λοιπόν να επιλέξουμε τη Δευτέρα 4/5/2009 που τυχαίνει να είναι και η πρώτη ημέρα των μετρήσεων, και να αρχίσουμε την ανάλυσή μας από την πρώτη μέτρηση που όπως είπαμε αντιστοιχεί στη χρονική στιγμή 12:19:29 και για 30 λεπτά μετά από αυτή τη στιγμή.

Από απλή παρατήρηση των μετρήσεων βλέπουμε εύκολα ότι οι δύο πρώτες φάσεις παραμένουν ουσιαστικά σταθερές στο διάστημα αυτό. Άρα οποιαδήποτε αιχμή οφείλεται καθαρά και μόνο στην τρίτη φάση. Γι' αυτό και περιορίζουμε την ανάλυσή μας στις τιμές ενέργειας της φάσης αυτής και μόνο αγνοώντας τις άλλες δύο. Φυσικά χρειάζονται μερικές τροποποιήσεις στον κώδικα τόσο για την προσπέλαση των στοιχείων από τη βάση δεδομένων όσο και για την ανάλυσή τους.

Τρέχουμε τον αλγόριθμο για τις 300 τιμές με μεταβλητές εισόδου (1241428769, 1241430675, 1, 50, 1.15) δηλαδή αρχικό και τελικό timestamp, 1 για data_interval δηλαδή επιλέγουμε όλες τις τιμές ανά μια, 50 μοίρες γωνία ανόδου και 1.15 λόγο κορυφής προς προηγούμενη τιμή. Μετά από δοκιμές βλέπουμε ότι αυτές οι τιμές δίνουν τα καλύτερα αποτελέσματα για το συγκεκριμένο dataset.

Στο διάγραμμα που ακολουθεί παρουσιάζονται όλες οι τιμές καθώς και οι κορυφές που ανιχνεύθηκαν.



Στη συνέχεια θα προχωρήσουμε στην περαιτέρω ανάλυση των δεδομένων μας. Για να το κάνουμε αυτό πρέπει να επεξεργαστούμε δεδομένα διαφορετικών και διαδοχικών ημερών. Και πάλι θα γίνει προσομοίωση μιας ημέρας από συγκεκριμένο αριθμό μετρήσεων του dataset. Για να ελέγξουμε πλήρως την ορθότητα του κώδικα θα πρέπει να δημιουργήσουμε ένα νέο μικρότερο dataset που να προσομοιώνει ένα μήνα μετρήσεων. Για να γίνει αυτό θα παίρνουμε 96 τιμές οι οποίες θα προσομοιώνουν μια ημέρα από συγκεκριμένο κομμάτι των μετρήσεων που έχουμε στη διάθεσή μας. Πιο συγκεκριμένα μετά από μια πρωτογενή ανάλυση των τιμών που έχουμε μπορούν να προκύψουν συνεχόμενες ημέρες από 26/5/2009 μέχρι και 5/6/2009. Για τις υπόλοιπες ημέρες δεν είναι πλήρεις οι τιμές και δεν υπάρχει κοινό επικαλυπτόμενο διάστημα. Επιλέγουμε να εργασθούμε σε ένα διάστημα που να υπάρχει φόρτο εργασίας ώστε να έχουμε μεταβαλλόμενη κατανάλωση.

Για το λόγο αυτό ξεκινάμε την ανάλυση από 12 το μεσημέρι και παίρνουμε από την κάθε ημέρα 96 τιμές για την προσομοίωσή μας. Για πιο σωστά αποτελέσματα παίρνουμε τη συνολική ισχύ και των τριών φάσεων, έτσι δημιουργείται ένας πίνακας που περιέχει timestamp και ισχύ τον οποίο και θα επεξεργαστούμε.

Πιο λεπτομερώς το πρώτο σετ τιμών είναι 96 μετρήσεις ξεκινώντας από τη χρονική στιγμή 12:19:30 στις 26/5/2009, και συνεχίζουμε για 10 ημέρες ξεκινώντας πάντα από τη συγκεκριμένη χρονική στιγμή. Έτσι έχουμε μια προσομοίωση 11 ημερών και θα ελέγξουμε τον αλγόριθμο πάνω σε αυτό το σετ τιμών. Για να μπορέσουμε να ελέγξουμε πλήρως τον αλγόριθμο μαζί με το κομμάτι που αφορά τη συσχέτιση διαφορετικών προφίλ πρέπει να χωρίσουμε αυτό το σετ τιμών στα δύο και να θεωρήσουμε ότι τα δύο κομμάτια αφορούν διαφορετικά προφίλ. Οι μετρήσεις μας καλύπτουν μια περίοδο από ημέρα Τρίτη μέχρι και τη μεθεπόμενη Παρασκευή. Για να έχουμε σωστή συσχέτιση λοιπόν μεταξύ ημερών μπορούμε να χωρίσουμε το σετ τιμών σε δύο κομμάτια τεσσάρων ημερών, Τρίτη έως και Παρασκευή, και να τρέξουμε 2 φορές τον αλγόριθμο θεωρώντας ότι πρόκειται για διαφορετικά προφίλ.

Οι μεταβλητές εισόδου που χρησιμοποιήσαμε για τον αλγόριθμο είναι (1243329571,1243330148,1,75,1.06,4) για την πρώτη τετράδα ημερών και (1243934373,1243934951,1,75,1.06,4) για τη δεύτερη. Το τέσσερα αντιπροσωπεύει τον αριθμό ημερών ενώ οι τιμές 75 για τη γωνία και 1.06 για το λόγο ανόδου επελέγησαν μετά από παρατήρηση των μετρήσεων.

Θα ελέγξουμε τον αλγόριθμο πάνω στους πίνακες που αφορούν το KPI_1(λόγος μέγιστης ισχύος προς μέση) και KPI_2(αριθμός αιχμών). Όλοι οι υπόλοιποι πίνακες αναλύονται με τον ίδιο ακριβώς τρόπο.

Για το πρώτο τετραήμερο οι πίνακες αυτοί είναι οι εξής:

1Α (μέγιστη ισχύς) :

	06:00-09:00	09:00-12:00	12:00-15:00	15:00-18:00	18:00-21:00	21:00-24:00
Ημέρα 1	0	1,4068	1,2386	1,2694	1,2753	1,3302
Ημέρα 2	1,3479	1,3609	0	1,1466	1,1253	1,3012
Ημέρα 3	1,162	1,2283	0	1,32	0	1,213
Ημέρα 4	0	0	1,2244	1,1785	0	1,1936

1B (αριθμός αγμών) :

	06:00-09:00	09:00-12:00	12:00-15:00	15:00-18:00	18:00-21:00	21:00-24:00
Ημέρα 1	0	2	1	2	2	1
Ημέρα 2	1	2	0	1	1	1
Ημέρα 3	1	1	0	2	0	2
Ημέρα 4	0	0	2	1	0	1

Οι αντίστοιχοι πίνακες για το δεύτερο τετραήμερο είναι οι ακόλουθοι:

2A (μέγιστη ισχύς) :

	06:00-09:00	09:00-12:00	12:00-15:00	15:00-18:00	18:00-21:00	21:00-24:00
Ημέρα 1	1,7105	1,2702	1,4762	1,285	0	1,2371
Ημέρα 2	0	1,371	1,3674	1,2546	1,5642	0
Ημέρα 3	1,2699	0	1,2936	1,2454	1,2289	1,544
Ημέρα 4	1,3525	1,3601	1,3775	1,2505	1,278	1,3621

2B (αριθμός αγμών) :

	06:00-09:00	09:00-12:00	12:00-15:00	15:00-18:00	18:00-21:00	21:00-24:00
Ημέρα 1	4	1	2	2	0	1
Ημέρα 2	0	2	1	1	1	0
Ημέρα 3	2	0	1	1	2	1
Ημέρα 4	1	1	2	1	1	1

Βρίσκουμε αρχικά τους τέσσερεις πίνακες συνδιακύμανσης οι οποίοι όπως έχει αναφερθεί προκύπτουν από τον πολλαπλασιασμό του εκάστωτε πίνακα με τον αντίστροφό του ($A \cdot A'$).

Οι πίνακες που προκύπτουν είναι οι ακόλουθοι:

$$\Pi 1 = I A * I A'$$

8,5205	6,5359	5,0172	4,6003
6,5359	7,9429	6,3296	2,9044
5,0172	6,3296	6,0727	3,0036
4,6003	2,9044	3,0036	4,3127

$$\Pi 2 = I B * I B'$$

14	9	8	5
9	8	7	2
8	7	10	4
5	2	4	6

$$\Pi 3 = 2 A * 2 A'$$

9,9001	5,3721	7,5923	9,3665
5,3721	7,7702	5,2536	7,3163
7,5923	5,2536	8,7313	8,7305
9,3665	7,3163	8,7305	10,6291

$$\Pi 4 = 2 B * 2 B'$$

26	6	13	12
6	7	4	6
13	4	11	8
12	6	8	9

Αυτοί οι πίνακες θα χρησιμοποιηθούν για την εξαγωγή συμπερασμάτων για τα δύο προφίλ. Για να γίνει αυτό πρέπει α υπολογίσουμε ιδιοτιμές και ιδιοδιανύσματα των τεσσάρων πινάκων. Αυτά είναι τα ακόλουθα:

Ιδιοτιμές και ιδιοδιανύσματα του Π1:

21,6868	0	0	0
0	3,412	0	0
0	0	1,5175	0
0	0	0	0,2326

0,5807	-0,4459	0,5732	-0,368
0,5675	0,4843	0,208	0,6326
0,4804	0,4049	-0,5346	-0,5652
0,3317	-0,6346	-0,5851	0,3807

Principal component του Π1 :

0,5807	0,5675	0,4804	0,3317
--------	--------	--------	--------

Ιδιοτιμές και ιδιοδιανύσματα του Π2:

28,9272	0	0	0
0	4,7828	0	0
0	0	3,7481	0
0	0	0	0,5419

0,6564	-0,0648	0,6143	0,4331
0,4801	-0,4632	-0,0374	-0,744
0,513	0,044	-0,7856	0,3432
0,2745	0,8828	0,0646	-0,3756

Principal component του Π2:

0,6564	0,4801	0,513	0,2745
--------	--------	-------	--------

Συγκρίνοντας τα 2 PC βλέπουμε ότι δεν υπάρχει σαφής συσχέτιση των 2 προφίλ καθώς δεν παρατηρούνται κοινές τιμές σε αντίστοιχες ημέρες. Μπορούμε όμως να πούμε ότι υπάρχει συσχέτιση μεταξύ της τρίτης ημέρας του πρώτου προφίλ και της δεύτερης ημέρας του δεύτερου προφίλ, όπως φαίνεται και εποπτικά εάν

εξετάσουμε τα 2 διαγράμματα των αντίστοιχων ημερών όπως παρουσιάζονται παρακάτω.

