



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Προηγμένα Συστήματα Πληροφορικής»

Μεταπτυχιακή Διατριβή

| | |
|-----------------------|--|
| Τίτλος Διατριβής | Φίλτρα συνεργασίας σε περιορισμένα Κοινωνικά Δίκτυα Collaborating Filtering in a Limited Social Network |
| Όνοματεπώνυμο Φοιτητή | Ευάγγελος Ταμπάκης |
| Πατρώνυμο | Δημήτριος |
| Αριθμός Μητρώου | ΜΠΣΠ/ 11019 |
| Επιβλέπων | Μαρία Βίρβου, Καθηγήτρια |

Ημερομηνία Παράδοσης **Οκτώβριος 2014**

Πανεπιστήμιο Πειραιώς

Τριμελής Εξεταστική Επιτροπή

(υπογραφή)

Μαρία Βίρβου
Καθηγήτρια

(υπογραφή)

Γεώργιος Τσιχριτζής
Καθηγητής

(υπογραφή)

Ευθύμιος Αλέπτης
Λέκτορας

Περιεχόμενα

| | |
|--|----|
| Περίληψη (abstract)..... | 7 |
| Εισαγωγή | 8 |
| ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ | |
| 1. Κοινωνικά Δίκτυα..... | 9 |
| 1.1 Γενικά..... | 9 |
| 1.2 Παραδείγματα δημοφιλών κοινωνικών δικτύων..... | 10 |
| 1.3 Πλεονεκτήματα κοινωνικών δικτύων..... | 12 |
| 1.4 Μειονεκτήματα κοινωνικών δικτύων..... | 13 |
| 1.5 Επιχειρηματικές ευκαιρίες και κοινωνικά δίκτυα..... | 15 |
| 2. Συστήματα συστάσεων (Recommendation systems)..... | 15 |
| 2.1 Γενικά..... | 15 |
| 2.2 Κατηγορίες συστημάτων συστάσεων..... | 17 |
| 2.2.1. Συστήματα collaborative filtering..... | 18 |
| 2.2.1.1. Item-based collaborative filtering..... | 18 |
| 2.2.1.2. Αξιολόγηση της μεθόδου collaborative filtering..... | 19 |
| 2.2.1.3. Αξιολόγηση της μεθόδου collaborative filtering..... | 20 |
| 2.2.1.4. Άμεση - έμμεση αξιολόγηση..... | 22 |
| 2.2.2. Συστήματα βασισμένα στην γνώση..... | 23 |
| 2.2.2.1. Critique-based recommendation..... | 24 |
| 2.2.2.2. Goal-based recommendation..... | 24 |
| 2.2.2.3. Αξιολόγηση των συστημάτων γνώσης..... | 25 |
| 2.2.3. Υβριδικά συστήματα συστάσεων..... | 26 |
| 2.2.3.1. Υβριδικά συστήματα τύπου «καταρράκτη»..... | 27 |
| 2.2.4. Παράμετροι επιτυχημένων τεχνικών συστάσεων..... | 27 |
| 2.2.4.1. Συνθήκες έλλειψης επαρκούς αριθμού αξιολογήσεων..... | 28 |
| 2.2.4.2. Αιτιολόγηση συστάσεων..... | 28 |
| 2.2.4.3. Χρήστες με ασυνήθιστες προτιμήσεις..... | 28 |
| 2.2.4.4. Ελλείψη επαρκών στοιχείων γνώσης..... | 29 |
| 2.2.4.5. Ανάγκη γρήγορης ανάπτυξης..... | 29 |
| 3. Εξόρυξη γνώσης από Δεδομένα (Knowledge Mining)..... | 29 |
| 3.1. Εισαγωγή..... | 29 |
| 3.2. Τεχνικές εξόρυξης δεδομένων..... | 31 |
| 3.3. Παράγοντες επιτυχίας εξόρυξης δεδομένων..... | 32 |
| 3.4. Εφαρμογές εξόρυξης δεδομένων..... | 33 |
| ΠΡΑΚΤΙΚΟ ΜΕΡΟΣ | |
| 4. Ανάλυση Τεχνολογιών..... | 35 |
| 4.1. Ανάλυση τεχνολογιών που χρησιμοποιήθηκαν..... | 35 |
| 4.1.1. Η γλώσσα προγραμματισμού C# & .NET 4.5 Framework..... | 35 |
| 4.1.2. Η γλώσσες JavaScript & JQuery..... | 36 |
| 4.1.3. Γλώσσα διαμόρφωσης περιεχομένου CSS..... | 37 |
| 4.1.4. Γλώσσα ερωτημάτων βάσης δεδομένων SQL..... | 38 |
| 4.1.5. Το Microsoft Visual Studio 2012..... | 38 |
| 4.1.6. Το MS SQL Studio 2008..... | 39 |
| 4.2. Βάση Δεδομένων..... | 40 |
| 4.2.1. Αναλυτική περιγραφή πινάκων..... | 40 |
| 5. Υλοποίηση συστήματος..... | 47 |
| 5.1. Σύντομη περιγραφή αλγορίθμων συστάσεων..... | 47 |
| 5.1.1. Σύντομη αναφορά βημάτων αλγορίθμου..... | 47 |
| 5.1.2. Σύντομη περιγραφή ανατροφοδότησης αλγορίθμου και μεταβολής προφίλ χρήστη..... | 48 |
| 5.2. Ανάλυση βημάτων αλγορίθμου συστάσεων στον χρήστη..... | 48 |