Προσθέτοντας τα 2 pc προκύπτει το εξής διάνυσμα:

1,1644	1,0476	0,9934	0,6062
--------	--------	--------	--------

Από αυτό εξάγουμε τα εξής συμπεράσματα:

1. Τα δύο προφίλ είναι περισσότερο συσχετισμένα την πρώτη ημέρα σε σχέση με τις υπόλοιπες (μιλώντας για αντίστοιχες ημέρες, δηλαδή πρώτη με πρώτη και ούτω καθεξής)
2. Η μικρότερη συσχέτιση παρουσιάζεται την τέταρτη ημέρα
3. Η τιμή 1.1644 θεωρείται σχετικά μικρή για μέγιστη τιμή συσχέτισης πράγμα που σημαίνει ότι τα δύο προφίλ ακόμα και την πρώτη ημέρα δεν παρουσιάζουν τόσο μεγάλη συσχέτιση όπως φαίνεται και από οπτική τους παρατήρηση.

Κάνουμε την ίδια ακριβώς ανάλυση και για τους άλλους 2 πίνακες και προκύπτουν τα εξής:

Ιδιοτιμές και ιδιοδιανύσματα του Π3:

31,5456	0	0	0
0	3,572	0	0
0	0	1,6843	0
0	0	0	0,2288

0,5196	0,4328	0,6362	0,3714
0,4021	-0,8593	0,0771	0,3066
0,486	0,2719	-0,7674	0,3176
0,5762	-0,02	0,0198	-0,8168

Principal component του Π3:

0,5196	0,4021	0,486	0,5762
--------	--------	-------	--------

Ιδιοτιμές και ιδιοδιανύσματα του Π4:

42,4811	0	0	0
0	6,341	0	0
0	0	3,3964	0
0	0	0	0,7815

0,7493	0,486	0,4332	0,1211
0,2482	-0,7812	0,3129	0,4797
0,4477	-0,129	-0,8432	0,2975
0,42	-0,3916	-0,0591	-0,8165

Principal component του Π4:

0,7493	0,2482	0,4477	0,42
--------	--------	--------	------

Ομοίως και εδώ η σύγκριση δε μας δίνει κάποια ισχυρή συσχέτιση σε αντίστοιχες ημέρες. Οι μόνες μέρες που φαίνεται να είναι συσχετισμένες είναι η δεύτερη του πρώτου προφίλ με την τέταρτη του δεύτερου προφίλ.

Προσθέτοντας τα 2 pc προκύπτει το εξής διάνυσμα:

1,2689	0,6503	0,9937	0,9962
--------	--------	--------	--------

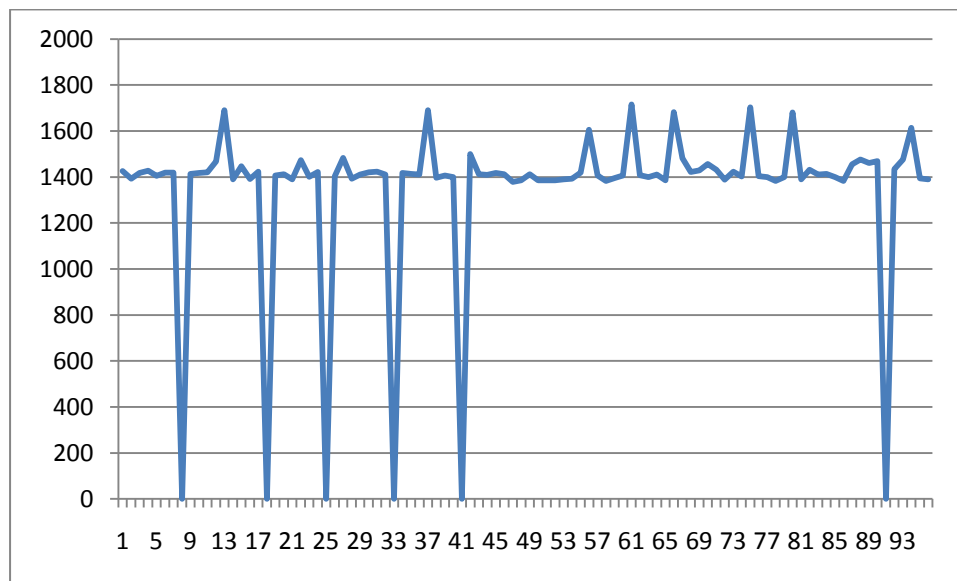
Ομοίως και εδώ προκύπτει ισχυρότερη συσχέτιση την πρώτη αντίστοιχη ημέρα για τα δύο προφίλ.

Αν και η κλίμακα του ελέγχου είναι πολύ μικρή για να γίνει προσπάθεια για pattern matching ή οτιδήποτε άλλο μπορούμε να δούμε ακόμα και από εδώ τον τρόπο με τον οποίο λειτουργεί ο μηχανισμός αυτός. Για παράδειγμα μπορούμε να πούμε ότι για τη μέγιστη ισχύ έχουμε παρόμοιες τιμές στο πρώτο pc για τη δεύτερη και Τρίτη ημέρα, κάτι το οποίο αν παρατηρούνταν για μεγαλύτερο αριθμό ημερών θα μπορούσε να μας βοηθήσει να τοποθετήσουμε τους δύο πελάτες στην ίδια κατηγορία.

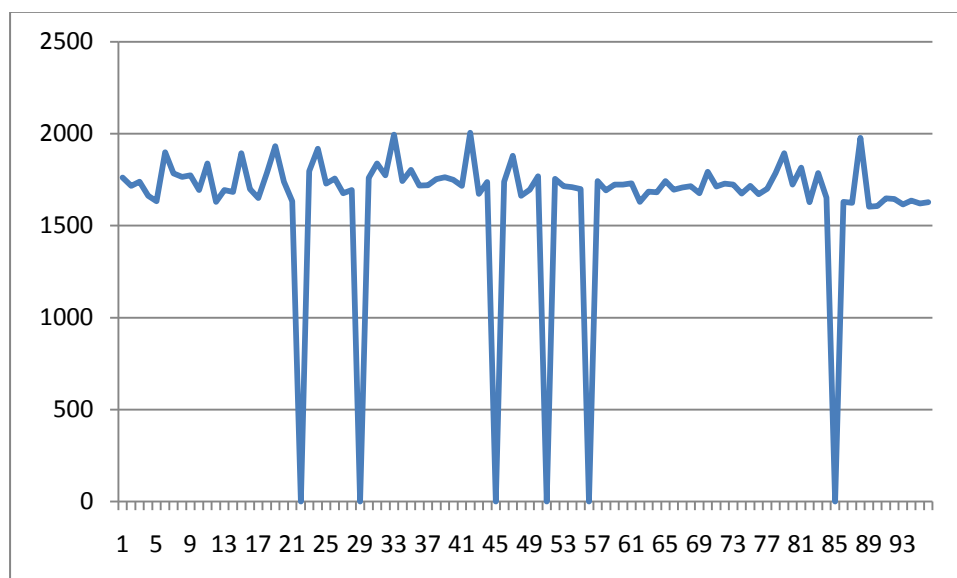
Ομοίως για τους πίνακες του αριθμού αιχμών, βλέπουμε ότι στο πρώτο pc έχουμε μια ολίσθηση τιμών από τη δεύτερη και Τρίτη ημέρα του πρώτου προφίλ στην Τρίτη και τέταρτη ημέρα του δεύτερου προφίλ. Αυτό επίσης σε μεγαλύτερη κλίμακα θα μπορούσε να φανεί πολύ χρήσιμο καθώς αναγνωρίζει μια συμπεριφορά των δύο πελατών η οποία για κάποιο λόγο εμφανίζεται με χρονική καθυστέρηση μιας ημέρας.

Ακολουθούν γραφικές παραστάσεις των προφίλ χωρισμένων ανά ημέρα.

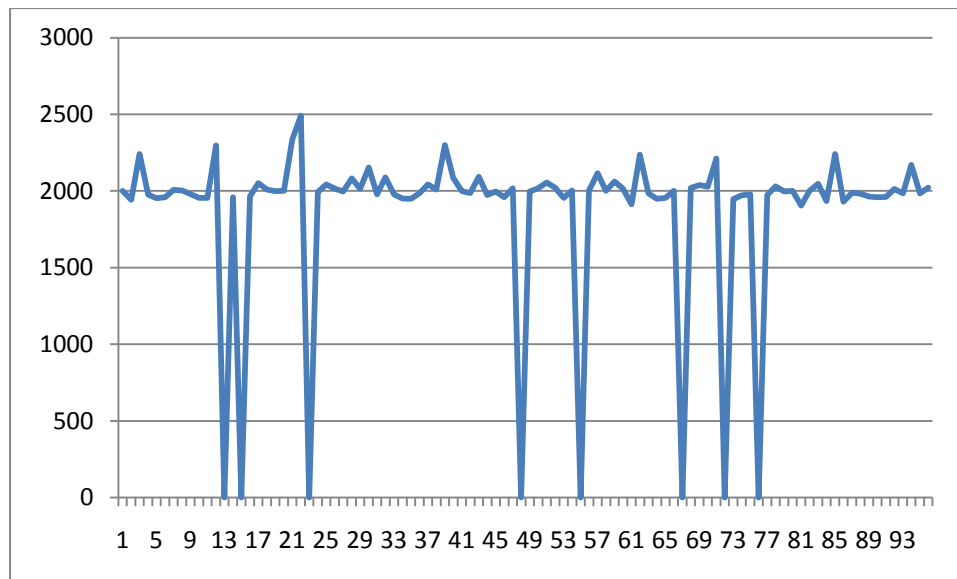
Προφίλ 1: ημέρα 1



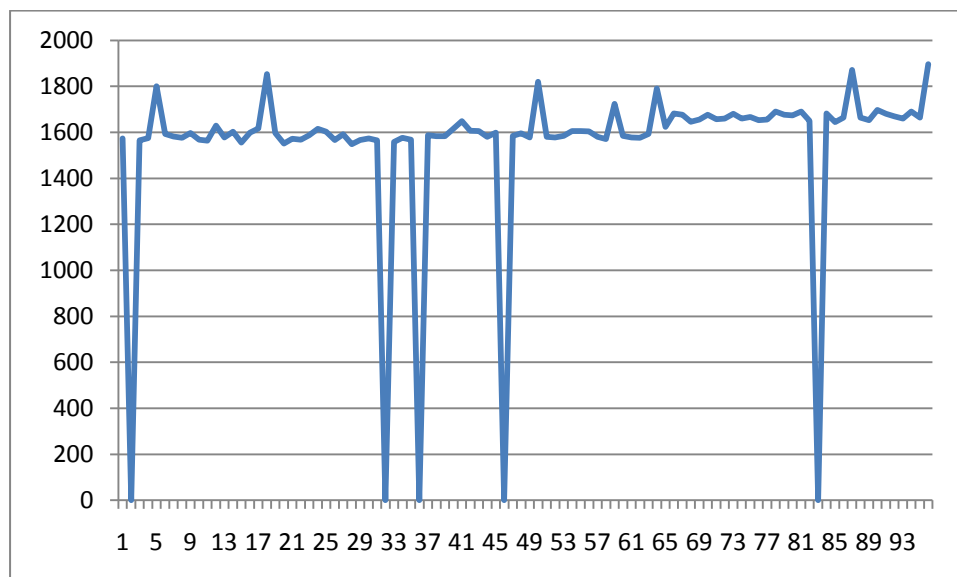
Προφίλ 1: ημέρα 2



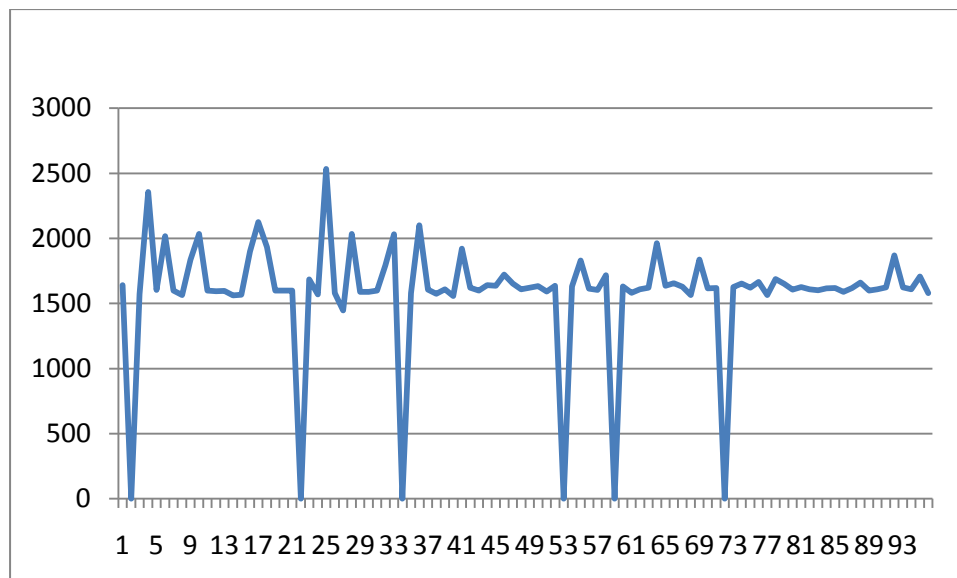
Προφίλ 1: ημέρα 3



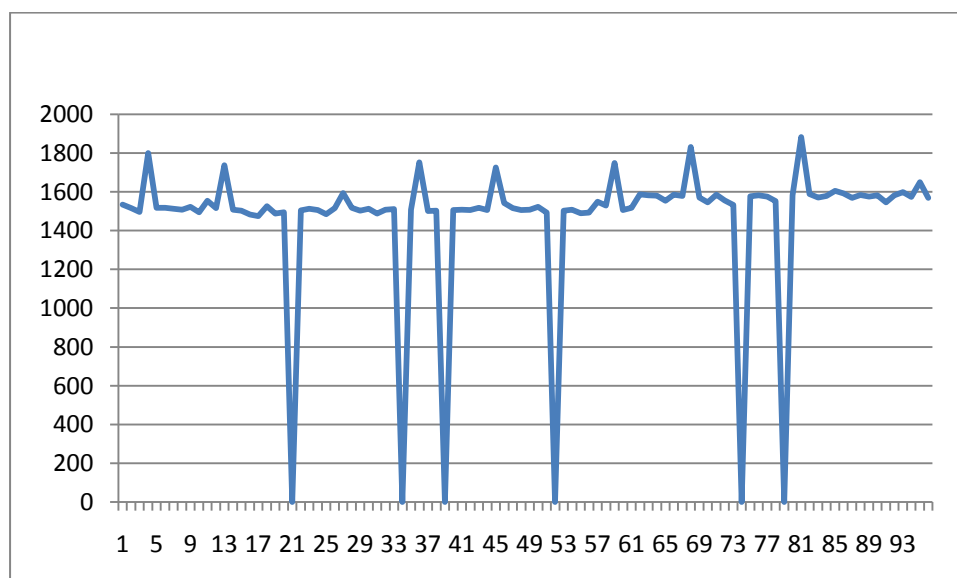
Προφίλ 1: ημέρα 4



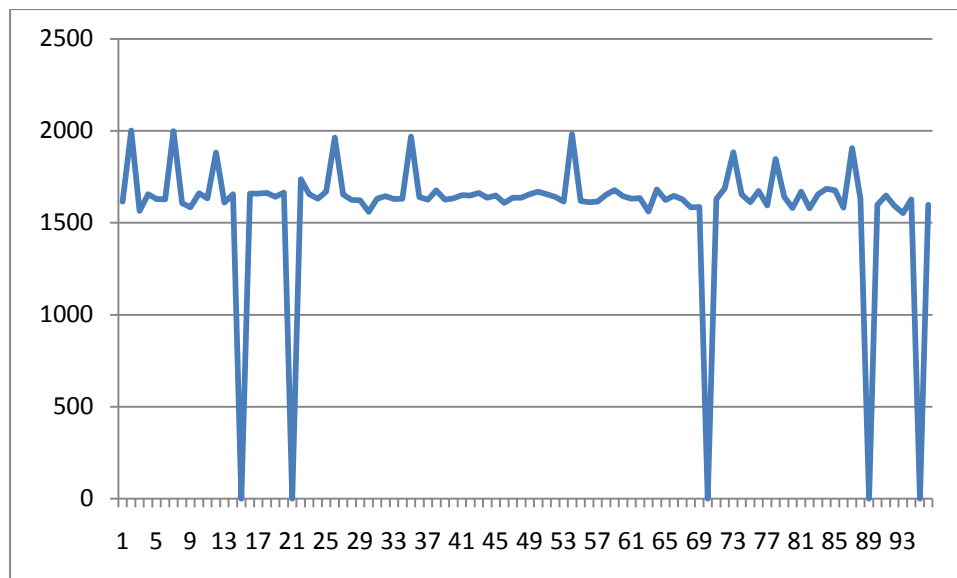
Προφίλ 2: ημέρα 1



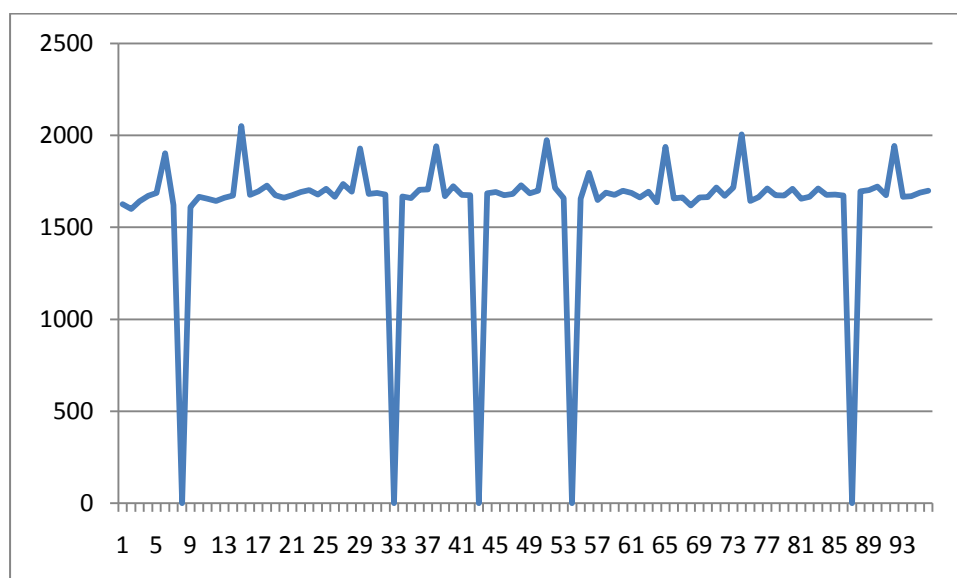
Προφίλ 2: ημέρα 2



Προφίλ 2: ημέρα 3



Προφίλ 2: ημέρα 4



6.2 Test case 2

Σαν επόμενο test case θεωρούμε 2 προφίλ καταναλωτών τα οποία είναι ουσιαστικά μετατοπισμένα το ένα ως προς το άλλο κατά 4 ημέρες. Χρησιμοποιούμε το ίδιο dataset όπως και στην προηγούμενη δοκιμή και ως πρώτο προφίλ παίρνουμε τις πρώτες 7 ημέρες ώστε να έχουμε μια ολόκληρη εβδομάδα και ως δεύτερο προφίλ παίρνουμε άλλες 7 ημέρες ξεκινώντας από την Πέμπτη ημέρα. Τα μετατοπισμένα αυτά προφίλ θα μας δώσουν σημαντικές πληροφορίες για τη μέθοδο που χρησιμοποιούμε. Μετά την εφαρμογή του αλγορίθμου προκύπτουν οι εξής πίνακες (θα εξετάσουμε εδώ τους πίνακες που αφορούν τη μέγιστη αιχμή, τον αριθμό των αιχμών και τα συγκεντρωτικά δεδομένα):

Προφίλ 1: μέγιστη αιχμή

	06:00-09:00	09:00-12:00	12:00-15:00	15:00-18:00	18:00-21:00	21:00-24:00
Ημέρα 1	0	1,273097575	1,195893033	1,394806277	1,436229587	1,266510775
Ημέρα 2	0	1,306632325	1,115923567	1,189120517	1,173403767	1,187346739
Ημέρα 3	1,228904525	1,241742362	0	1,050175713	1,125334734	1,17253991
Ημέρα 4	1,066732698	1,230837869	0	1,317820324	0	1,143197077
Ημέρα 5	0	0	1,124845488	1,079132356	0	1,193620415
Ημέρα 6	1,178291194	1,265594397	1,177415805	1,283259522	1,161619382	1,166883963
Ημέρα 7	0	0	0	0	0	0

Προφίλ 1: αριθμός αιχμών

	06:00-09:00	09:00-12:00	12:00-15:00	15:00-18:00	18:00-21:00	21:00-24:00
Ημέρα 1	0	2	1	1	1	1
Ημέρα 2	0	2	1	2	2	1
Ημέρα 3	1	2	0	1	1	1
Ημέρα 4	1	1	0	2	0	2
Ημέρα 5	0	0	2	1	0	1
Ημέρα 6	3	3	2	3	3	3
Ημέρα 7	0	0	0	0	0	0

Προφίλ 1 : Συγκεντρωτικά

	KPI_1	KPI_2	KPI_3	KPI_4
Ημέρα 1	4	1	4	0
Ημέρα 2	1	1	1	5
Ημέρα 3	1	1	0	5
Ημέρα 4	3	3	3	4
Ημέρα 5	5	2	5	3
Ημέρα 6	3	0	3	3
Ημέρα 7	0	0	0	4

Προφίλ 2: μέγιστη αιχμή

	06:00-09:00	09:00-12:00	12:00-15:00	15:00-18:00	18:00-21:00	21:00-24:00
Ημέρα 1	0	0	1,124903425	1,084687595	0	1,097628942
Ημέρα 2	1,178291194	1,265594397	1,177415805	1,283259522	1,161619382	1,166883963
Ημέρα 3	0	0	0	0	0	0
Ημέρα 4	1,529743332	1,166759508	1,3304253	1,284966179	0	1,138015019
Ημέρα 5	0	1,248410671	1,245162056	1,152297042	1,410001873	0
Ημέρα 6	1,167029057	0	1,187484405	1,127975858	1,12574731	1,393784278
Ημέρα 7	1,229984051	1,235453244	1,253664991	1,147367902	1,173486807	1,239113096

Προφίλ 2: αριθμός αιχμών

	06:00-09:00	09:00-12:00	12:00-15:00	15:00-18:00	18:00-21:00	21:00-24:00
Ημέρα 1	0	0	2	1	0	1
Ημέρα 2	3	3	2	3	3	3
Ημέρα 3	0	0	0	0	0	0
Ημέρα 4	4	1	2	2	0	1
Ημέρα 5	0	2	1	1	1	0
Ημέρα 6	2	0	1	1	2	1
Ημέρα 7	1	1	2	1	1	1

Προφίλ 2: Συγκεντρωτικά

	KPI_1	KPI_2	KPI_3	KPI_4
Ημέρα 1	2	2	5	5
Ημέρα 2	3	0	3	3
Ημέρα 3	0	0	0	4
Ημέρα 4	0	0	0	2
Ημέρα 5	4	1	4	0
Ημέρα 6	5	0	5	3
Ημέρα 7	2	2	5	1

Από μια απλή οπτική παρατήρηση των πινάκων φαίνεται εύκολα η χρονική μετατόπιση που έχει αναφερθεί. Για να συνεχίσουμε με την ανάλυσή μας βρίσκουμε ξανά τους πίνακες συνδιασποράς και στη συνέχεια ιδιοτιμές και ιδιοδιανύσματα δηλαδή εφαρμόζουμε τη μέθοδο PCA για να βρούμε τα principal components.