| | |
|---|----|
| 5.2.1. Ανάκτηση δεδομένων χρηστών και εξαγωγή πίνακα ομοιότητας τους..... | 48 |
| 5.2.2. Εύρεση μη αξιολογήσεων καταστημάτων από το χρήστη μας..... | 51 |
| 5.2.3. Εύρεση μη αξιολογήσεων καταστημάτων από το χρήστη μας..... | 52 |
| 5.2.4. Εμπλουτισμός προτάσεων με την πληροφορία των φίλων του χρήστη..... | 53 |
| 5.2.5. Εφαρμογή φίλων χρήστη..... | 53 |
| 5.2.6. Ανατροφοδότηση αλγορίθμου και προφίλ χρήστη..... | 55 |
| 5.2.6.1. Ανατροφοδότηση αλγορίθμου..... | 55 |
| 5.2.6.2. Προφίλ χρήστη..... | 56 |
| 5.2.6.3. Ενημέρωση προφίλ χρήστη..... | 58 |
| 6. Ανάλυση και περιγραφή της ιστοσελίδας exodus..... | 61 |
| 6.1. Περιγραφή σελίδων..... | 61 |
| 6.1.1. Register, εγγραφή χρήστη..... | 61 |
| 6.1.2. Login, σύνδεση χρήστη στην ιστοσελίδα..... | 63 |
| 6.1.3. Index, αρχική σελίδα exodus..... | 63 |
| 6.1.4. Stores. Σελίδα αναζήτησης καταστημάτων..... | 68 |
| 6.1.5. Specials. Σελίδα προτάσεων καταστημάτων στον χρήστη..... | 70 |
| 6.1.6. UserProfile. Σελίδα πληροφοριών φίλων χρήστη..... | 72 |
| 6.1.7. StoresDetails. Σελίδα αναλυτικών πληροφοριών καταστήματος..... | 73 |
| 6.1.8. ManageAccount. Σελίδα αλλαγή προφίλ χρήστη..... | 75 |
| 7. Συμπεράσματα..... | 76 |
| 7.1. Εισαγωγή..... | 76 |
| 7.2. Συμπεράσματα..... | 76 |
| 7.3. Εξέλιξη στο μέλλον..... | 77 |
| 8. Βιβλιογραφία..... | 78 |
| 9. Πίνακας εικόνων..... | 80 |

Περίληψη (Abstract)

The creation of a wide network that is today known as the internet gave commercial corporations the solid base to look for customers. This led the corporations to focus in the development of websites and web applications that have as main purpose to present the merchandisable products and the services to the users of that network. Soon it became a fact that the simple presentation of those products is not good enough to convince the users to prefer their products from the growing completion's similar ones. The recommendation systems took their part to help find the solution to such problem and attract a larger amount of users. Recommendations systems main goal is to provide users a more targeted amount of suggestions. This way users are provided with products that are more suitable to them, something that increases the probability of them purchasing those products. The modern trend shows an increasing interest of users in using social media networks. As a result an increased number of those networks were created in which users are not an individual entity but members of a wide group. A user that belongs to such a network is more prone to revealing his consumer preferences in products and services. This user is always seeking his connected fellow users and takes their opinion rather seriously.

The main purpose of this diploma thesis is to create a recommendations algorithm which is based on user interaction with products. In addition with this algorithm we will create a limited social network in which the users that belong to it will be a part of it and interact with the algorithm. The result of the combination of the recommendations algorithm with that limited social network will create a unified recommendations system that will include the advantages of both those technologies in order to maximize the quality of the recommended items.

Η δημιουργία του ευρύτερου δικτύου που σήμερα απαντά στο όνομα διαδίκτυο έδωσε το έναυσμα στις επιχειρήσεις να αναζητήσουν το αγοραστικό τους κοινό στους χρήστες αυτού του δικτύου. Αυτό οδήγησε τις εταιρείες αυτές να προσανατολιστούν στην δημιουργία ιστοσελίδων και web εφαρμογών που έχουν σαν σκοπό να παρουσιάσουν τα προϊόντα αλλά και τις υπηρεσίες που εμπορεύονται στο νέο αυτό μέσο, τον χώρο του διαδικτύου. Γρήγορα διαφάνηκε ότι η απλή παρουσίαση των προϊόντων τους στον χρήστη δεν αρκεί να τον πείσει ώστε να προτιμήσει τα δικά τους προϊόντα έναντι στα αντίστοιχα που προσέφερε ο όλο και αυξανόμενος ανταγωνισμός. Αυτό οδήγησε τις εταιρείες να προσανατολιστούν και να επενδύσουν σε ποιο έξυπνες μεθόδους ώστε να προσελκύσουν μεγαλύτερο όγκο πελατών. Τα συστήματα προτάσεων (recommendation systems) ήρθαν να δώσουν λύση στο πρόβλημα αυτό. Στόχος των συστημάτων αυτών είναι να αξιοποιήσουν την αλληλεπίδραση των χρηστών με τα προϊόντα τους έτσι ώστε να παρέχουν στοχευμένες προτάσεις σε αυτούς. Με αυτό τον τρόπο οι χρήστες λαμβάνουν προτάσεις για προϊόντα τα οποία είναι πιο ελκυστικά σε αυτούς κάτι που αυξάνει τις πιθανότητες να προτιμήσει τα συγκεκριμένα προϊόντα. Η σύγχρονη τάση των χρηστών δείχνει μία στροφή στα κοινωνικά δίκτυα. Αυτό οδήγησε στην δημιουργία πολλών κοινωνικών δικτύων στα οποία οι χρήστες δεν είναι πλέον μεμονωμένες οντότητες αλλά μέλος μια ευρύτερης ομάδας. Ο χρήστης που είναι ενταγμένος σε μια τέτοια ομάδα εκφράζει με μεγαλύτερη ευκολία τις προτιμήσεις του απέναντι σε προϊόντα και υπηρεσίες. Αναζητά τους συνδεδεμένους με αυτόν χρήστες και λαμβάνει σοβαρά υπόψη του την γνώμη τους.

Στόχος αυτής της εργασίας είναι να υλοποιήσει ένα σύστημα προτάσεων το οποίο στηρίζεται στην αλληλεπίδραση του χρήστη με τα προϊόντα. Ταυτόχρονα με τον αλγόριθμο αυτόν θα δημιουργήσουμε ένα κοινωνικό δίκτυο περιορισμένης κλίμακας στο οποίο οι χρήστες μπορούν να ενταχθούν και αλληλεπιδράσουν μεταξύ τους. Το αποτέλεσμα του κοινωνικού δικτύου σε συνδυασμό με τον αλγόριθμο προτάσεων θα δημιουργήσουν ένα σύστημα προτάσεων το οποίο θα περιέχει τα πλεονεκτήματα και των δύο αυτών τεχνολογιών με σκοπό την καλύτερη αποτελεσματικότητα των συστάσεων που θα εξάγει.