Πρώτο PC του πίνακα 1 για το προφίλ 1:

-0,4771	-0,4348	-0,4033	-0,3446	-0,2504	-0,4891	0
---------	---------	---------	---------	---------	---------	---

Κάνοντας ακριβώς το ίδιο για τον πρώτο πίνακα του προφίλ 2 παίρνουμε το εξής πρώτο PC:

-0,2458	-0,4801	0	-0,4423	-0,3365	-0,4088	-0,4827
---------	---------	---	---------	---------	---------	---------

Χωρίς καμία περαιτέρω επεξεργασία η μετατόπιση που έχει ήδη αναφερθεί παραπάνω είναι εμφανής στα 2 αυτά PCs καθώς οι 3 τελευταίες τιμές του πρώτου είναι σχεδόν ίδιες με τις 3 πρώτες του δεύτερου. Αυτό είναι δείγμα ομοιότητας των δύο προφίλ, ενώ αν μιλούσαμε για το ίδιο προφίλ σε διαφορετικές εβδομάδες θα είχαμε σαφή ένδειξη περιοδικότητας. Στα ίδια ακριβώς συμπεράσματα καταλήγουμε και από την ανάλυση των άλλων πινάκων.

Προσθέτοντας τα δύο PC προκύπτει το επόμενο διάνυσμα από το οποίο προκύπτει ότι τη δεύτερη ημέρα έχουμε την υψηλότερη συσχέτιση μεταξύ των ημερών. Παρόλα αυτά βλέπουμε ότι η συσχέτιση δεν είναι ισχυρή καθώς η τιμή είναι μικρότερη της μονάδας, και παραπλήσια με άλλες τιμές σε διαφορετικές ημέρες, όπως για παράδειγμα στην έκτη.

-0,7229	-0,9149	-0,4033	-0,7869	-0,5869	-0,8979	-0,4827
---------	---------	---------	---------	---------	---------	---------

Ως προς τους πίνακες του αριθμού αιχμών έχουμε τα εξής 2 PCs:

-0,2838	-0,3804	-0,288	-0,291	-0,1619	-0,7622	0
---------	---------	--------	--------	---------	---------	---

Και

-0,1689	-0,7233	0	-0,4731	-0,199	-0,3101	-0,298
---------	---------	---	---------	--------	---------	--------

Ομοίως και εδώ παρατηρούμε εύκολα τη μετατόπιση ενώ μετά από πρόσθεση έχουμε

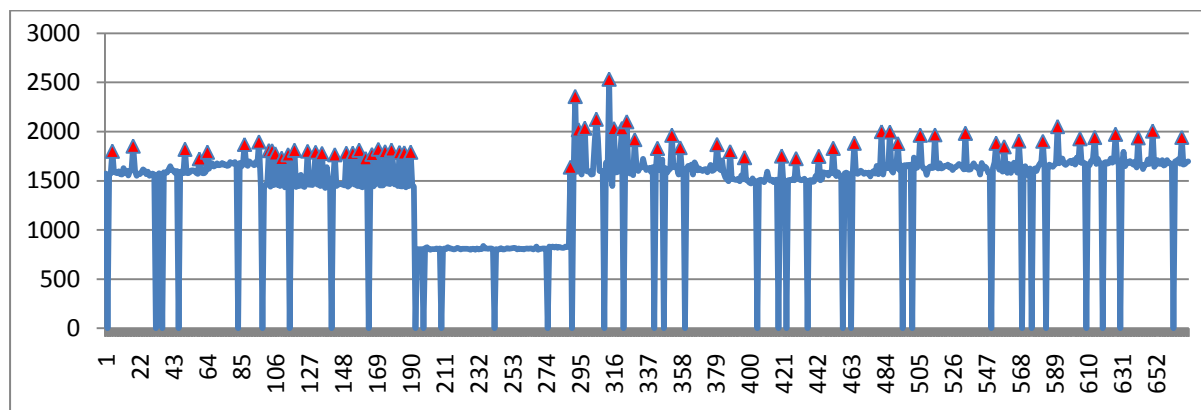
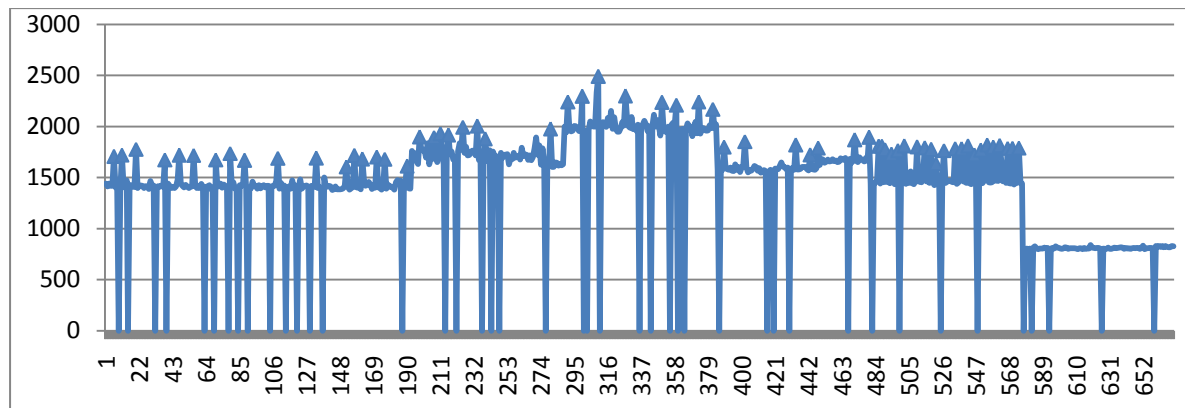
-0,4527	-1,1037	-0,288	-0,7641	-0,3609	-1,0723	-0,298
---------	---------	--------	---------	---------	---------	--------

Σε αυτό το διάνυσμα προκύπτει μεγαλύτερη συσχέτιση και πάλι ως προς τη δεύτερη ημέρα.

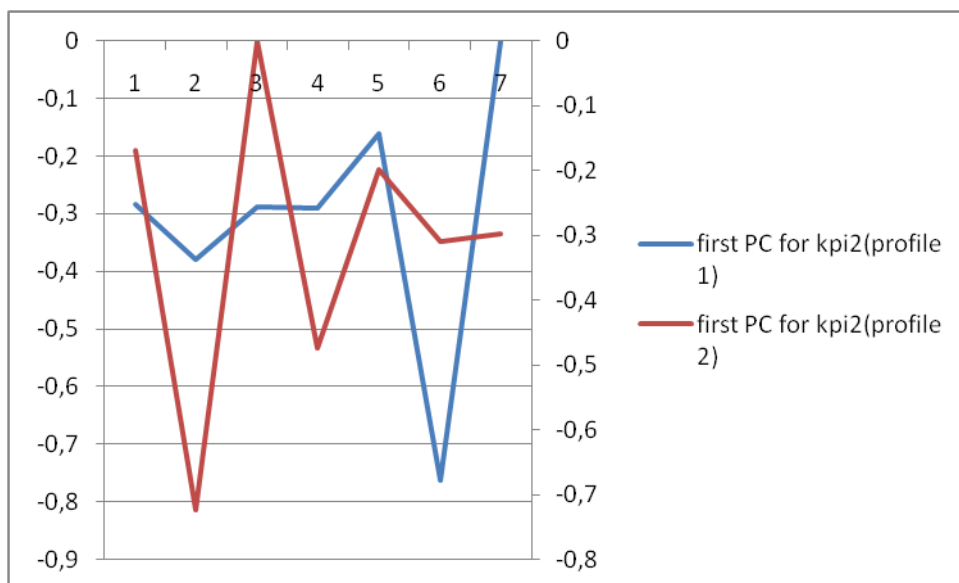
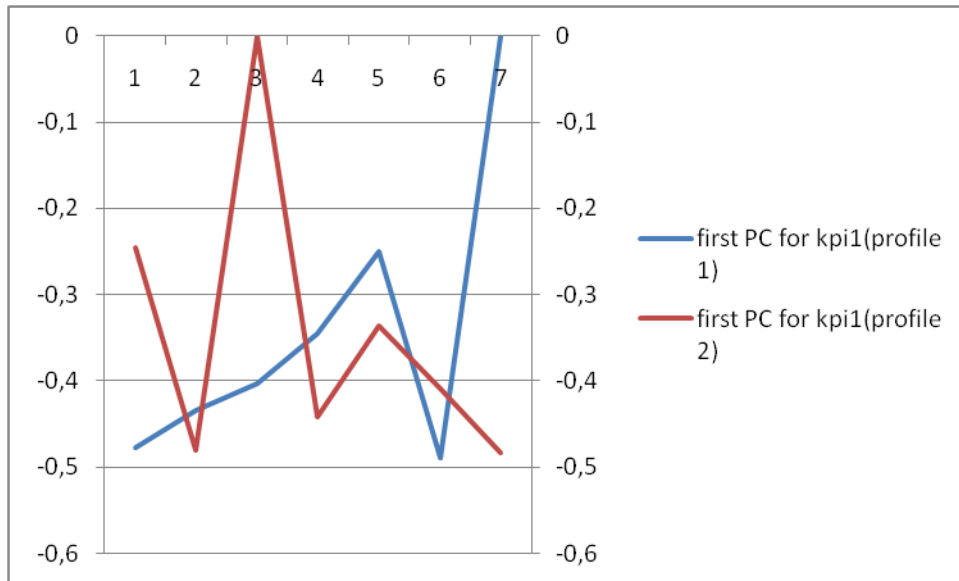
Αν θέλουμε επομένως μια συνολική εικόνα της συσχέτισης των 2 προφίλ μπορούμε να καταρτίσουμε ένα κοινό πίνακα ο οποίος να περιέχει τα αποτελέσματα