Εισαγωγή

Στην παρούσα εργασία θα δημιουργήσουμε μία ιστοσελίδα η οποία σκοπό έχει να παρουσιάσει στον χρήστη αυτής, καταστήματα εστίασης και ψυχαγωγίας τα οποία μπορεί να επισκεφθεί. Τα καταστήματα αυτά περιέχουν κάποια διακριτά χαρακτηριστικά τα οποία τα κατατάσσουν σε κατηγορίες. Τα χαρακτηριστικά αφορούν την κατηγορία του καταστήματος και το μουσικό είδος στο οποίο ανήκει το κατάστημα. Τα παραπάνω στοιχεία θα συνδυαστούν έτσι ώστε να δημιουργήσουμε ένα σύστημα προτάσεων το οποίο θα παράγει προτάσεις στους χρήστες με βάση την πρότερη γνώση που έχουμε για αυτούς καθώς και την αλληλεπίδραση τους με τα καταστήματα. Η αλληλεπίδραση αυτή αφορά την αξιολόγηση που ο κάθε χρήστης δίνει στο συγκεκριμένο κατάστημα.

Για να εξάγουμε τα δεδομένα των προτάσεων στον χρήστη θα υλοποιήσουμε ένα αλγόριθμο που βασίζεται στα συνεργατικά φίλτρα (collaborating filtering). Ο αλγόριθμος αυτός έχει σαν σκοπό να αξιοποιήσει τις βαθμολογήσεις των χρηστών για τα καταστήματα συνδυάζοντας αυτά τα δεδομένα με το προφίλ του χρήστη. Το προφίλ του χρήστη αποτελεί ένα στερεότυπο για τις προτιμήσεις του κάθε χρήστη σχετικά με τα δύο είδη χαρακτηριστικών που αναλύσαμε παραπάνω. Το προφίλ αυτό αποτελεί ένα δυναμικό γνώμονα για τις προτιμήσεις του κάθε χρήστη το οποίο μεταβάλλεται όσο ο χρήστης αλληλεπιδρά με τα καταστήματα της ιστοσελίδας μας. Παράλληλα θα δημιουργήσουμε ένα περιορισμένο κοινωνικό δίκτυο (limited social network) το οποίο αποτελεί μια μικρή κοινότητα για τους χρήστες της ιστοσελίδας μας. Το κοινωνικό αυτό δίκτυο επιτρέπει στους χρήστες την δημιουργία μικρών ή και μεγαλύτερων γκρουπ συνδεδεμένων χρηστών(ομάδα φίλων). Ο συνδυασμός του αλγορίθμου προτάσεων σε συνεργασία με το κοινωνικό δίκτυο τη ιστοσελίδας θα μας επιτρέψει να εξάγουμε δεδομένα για τις προτιμήσεις της ομάδας χρηστών αλλά και των μεμονωμένων χρηστών. Τα δεδομένα αυτά αποτελούν μία πιο ενδεικτική εικόνα για τις προτιμήσεις κάθε χρήστη και μας επιτρέπουν την εξαγωγή πιο ποιοτικών προτάσεων.