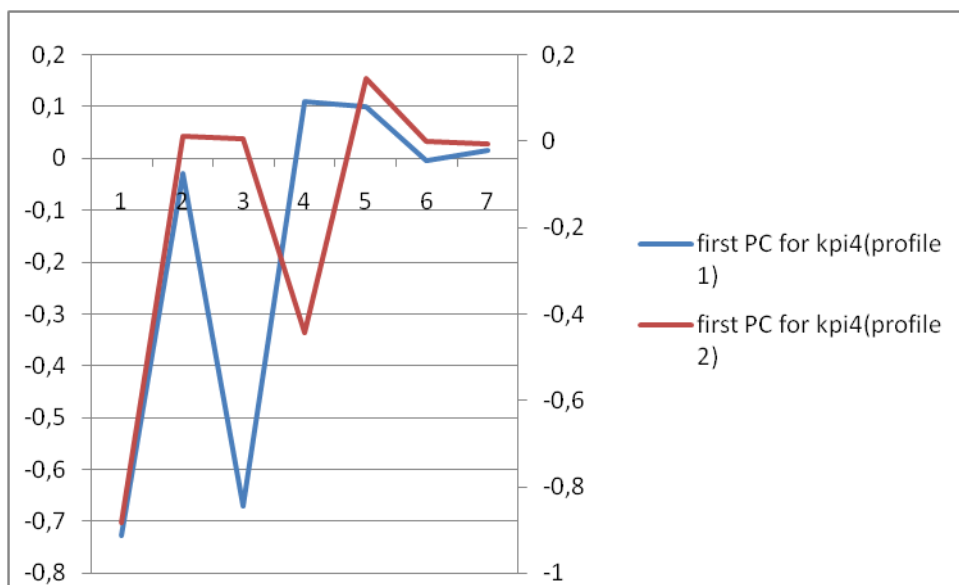
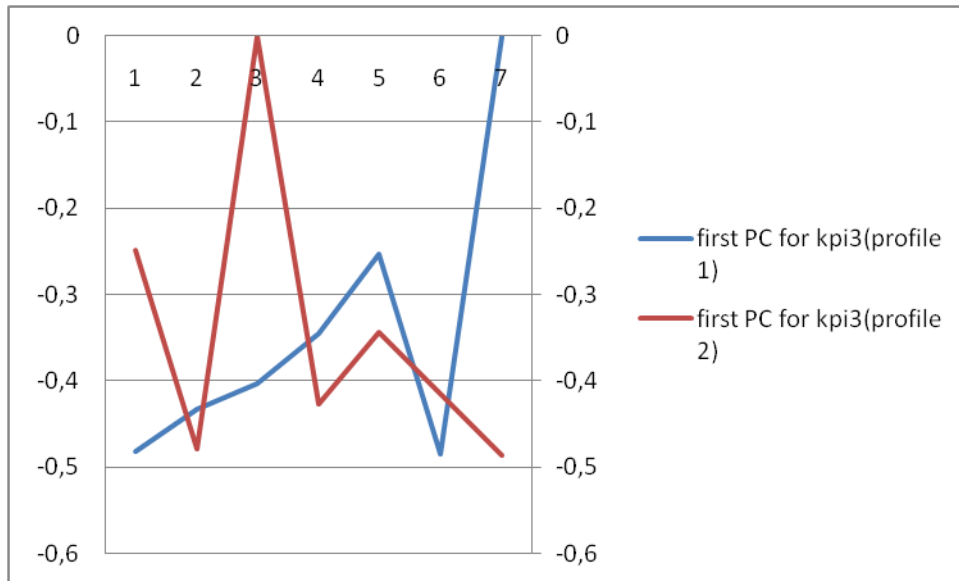
της πρόσθεσης των πρώτων PCs. Θα έχουμε δηλαδή ένα πίνακα με 5 γραμμές (όσοι και οι πίνακες των KPIs) και από αυτό τον πίνακα ελέγχοντας ανά στήλες μπορούμε να δούμε την τις ημέρες με τη μεγαλύτερη συνολική συσχέτιση απλά ελέγχοντας τις τιμές και βρίσκοντας τις ημέρες που παρουσιάζουν τα περισσότερα μέγιστα ανά στήλη. Στο προηγούμενο παράδειγμα η ημέρα με τη μεγαλύτερη συσχέτιση συνδυάζοντας τα PCs των πινάκων μέγιστων αιχμών και αριθμού αιχμών είναι η δεύτερη. Με αυτό το συνολικό πίνακα μπορούμε να δούμε ολικές συσχετίσεις και περιοδικότητα. Για παράδειγμα μπορούμε να τρέξουμε τον αλγόριθμο για έναν ολόκληρο μήνα για 2 προφίλ, και έτσι να έχουμε ένα πίνακα 5*30 που θα μας δίνει τις επιθυμητές πληροφορίες.

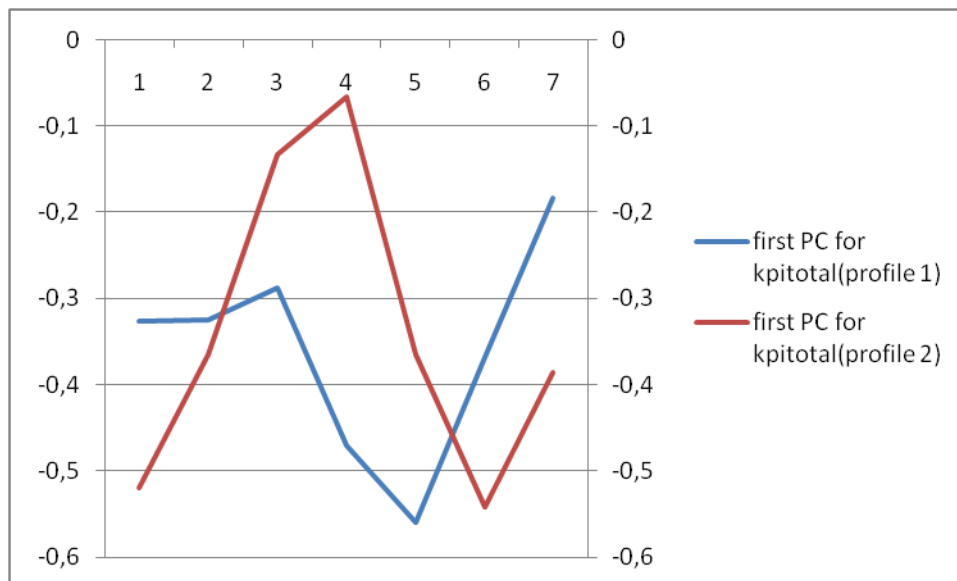
Οι ολικές καμπύλες στις οποίες έγινε η ανάλυση ακολουθούν:



Αν δούμε τα πρώτα PC των προφίλ για το κάθε KPI σε συγκριτικά διαγράμματα βλέπουμε απευθείας την ομοιότητά τους και άρα τη χρησιμότητα της μεθόδου την οποία χρησιμοποιήσαμε. Ακολουθούν τα διαγράμματα αυτά:







Σε όλα αυτά τα διαγράμματα των πρώτων PCs για τον κάθε έναν από τους πίνακες των 2 προφίλ φαίνεται καθαρά ότι η καμπύλη του προφίλ 2 συμπίπτει πλήρως με την καμπύλη του προφίλ 1 αν γίνει μια μετατόπιση της προς τα δεξιά κατά 3 μονάδες. Αυτό μας υποδεικνύει σαφώς ότι πρόκειται για 2 προφίλ τα οποία έχουν ίδια συμπεριφορά σε διαφορετικές όμως ημέρες, οπότε θα μπορούσαν να τοποθετηθούν στην ίδια κατηγορία πελατών αν μπορούσαν να μετακινήσουν τα προφίλ τους χρονικά ώστε να συμπέσουν. Ένα άλλο πολύ χρήσιμο συμπέρασμα αφορά την περιοδικότητα ενός προφίλ. Ελέγχοντας τα διαγράμματα διαφορετικών εβδομάδων ενός προφίλ με περιοδική συμπεριφορά θα δούμε πανομοιότυπες αυξομειώσεις στα ιδιοδιανύσματά τους. Έτσι παρατηρώντας απλά τα διαγράμματα μπορούμε να αντιληφθούμε σημαντικά χαρακτηριστικά ενός προφίλ όπως περιοδικότητα, επαναληπτικότητα και την ύπαρξη patterns ανά διαστήματα.

Ένας άλλος τρόπος για να παρατηρούμε συσχετίσεις δύο προφίλ είναι όπως έχει ήδη αναφερθεί η δημιουργία ενός πίνακα που θα περιέχει 5 διανύσματα τα οποία θα προέρχονται από την πρόσθεση των πρώτων PCs του κάθε πίνακα για το κάθε προφίλ. Για τα προφίλ τα οποία παρουσιάζονται στα διαγράμματα πιο πάνω ο πίνακας αυτός έχει ως εξής:

kpi1	kpi2	kpi3	kpi4	kpitotal
-0,7229	-0,4526	-0,2325	-1,6111	-0,8461
-0,9150	-1,1037	0,0472	-0,0168	-0,6900
-0,4033	-0,2880	-0,4037	-0,6627	-0,4210
-0,7869	-0,7641	0,0824	-0,3334	-0,5376
-0,5870	-0,3609	0,0907	0,2453	-0,9243
-0,8979	-1,0723	-0,0712	-0,0040	-0,9097
-0,4827	-0,2980	0,4857	0,0104	-0,5699

Βλέπουμε ότι τα περισσότερα μέγιστα ανά στήλη βρίσκονται στη δεύτερη γραμμή, κάτι το οποίο δείχνει ότι τα 2 προφίλ είναι πλέον συσχετισμένα τη δεύτερη ημέρα. Με τον τρόπο αυτό μπορούμε να έχουμε μια εποπτική εικόνα της συσχέτισης των 2 προφίλ σε σχέση με όλα τα KPIs.

Αν η παραπάνω ανάλυση γίνει για περισσότερα προφίλ πελατών τότε μπορούμε να καταρτίσουμε πίνακες ανά κατηγορία KPI. Έστω για παράδειγμα ότι έχουμε 10 προφίλ διαφορετικών πελατών και θέλουμε να τα συσχετίσουμε ως προς τη μέγιστη αιχμή του καθενός (KPI_1). Σχηματίζουμε ένα πίνακα με τόσες γραμμές όσα και τα προφίλ, και με κάθε γραμμή να περιέχει το πρώτο principal component του πίνακα KPI_1 του εκάστοτε προφίλ. Ελέγχοντας τον πίνακα αυτό ανά γραμμές μπορούμε να ψάξουμε για επαναλαμβανόμενες τιμές, είτε μεμονωμένες είτε σε ομάδες. Αυτή η επαναληπτικότητα υποδεικνύει περιοδικότητα του συγκεκριμένου προφίλ στις συγκεκριμένες ημέρες στις οποίες εμφανίζονται οι επαναλαμβανόμενες τιμές.

Αναλύοντας τώρα τον ίδιο πίνακα ανά στήλες αυτή τη φορά μπορούμε να εντοπίσουμε συσχετισμένα προφίλ. Αυτό γίνεται ως εξής: υψηλά συσχετισμένοι πελάτες θα έχουν σχεδόν πανομοιότυπες τιμές στα principal components τους στις ημέρες που είναι συσχετισμένες. Εφόσον κάθε γραμμή αντιστοιχεί σε διαφορετικό πελάτη ελέγχοντας μια στήλη του πίνακα αναμένουμε κοινές τιμές σε στήλες που αντιστοιχούν σε ημέρες με κοινές καμπύλες κατανάλωσης. Έτσι όσες τιμές παρουσιάζουν μικρή απόκλιση μεταξύ τους τόσα είναι και τα προφίλ που είναι συσχετισμένα μεταξύ τους.

7

Συμπεράσματα και μελλοντικές εφαρμογές μελέτης

7 Συμπεράσματα και μελλοντικές εφαρμογές μελέτης

Γίνεται φανερό από όλα τα παραπάνω ότι η ανάλυση δεδομένων ενέργειας με αλγοριθμικό τρόπο μας δίνει πολύ σημαντικά συμπεράσματα σχετικά με τη συνολική κατανάλωση του ενεργειακού συστήματος.

Η ανάλυση αιχμών, που περιλαμβάνει την ανίχνευση και συσχέτισή τους, είναι πολύ σημαντικό εργαλείο για τις εταιρείες παροχής ηλεκτρικής ενέργειας καθώς δίνει τη δυνατότητα για καλύτερη διαχείριση ζήτησης από την πλευρά τους. Γνωρίζοντας τη συμπεριφορά καταναλωτών ως προς τα μέγιστα φορτία τους, το σημαντικότερο σημείο δηλαδή για το ενεργειακό δίκτυο, η εταιρεία μπορεί να αποφεύγει κινδύνους κατάρρευσης λόγω επικαλυπτόμενων αιχμών την ίδια χρονική στιγμή.

Η συσχέτιση ενεργειακών πελατών με βάση πραγματικά δεδομένα και όχι προβλέψεις κατανάλωσης επιτρέπει τη δημιουργία σαφώς διαχωρισμένων, μη επικαλυπτόμενων κλάσεων πελατών. Η κάθε κλάση περιλαμβάνει πελάτες οι οποίοι παρουσιάζουν συγγενή ενεργειακή συμπεριφορά. Η συμπεριφορά αυτή συνίσταται στη χρονική συσχέτιση αιχμών, δηλαδή την παρουσία αιχμών σε κοινά σημεία της ημέρας, η και σε άλλα χαρακτηριστικά που μας ενδιαφέρουν, όπως τον αριθμό των αιχμών, το ύψος της κατανάλωσης και το πόσο γρήγορα ή μη φτάνουμε σε αυτό το μέγιστο της κατανάλωσης. Όλοι αυτοί οι παράγοντες είναι βασικοί για τη συμπεριφορά του συστήματος ειδικά στις ώρες αιχμής.

Η συσταδοποίηση πελατών σε ενεργειακές κλάσεις οδηγεί σε προσωποποιημένες τιμολογιακές πολιτικές, προσαρμοσμένες πλήρως στα δεδομένα του κάθε πελάτη. Ακόμα πιο σημαντικό είναι το γεγονός ότι η πραγματικού χρόνου λειτουργία του πληροφοριακού συστήματος επιτρέπει και την πραγματικού χρόνου διαχείριση ζήτησης. Η εταιρεία μπορεί πλέον να προσφέρει κίνητρα σε πελάτες, κυρίως με τη μορφή μειωμένων χρεώσεων σε συγκεκριμένες χρονικές στιγμές, ώστε να τροποποιεί την κατανάλωση πελατών, με απώτερο στόχο την τροποποίηση της συνολικής καμπύλης κατανάλωσης του όλου συστήματος, ώστε αυτό να λειτουργεί όσο πιο μακριά γίνεται από το σημείο κινδύνου. Τα κίνητρα αυτά ανακοινώνονται στους ενδιαφερόμενους πελάτες είτε μέσω των αυτόματων μετρητών σε πραγματικό χρόνο, είτε στους μηνιαίους λογαριασμούς. Έτσι μπορούν και οι πελάτες να τα εκμεταλλευτούν πλήρως προς όφελός τους.

Ένα άλλο πλεονέκτημα της μεθόδου που περιγράφηκε είναι η δημιουργία χαρακτηριστικών προφίλ πελατών. Με την αποθήκευση δεδομένων του κάθε πελάτη για μεγάλα χρονικά διαστήματα έχουμε μια σαφή εικόνα της συμπεριφοράς του. Έτσι μπορούμε να δημιουργήσουμε μια μέση κατανάλωση από όλες τις μετρήσεις. Αυτή η μέση συνολική κατανάλωση του πελάτη είναι το ενεργειακό προφίλ του, και μπορεί να χρησιμοποιηθεί για την πρόβλεψη της συμπεριφοράς του σε μελλοντικές χρονικές στιγμές.

Τα ενεργειακά πληροφοριακά συστήματα εξελίσσονται συνεχώς και νέες μέθοδοι ανάλυσης προστίθενται σε αυτά. Στο μέλλον και με την προϋπόθεση ότι έχουν τοποθετηθεί μετρητές στην πλειοψηφία των πελατών θα μπορούσε να δημιουργηθεί ένα δίκτυο ομότιμων μετρητών που να επεξεργάζεται τα δεδομένα απευθείας χωρίς την αποθήκευσή τους σε βάσεις δεδομένων. Με τις αυξημένες δυνατότητες που μας προσφέρει το διαδίκτυο οι μετρήσεις θα αποστέλλονται σε διακομιστές οι οποίοι τα επεξεργάζονται άμεσα και επιλέγουν αυτόματα τους πιο συσχετισμένους πελάτες, αποστέλλοντας τους πληροφορίες για τις χρεώσεις τους.

Ακόμη το σύστημα επιτρέπει τη δημιουργία ολοκληρωμένων προγραμμάτων οικιακής χρήσης, μέσω των οποίων θα μπορεί ο πελάτης μόνος του να ελέγχει ανά πάσα στιγμή την κατανάλωσή του και να παρεμβαίνει σε αυτή, μέσα από ένα απλό user interface. Έτσι πλέον η επικοινωνία πελάτη – εταιρείας καθίσταται αμφίδρομη και άρα πιο αποδοτική.

Σε τελικό στάδιο ο στόχος είναι η ύπαρξη ενός πλήρως αυτοματοποιημένου συστήματος διαχείρισης ενέργειας, το οποίο θα μπορεί μέσα από συνεχή ανάλυση παλαιών και νέων μετρήσεων να διαχειρίζεται πλήρως και με τον πιο αποδοτικό τρόπο την παραγωγή, μεταφορά, διανομή και κατανάλωση της ενέργειας.

8 ΑΝΑΦΟΡΕΣ

- [1] Akihiro Koretsune, Shingo Aoki, Takefumi Konzo, Hiroshi Tsuji , Osaka Prefecture University 1-1 Gakuencho , Sakai, Japan, Shinji Shimano, Eiji Mimura , Energy Use R&D Center, Kansai Electronic Power Co. Inc. 3-1 1-20, Nakuoji, Amagasaki, Japan , DEA – based Data Mining for Energy Consumption, IEEE (2005) Volume 1, pages 1073 - 1076
- [2] A. H. Nizar, Z. Y. Dong, J. H. Zhao, Load Profiling and Data Mining Techniques in Electricity Deregulated Market, IEEE 2006
- [3] Jose A. Dominguez Navarro, Jose L. Bernal-Agustin, Rodolfo Dufo-Lopez, Data mining methodology for disaggregation of load demand, Electric Power Systems Research 79 (2009) 1393–1399
- [4] J. K. Lin, S. K. Tso, H. K. Ho, C. M. Mak, K. M. Yung, Y. K. Ho, Study of climatic effects on peak load and regional similarity of load profiles following disturbances based on data mining, Electrical Power and Energy Systems 28 (2006) 177–185
- [5] D. A. Hull and T. A. Reddy, Center for Energy and Environmental Studies, Princeton University, Princeton, USA, A procedure to group residential air-conditioning load profiles during the hottest days in summer, Energy Vol. 15, No. 12, pp 1085-1097, 1990, Pergamon Press
- [6] Gianfranco Chicco, Roberto Napoli, Petru Postolache, Mircea Scutariu, Cornel Toader, Electric Energy Customer Characterization for Developing Dedicated Market Strategies, 2001 IEEE Porto Power Tech Conference, 10-13 September, Porto, Portugal
- [7] Lindsay I. Smith , A tutorial on Principal Component Analysis, February 26, 2002
- [8] http://en.wikipedia.org/wiki/Demand_response
- [9] Κοινοτική Οδηγία, 32006L0032, 2006/32/ΕΚ Επίσημη Εφημερίδα αριθ. L 114 της 27/04/2006 σ. 0064 – 0085
- [10] http://en.wikipedia.org/wiki/Energy_management_system
- [11] http://en.wikipedia.org/wiki/Smart_grid
- [12] http://en.wikipedia.org/wiki/Smart_meter
- [13] http://en.wikipedia.org/wiki/Principal_component_analysis

[14] European Comision , European SmartGrids Technology Platform
http://ec.europa.eu/research/energy/pdf/smartgrids_en.pdf

[15] Vassilis Nikolopoulos, PhD Candidate NTUA, Theoretical and Technical Analysis of E.M.I.R. System Energy Management and Intelligent Reporting, <http://www.intelen.gr>

[16] Intelen Group, Smart Profiling Impact στη διαχείριση ζήτησης (DSM),
http://www.plugme.eu/portal/index.php?option=com_content&view=article&id=139:dsr3&catid=65:case-studies&Itemid=15

[17] Βασίλειος Νικολόπουλος, Ισίδωρος Κοκκός, Δέσποινα Κουκουλά, Προχωρημένα Ενεργειακά Πληροφοριακά Συστήματα – Εφαρμογές σε έξυπνο web profiling ενεργειακών πελατών, Δελτίο ΠΣΔΜ-Η Φεβρουάριος 2009, σελ. 24-36

Παρόρσημα

A

9 Παράρτημα

9.1 Καθορισμός μεταβλητών εισόδου:

\$start_time: αρχικό timestamp για προσπέλαση δεδομένων

\$end_time: τελικό timestamp

\$data_interval: διάστημα μεταξύ διαδοχικών μετρήσεων που χρησιμοποιούμε (παίρνουμε από τη βάση δεδομένων στοιχεία ανά \$data_interval μετρήσεις)

\$days_count: αριθμός ημερών για τον οποίο θέλουμε να τρέξει ο αλγόριθμος

\$deg_threshold: κατώφλι για την τιμή του ρυθμού ανόδου, καθόδου της γραφικής παράστασης (όριο για την ανίχνευση ή όχι αιχμής)

\$val_threshold: κατώφλι για την ποσοστιαία διαφορά της τιμής μιας κορυφής από την προηγούμενη τιμή (όριο για την ανίχνευση ή όχι τιμής)

\$timestamp_threshold: κατώφλι για το χρονικό παράθυρο μέσα στο οποίο δύο αιχμές διαφορετικών ημερών μπορούν να συσχετιστούν χρονικά

\$customer_ID: κωδικός του συγκεκριμένου πελάτη

9.2 Περιγραφή συναρτήσεων:

connectToDB: σύνδεση στη βάση δεδομένων με χρήση του αρχείου config

getdata: ανάκτηση από τη βάση δεδομένων των επιθυμητών στοιχείων

algo: ανίχνευση αιχμών σε καθορισμένο διάστημα

week: εύρεση αιχμών για καθορισμένο αριθμό ημερών και αποθήκευσή τους σε κοινό πίνακα

associate_peaks: χρονική συσχέτιση αιχμών μέσα σε δοθέν διάστημα ημερών και εύρεση μέσης τιμής τους