Για την διευκόλυνση του αναγνώστη θα πρέπει σε αυτό το σημείο να αναφερθούμε στην διάρθρωση της εργασίας. Αρχικά θα αναφερθούμε στα κοινωνικά δίκτυα αναφέροντας την ιστορική τους ανοδική πορεία. Θα αναλύσουμε τα πλεονεκτήματα τους αλλά και τα μειονεκτήματα τους καθώς και την χρησιμότητα τους στα συστήματα προτάσεων. Επίσης θα αναφέρουμε τα κυριότερα είδη αλγορίθμων συστάσεων που υπάρχουν στην αγορά. Θα αναλύσουμε κάποια από αυτά παρουσιάζοντας μειονεκτήματα και πλεονεκτήματα αλλά και την κύρια χρήση τους. Στο πρακτικό μέρος του παρόντος θα αναλύσουμε τις τεχνικές υλοποίησης της ιστοσελίδας Exodus, παρουσιάζοντας τις τεχνολογίες που χρησιμοποιήθηκαν. Θα αναλύσουμε τα επιμέρους υποσυστήματα που αποτελούν την ιστοσελίδα μας όπως την βάση δεδομένων, την επικοινωνία με τον χρήστη κτλ. Επίσης θα δούμε βήμα-βήμα την συμπεριφορά του αλγορίθμου και την διαδικασία που ακολουθείται για την εξαγωγή των συστάσεων στον χρήστη. Θα δούμε επίσης την ενσωμάτωση του κοινωνικού δικτύου που υλοποιήσαμε με τον αλγόριθμο συστάσεων και θα παρακολουθήσουμε από την πλευρά του χρήστη πώς το κοινωνικό δίκτυο έχει αντίκτυπο στις επιλογές του χρήστη.

Τέλος, θα αναλύσουμε το user interface της εφαρμογής το οποίο έχει σχεδιαστεί έτσι ώστε να παρέχει στον χρήστη με συντομία και με αποτελεσματικότητα τον πυρήνα αυτής της εργασίας το οποίο είναι οι προτάσεις στον χρήστη. Θα αναλύσουμε κατά σελίδα όλο τον ιστότοπο που έχουμε δημιουργήσει και θα παρουσιάσουμε τις εξελιγμένες τεχνικές οι οποίες χρησιμοποιήθηκαν έτσι ώστε να αντλήσουμε πληροφορία από τον χρήστη για να ανατροφοδοτήσουμε τον αλγόριθμο με σκοπό να παρέχουμε καλύτερες προτάσεις σε αυτόν.

ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ

1. Κοινωνικά Δίκτυα

1.1 Γενικά

Όταν χρησιμοποιούμε τον όρο κοινωνικό δίκτυο, θεωρούμε ότι αναφερόμαστε στο σύνολο των σχέσεων που αναπτύσσονται σε μία συγκεκριμένη ομάδα. Στο πλαίσιο της θεωρίας των κοινωνικών δικτύων, κάθε διαφορετική σχέση, οποιαδήποτε μορφή και αν λαμβάνει, εκφράζει αυτόματα και τη σύνδεση δύο υποκειμένων. Σε μία ομάδα που ερευνάται με τις τεχνικές και μεθόδους της ανάλυσης κοινωνικών δικτύων το σύνολο των καταγεγραμμένων σχέσεων αποτυπώνονται γραφικά με κόμβους (nodes) που εκφράζουν τα άτομα και γραμμές που εκφράζουν τις συνδέσεις μεταξύ τους. Αυτή η σχεδόν απλοϊκή σχηματική απεικόνιση αποκτά νόημα με ποσοτικούς όρους ανάλυσης στην κοινωνική έρευνα, αποκαλύπτοντας τη «γεωμετρία» της δομής της υπό έρευνα ομάδας, όπως αυτή γίνεται αντιληπτή και κατανοητή από το «πλέγμα» των καταγεγραμμένων σχέσεων που αναπτύσσουν τα διακριτά μέλη που την αποτελούν.

Στα πρώτα χρόνια λειτουργίας του Διαδικτύου, από τα μέσα και έως τα τέλη της δεκαετίας του 1990, πολλές επιχειρήσεις φοβούνταν ότι οι εργαζόμενοι θα έχαναν χρόνο σερφάροντας άσκοπα, για αυτό προσπάθησαν να ελέγξουν την πρόσβασή τους σε αυτό. Παρότι ορισμένοι εργαζόμενοι όντως σέρφαραν άσκοπα, πολλοί ανακάλυψαν τρόπους να αξιοποιήσουν τη δύναμη του Internet ώστε να βελτιώσουν τη δουλειά τους, και το διαδίκτυο γρήγορα εξελίχθηκε σε πολύτιμη πηγή πληροφοριών, εξαιρετικό εργαλείο έρευνας και μέσο παρακολούθησης του ανταγωνισμού. Η τελευταία τεχνολογική καινοτομία που εμφανίστηκε στον χώρο εργασίας είναι τα social media, Facebook, LinkedIn, Twitter, Ning, Plaxo, Hi5 και Second Life, τα οποία έφεραν επανάσταση στον τρόπο που επικοινωνούμε, δημιουργούμε δίκτυα και ανταλλάσσουμε πληροφορίες.