Pavg: εύρεση μέσης τιμής των μετρήσεων ενός τριώρου

KPI_1: εύρεση Pmax/Pavg για 6 τριώρα μέσα σε κάθε ημέρα για καθορισμένο αριθμό ημερών

KPI_2: εύρεση αριθμού αιχμών για 6 τριώρα μέσα σε κάθε ημέρα για καθορισμένο αριθμό ημερών

KPI_3: εύρεση P_{pavg}/P_{avg} (P_{pavg} η μέση τιμή των αιχμών του τρίωρου) για 6 τρίωρα μέσα σε κάθε ημέρα για καθορισμένο αριθμό ημερών

KPI_4: εύρεση $(P_{end}-P_{start})/P_{avg}$ (τελική τιμή – αρχική τιμή τρίωρου /μέση τιμή) δηλαδή σταθμισμένου ρυθμού ανόδου ή καθόδου γραφικής παράστασης για 6 τρίωρα μέσα σε κάθε ημέρα για καθορισμένο αριθμό ημερών

KPI_total: εύρεση για το κάθε KPI του τρίωρου στο οποίο εμφανίζεται η μέγιστη τιμή και αποθήκευση των δεικτών αυτών σε κοινό πίνακα

output: έξοδος αλγορίθμων για KPIs σε αρχεία dat (tab delimited format) για είσοδο σε matlab

9.3 Κώδικας PHP

Config.php

```
<?php  
  
//configuration file  
  
$dbhost = 'localhost';  
  
$dbuser = 'localhost';  
  
$dbpass = '*****';  
  
$dbname = 'testdb';  
  
?>
```

connectToDb.php

```
<?php

//sundesh sth vash dedomenwn

function connectToDB() {
    include("config.php");
    global $mysqli;
    $mysqli = mysqli_connect($dbhost, $dbuser, $dbpass, $dbname);
    if (mysqli_connect_errno()) {
        echo("fault");
        echo "<br/>";
        printf("Connect failed: %s\n", mysqli_connect_error());
        exit();
    }
    return $mysqli;
}

?>
```

Getdata.php

```
<?php

//anakthsh sygekrimenwn dedomenwn apo th vash

function getdata($initial_timestamp,$final_timestamp,$interval) {
    include_once("connectToDB.php");
    $mysqli=connectToDB();
```

```

$count=0;

$i=0;

$get_data_sql="SELECT timestamp,data FROM live WHERE timestamp
BETWEEN $initial_timestamp AND ($final_timestamp - 1) ORDER BY timestamp";

$get_data_res=mysqli_query($mysqli,$get_data_sql) or
die(mysqli_error($mysqli));

if (mysqli_num_rows($get_data_res)<1) {

    echo("No records selected");

    echo "<br/>";

}

else {

    while ($data = mysqli_fetch_array($get_data_res)) {

        if (($i)%($interval)==0) {

            $data_array[$count++]=array('timestamp'=>$data["timestamp"],
'KWh'=>$data["data"]);

            }

            $i++;

        }

    }

    mysqli_close($mysqli);

    return $data_array;

}

?>

```

Algo.php

```
<?php
```

```
//evresh koryfwn se sygekrimmeno dothen diasthma
```

```
function algo($start_time,$end_time,$data_interval,$deg_threshold,$val_threshold) {  
    include_once("getdata.php");  
  
    $count=0;  
  
    //    $next_day_start=strtotime("+$days_count-1 days",$start_time);  
    //    $next_day_end=strtotime("+$days_count-1 days",$end_time);  
  
    $data_array=getdata($start_time,$end_time,$data_interval);  
  
    for($i=0;$i<86400/($data_interval*15*60)-3;$i++) {  
        //86400/($data_interval*15*60)-3 einai o arithmos twn diasthmatwn ana hmera opws  
        prokypetei apo to ekastote data_interval(afairoume 3 epeidh doulevoume ana triades  
        gia na meinoume entos metrhsewn  
  
        $left=$data_array[$i]["KWh"];  
  
        $middle = $data_array[$i+1]["KWh"];  
  
        $right=$data_array[$i+2]["KWh"];  
  
        $rise_deg = rad2deg(atan(($middle-$left)/($data_interval*15)));  
  
        $drop_deg = rad2deg(atan(($right-$middle)/($data_interval*15)));  
  
        if ($left!=0) {  
            if (((($rise_deg>=$deg_threshold)&&($drop_deg<=-  
$deg_threshold))  
  
                &&(($middle/$left)>$val_threshold)) {  
  
                //    echo "peak found at ".$data_array[$i+1]["timestamp"]."  
                with value ".$middle."<br/>";  
  
                $peaks_matrix[$count++]=array('timestamp'=>$data_array[$i+1]["timestamp"  
                ],'KWh'=>$middle);  
  
                }  
    }  
}
```

```

    }

    else {

        if (($rise_deg>=$deg_threshold)&&($drop_deg<=-
$deg_threshold)) {

//            echo "peak found at ".$data_array[$i+1]["timestamp"]."
with value ".$middle."<br/>";

            $peaks_matrix[$count++]=array('timestamp'=>$data_array[$i+1]["timestamp"
], 'KWh'=>$middle);

        }

    }

}

//    echo '<pre>', print_r($peaks_matrix, true), '</pre>';

    return $peaks_matrix;

}

?>

```

Week.php

```
<?php
```

```
//evresh koryfvn gia mia evdomada kai apothikeysh tous se koino pinaka
```

```
function
```

```
week($start_time,$end_time,$data_interval,$days_count,$deg_threshold,$val_thresho
ld,$timestamp_threshold) {
```

```
    include_once ("algo.php");
```

```
    $first_day_end=strtotime("+1 day",$start_time);
```

```

    $peaks_matrix =
algo($start_time,$first_day_end,$data_interval,$deg_threshold,$val_threshold);

    $time_offset=0;

    $next_day_start=$start_time;

    for ($j=0;$j<sizeof($peaks_matrix);$j++) {

        $week_matrix[$j]["time_offset"]=($peaks_matrix[$j]["timestamp"]-
$start_time)/900;

        $week_matrix[$j][1]=$peaks_matrix[$j]["KWh"];

    }

//    echo '<pre>', print_r($peaks_matrix, true), '</pre>';

    for ($count=2;$count<=$days_count;$count++) {

        $next_day_start=strtotime("+1 day",$next_day_start);

        $next_day_end=strtotime("+2 days",$next_day_start);

        $peaks_matrix=algo($next_day_start,$next_day_end,$data_interval,$deg_thre
shold,$val_threshold);

//        echo '<pre>', print_r($peaks_matrix, true), '</pre>';

        for ($i=0;$i<sizeof($week_matrix);$i++) {

            $week_matrix[$i][$count]=0;

        }

        for ($j=0;$j<sizeof($peaks_matrix);$j++) {

            $time_offset=($peaks_matrix[$j]["timestamp"]-
$next_day_start)/900;

            for ($i=0;$i<sizeof($week_matrix);$i++) {

                $timestamp=$next_day_start+$week_matrix[$i]["time_offset"]*900;

                if (($peaks_matrix[$j]["timestamp"] <= ($timestamp +
$timestamp_threshold))

                    &&($peaks_matrix[$j]["timestamp"] >=
($timestamp - $timestamp_threshold))) {

```

```

        $week_matrix[$i][$count]=$peaks_matrix[$j]["KWh"];

        break;

    }

}

$week_matrix[$i]["time_offset"]=$time_offset;

$week_matrix[$i][$count]=$peaks_matrix[$j]["KWh"];

$stempcount=$count-1;

while
((($stempcount>0)&&!(isset($week_matrix[$i][$stempcount])))) {

    $week_matrix[$i][$stempcount]=0;

    $stempcount--;

}

}

}

return $week_matrix;

}

?>

```

Associate_peaks.php

```
<?php
```

```
//sysxetish koryfwn pou emfanizontai kathe hmera se sygekrimmenh wra
```

```
function
```

```
associate_peaks($start_time,$end_time,$data_interval,$days_count,$deg_threshold,$
val_threshold,$timestamp_threshold) {
```



```

include_once ("week.php");

$week_matrix=week($start_time,$end_time,$data_interval,$days_count,$deg
_threshold,$val_threshold,$timestamp_threshold);

$count=0;

$j=1;

$sum=0;

for ($i=0;$i<sizeof($week_matrix);$i++) {
    for ($j=1;$j<=$days_count;$j++) {
        if ($week_matrix[$i][$j]!=0) {
            $sum+=$week_matrix[$i][$j];
        }
        else {
            $sum=0;
            break;
        }
    }
    if ($sum!=0) {

$peak_avg[$count]["time_offset"]=$week_matrix[$i]["time_offset"];

        for ($k=1;$k<=$days_count;$k++) {
            $peak_avg[$count][$k]=$week_matrix[$i][$k];
        }

        $peak_avg[$count++]["average"]=$sum/$days_count;

        $sum=0;

    }
}

return $peak_avg;
}

```

KPI_1.php

```
<?php
```

```
//kathorismos Pmax/Pavg gia mia sygekrimmenh hmera ana triwro
```

```
function
```

```
KPI_1($start_time,$end_time,$data_interval,$deg_threshold,$val_threshold,$days_count) {
```

```
    include_once ("algo.php");
```

```
    include_once ("Pavg.php");
```

```
    $start=$start_time;
```

```
    for ($count=0;$count<$days_count;$count++) {
```

```
        $end=strtotime("+1 day",$start);           //allagh
```

```
//        $end=strtotime("+653 seconds",$start);
```

```
        $peaks=algo($start,$end,$data_interval,$deg_threshold,$val_threshold);
```

```
        $avg=Pavg($start,$end,$data_interval);
```

```
        $max_peak=array(0,0,0,0,0,0);
```

```
        for ($j=0;$j<sizeof($peaks);$j++) {
```

```
            $time_offset=($peaks[$j]["timestamp"]-$start)/900;
```

```
//neo time offset gia edw mono allazw to 900 se 6(gia ta test cases epeidh exoume  
metrhseis ana 6 sec diairoume me to 6,ston kanoniko kwdika ginetai diaresh me  
15*60=900)
```

```
//            echo $time_offset."</br>";
```

```
            switch ($time_offset) {
```

```
                case (($time_offset>=24)&&($time_offset<36)):
```

```
                    if ($peaks[$j]["KWh"]>$max_peak[0]) {
```

```
                        $max_peak[0]=$peaks[$j]["KWh"];
```

```
                    }
```

```

        break;

        case (($time_offset>=36)&&($time_offset<48)):

            if ($peaks[$j]["KWh"]>$max_peak[1]) {

$max_peak[1]=$peaks[$j]["KWh"];

            }

            break;

        case (($time_offset>=48)&&($time_offset<60)):

            if ($peaks[$j]["KWh"]>$max_peak[2]) {

$max_peak[2]=$peaks[$j]["KWh"];

            }

            break;

        case (($time_offset>=60)&&($time_offset<72)):

            if ($peaks[$j]["KWh"]>$max_peak[3]) {

$max_peak[3]=$peaks[$j]["KWh"];

            }

            break;

        case (($time_offset>=72)&&($time_offset<84)):

            if ($peaks[$j]["KWh"]>$max_peak[4]) {

$max_peak[4]=$peaks[$j]["KWh"];

            }

            break;

        case (($time_offset>=84)&&($time_offset<96)):

            if ($peaks[$j]["KWh"]>$max_peak[5]) {

$max_peak[5]=$peaks[$j]["KWh"];

```

```

        }
        break;
    }
}

//      echo '<pre>', print_r($max_peak, true), '</pre>';
        for ($i=0;$i<=5;$i++) {
            $KPI_1[$count][$i]=$max_peak[$i]/$avg[$i];
        }
        $start=$end;
    }
    return $KPI_1;
}

```

KPI_2.php

```

<?php

//kathorismos arithmou koryfwn se kathe triwro ths hmeras

function
KPI_2($start_time,$end_time,$data_interval,$deg_threshold,$val_threshold,$days_co
unt) {

    include_once ("algo.php");

    $start=$start_time;

    for ($count=0;$count<$days_count;$count++) {

        $end=strtotime("+1day",$start);           //allagh

//            $end=strtotime("+653 seconds",$start);

```

```

$peaks=algo($start,$end,$data_interval,$deg_threshold,$val_threshold);
    for ($i=0;$i<=5;$i++) {
        $number_of_peaks[$count][$i]=0;
    }
    for ($j=0;$j<sizeof($peaks);$j++) {
        $time_offset=($peaks[$j]["timestamp"]-$start)/900;    //allagh
900 me 6

        switch ($time_offset) {
            case (($time_offset>=24)&&($time_offset<36)):
                $number_of_peaks[$count][0]++;
                break;
            case (($time_offset>=36)&&($time_offset<48)):
                $number_of_peaks[$count][1]++;
                break;
            case (($time_offset>=48)&&($time_offset<60)):
                $number_of_peaks[$count][2]++;
                break;
            case (($time_offset>=60)&&($time_offset<72)):
                $number_of_peaks[$count][3]++;
                break;
            case (($time_offset>=72)&&($time_offset<84)):
                $number_of_peaks[$count][4]++;
                break;
            case (($time_offset>=84)&&($time_offset<96)):
                $number_of_peaks[$count][5]++;
                break;
        }
    }

```

```

        }

        $start=$end;
    }

    return $number_of_peaks;
}

```

KPI_3.php

```
<?php
```

```
//kathorismos Ppavg/Pavg ana triwro gia mia hmera
```

```
function
```

```
KPI_3($start_time,$end_time,$data_interval,$deg_threshold,$val_threshold,$days_count) {
```

```
    include_once ("algo.php");
```

```
    include_once ("Pavg.php");
```

```
    $start=$start_time;
```

```
    for ($count=0;$count<$days_count;$count++) {
```

```
        $end=strtotime("+1 day",$start);    //allagh
```

```
//        $end=strtotime("+653 seconds",$start);
```

```
    $peaks=algo($start,$end,$data_interval,$deg_threshold,$val_threshold);
```

```
    $avg=Pavg($start,$end,$data_interval);
```

```
    $num_of_peaks=array(0,0,0,0,0,0);
```

```
    $peak_sum=array(0,0,0,0,0,0);
```

```
    for ($j=0;$j<sizeof($peaks);$j++) {
```

900 me 6

```
$time_offset=($peaks[$j]["timestamp"]-$start)/900; //allagh
```

```
switch ($time_offset) {  
    case (($time_offset>=24)&&($time_offset<36)):  
        $peak_sum[0]+=$peaks[$j]["KWh"];  
        $num_of_peaks[0]++;  
        break;  
    case (($time_offset>=36)&&($time_offset<48)):  
        $peak_sum[1]+=$peaks[$j]["KWh"];  
        $num_of_peaks[1]++;  
        break;  
    case (($time_offset>=48)&&($time_offset<60)):  
        $peak_sum[2]+=$peaks[$j]["KWh"];  
        $num_of_peaks[2]++;  
        break;  
    case (($time_offset>=60)&&($time_offset<72)):  
        $peak_sum[3]+=$peaks[$j]["KWh"];  
        $num_of_peaks[3]++;  
        break;  
    case (($time_offset>=72)&&($time_offset<84)):  
        $peak_sum[4]+=$peaks[$j]["KWh"];  
        $num_of_peaks[4]++;  
        break;  
    case (($time_offset>=84)&&($time_offset<96)):  
        $peak_sum[5]+=$peaks[$j]["KWh"];  
        $num_of_peaks[5]++;  
        break;  
}
```

```

        }
    }
//    echo '<pre>', print_r($peak_sum, true), '</pre>';
//    echo '<pre>', print_r($num_of_peaks, true), '</pre>';
    for ($i=0;$i<=5;$i++) {
        if ($num_of_peaks[$i]!=0) {
            $Ppavg[$i]=$peak_sum[$i]/$num_of_peaks[$i];
            $KPI_3[$count][$i]=$Ppavg[$i]/$avg[$i];
        }
        else {
            $KPI_3[$count][$i]=0;
        }
    }
    $start=$end;
}
return $KPI_3;
}

```

KPI_4.php

```
<?php
```

```
//ypologismos rythmou anodou h kathodou
```

```

function KPI_4($start_time,$end_time,$data_interval,$days_count) {
    include_once ("getdata.php");

```



```

include_once ("Pavg.php");

$start=$start_time;

for ($count=0;$count<$days_count;$count++) {

    $send=strtotime("+1 day",$start);    //allagh
//    $send=strtotime("+653 seconds",$start);

    $data=getdata($start,$send,$data_interval);

    $Pavg=Pavg($start,$send,$data_interval);

    for ($i=2;$i<=7;$i++) {

        $begin=$i*12;

        $finish=($i+1)*12 - 1;

        $KPI_4[$count][$i-2]=($data[$finish]['KWh']-
$data[$begin]['KWh'])/$Pavg[$i-2];

    }

    $start=$send;

}

return $KPI_4;

}

?>

```

KPI_total.php

```

<?php

function
KPI_total($start_time,$send_time,$data_interval,$deg_threshold,$val_threshold,$days
_count) {

    include_once ("KPI_1.php");

    include_once ("KPI_2.php");

```

```

include_once ("KPI_3.php");

include_once ("KPI_4.php");

$KPI_1=KPI_1($start_time,$end_time,$data_interval,$deg_threshold,$val_threshold,$days_count);

$KPI_2=KPI_2($start_time,$end_time,$data_interval,$deg_threshold,$val_threshold,$days_count);

$KPI_3=KPI_3($start_time,$end_time,$data_interval,$deg_threshold,$val_threshold,$days_count);

$KPI_4=KPI_4($start_time,$end_time,$data_interval,$days_count);

for ($count=0;$count<$days_count;$count++) {

    for ($i=0;$i<=5;$i++) {

        $line1[$i]=$KPI_1[$count][$i];

        $line2[$i]=$KPI_2[$count][$i];

        $line3[$i]=$KPI_3[$count][$i];

        $line4[$i]=$KPI_4[$count][$i];

    }

    $element1=array_keys($line1,max($line1));

    $element2=array_keys($line2,max($line2));

    $element3=array_keys($line3,max($line3));

    $element4=array_keys($line4,max($line4));

    $KPI_total[$count][0]=$element1[0];

    $KPI_total[$count][1]=$element2[0];

    $KPI_total[$count][2]=$element3[0];

    $KPI_total[$count][3]=$element4[0];

}

return $KPI_total;

}

?>

```

Output.php

```
<?php

function
output($start_time,$end_time,$data_interval,$deg_threshold,$val_threshold,$days_co
unt) {

    include_once ("KPI_1.php");

    include_once ("KPI_2.php");

    include_once ("KPI_3.php");

    include_once ("KPI_4.php");

    include_once ("KPI_total.php");

    $KPI_1=KPI_1($start_time,$end_time,$data_interval,$deg_threshold,$val_thr
eshold,$days_count);

    $KPI_2=KPI_2($start_time,$end_time,$data_interval,$deg_threshold,$val_thr
eshold,$days_count);

    $KPI_3=KPI_3($start_time,$end_time,$data_interval,$deg_threshold,$val_thr
eshold,$days_count);

    $KPI_4=KPI_4($start_time,$end_time,$data_interval,$days_count);

    $KPI_total=KPI_total($start_time,$end_time,$data_interval,$deg_threshold,$
val_threshold,$days_count);

    $tab = "\t";

    $cr = "\r\n";

    $data1=$data2=$data3=$data4=$data5="";

    for ($count=0;$count<$days_count;$count++) {

        for ($i=0;$i<5;$i++) {

            $data1.=$KPI_1[$count][$i].$tab;

            $data2.=$KPI_2[$count][$i].$tab;

            $data3.=$KPI_3[$count][$i].$tab;

            $data4.=$KPI_4[$count][$i].$tab;

        }

    }

}
```

```

        $data1.=$KPI_1[$count][5].$cr;
        $data2.=$KPI_2[$count][5].$cr;
        $data3.=$KPI_3[$count][5].$cr;
        $data4.=$KPI_4[$count][5].$cr;
        for ($j=0;$j<3;$j++) {
            $data5.=$KPI_total[$count][$j].$tab;
        }
        $data5.=$KPI_total[$count][3].$cr;
    }
    $fp1=fopen("A1.dat","w");
    fwrite($fp1,$data1);
    fclose($fp1);
    $fp2=fopen("A2.dat","w");
    fwrite($fp2,$data2);
    fclose($fp2);
    $fp3=fopen("A3.dat","w");
    fwrite($fp3,$data3);
    fclose($fp3);
    $fp4=fopen("A4.dat","w");
    fwrite($fp4,$data4);
    fclose($fp4);
    $fp5=fopen("Atotal.dat","w");
    fwrite($fp5,$data5);
    fclose($fp5);
}
?>

```

Main.php

```
<?php  
  
include_once ("output.php");  
  
$output=output(1243243173,1243330148,1,75,1.06,7);  
  
?>
```

9.4 Κώδικας MATLAB

```
load 'C:\Program Files\EasyPHP 3.0\www\A1.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\A2.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\A3.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\A4.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\Atotal.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\B1.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\B2.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\B3.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\B4.dat'  
load 'C:\Program Files\EasyPHP 3.0\www\Btotal.dat'  
  
[coeffA1 latentA1 explainedA1]=pcacov(A1*A1');  
[coeffA2 latentA2 explainedA2]=pcacov(A2*A2');  
[coeffA3 latentA3 explainedA3]=pcacov(A3*A3');  
[coeffA4 latentA4 explainedA4]=pcacov(A4*A4');  
[coeffAtotal latentAtotal explainedAtotal]=pcacov(Atotal*Atotal');  
[coeffB1 latentB1 explainedB1]=pcacov(B1*B1');  
[coeffB2 latentB2 explainedB2]=pcacov(B2*B2');  
[coeffB3 latentB3 explainedB3]=pcacov(B3*B3');  
[coeffB4 latentB4 explainedB4]=pcacov(B4*B4');  
[coeffBtotal latentBtotal explainedBtotal]=pcacov(Btotal*Btotal');
```