Η χρήση των κοινωνικών δικτύων από τους εργαζομένους στο πλαίσιο της εργασίας τους έχει τη δυναμική να μεταμορφώσει συνολικά τον κόσμο της εργασίας. Πολλές γνωστές εταιρείες αξιοποιούν τις δυνατότητες διασύνδεσης που προσφέρουν τα social media για να ενισχύσουν την παραγωγικότητα, την καινοτομία, τη φήμη, τη συνεργασία και τη δέσμευση των εργαζομένων τους με την εταιρεία. Το 75% των επιχειρήσεων διεθνώς δεν διαθέτουν επίσημη πολιτική για τη χρήση ιστοχώρων κοινωνικής δικτύωσης σε ώρα εργασίας. Στην Ελλάδα το ποσοστό αυτό ανέρχεται στο 86%, στην περιοχή EMEA είναι στο 87% και στην Αμερική στο 69%.

Στην πλειοψηφία τους, λοιπόν, οι εταιρείες τηρούν στάση αναμονής απέναντι στο φαινόμενο, πριν προχωρήσουν στην ανάπτυξη επίσημων πολιτικών για τη χρήση των κοινωνικών δικτύων. Οι επικεφαλές θα πρέπει να αναζητήσουν τρόπους αξιοποίησης της δημοτικότητας και της αξίας που μπορούν να τους προσθέσουν τα social media, ενισχύοντας έτσι την απόδοση των οργανισμών τους και προάγοντας τους εταιρικούς στόχους. Ωστόσο, οι προσπάθειες αυτές δεν θα πρέπει να εστιάζουν στον έλεγχο της συμπεριφοράς των εργαζομένων στα κοινωνικά δίκτυα, αλλά στον προσανατολισμό της χρήσης τους προς κατευθύνσεις ωφέλιμες τόσο για τους ίδιους τους οργανισμούς όσο και για τους εργαζομένους.

Αν και social media συνηθίζεται να ονομάζονται μόνο τα μέσα κοινωνικής δικτύωσης (social networks), όπως το facebook και το LinkedIn, τα συγκεκριμένα δεν είναι παρά μόνο δύο από αυτά. Στα social media περιλαμβάνονται και πολλά άλλα όπως τα blogs και media sharing sites, και χωρίζονται στις παρακάτω κατηγορίες:

- Social news and recommendations (π.χ digg.com)
- Social book marking sites (π.χ delicious.com)
- Micro blogging services (π.χ. twitter)
- Blogging systems (π.χ. blogger.com)
- Social networks (π.χ. facebook, linkedin)
- Social sharing (π.χ. youtube, flickr)
- Wikis (π.χ. mediawiki.org)

Κοινωνική δικτύωση είναι η συγκέντρωση ή συμμετοχή των ατόμων σε συγκεκριμένες ομάδες. Εμφανίζονται διάφοροι ορισμοί για το τι είναι ένα κοινωνικό δίκτυο και μερικοί από αυτούς είναι:

- Κοινωνικά δίκτυα είναι τα «πολυδιάστατα συστήματα επικοινωνίας και διαμόρφωσης της ανθρώπινης πρακτικής και της κοινωνικής ταυτότητας».
- Κοινωνικά δίκτυα είναι το άθροισμα των προσωπικών επαφών μέσω των οποίων το άτομο διατηρεί την κοινωνική του ταυτότητα, λαμβάνει συναισθηματική υποστήριξη, υλική ενίσχυση και συμμετοχή στις υπηρεσίες, έχει πρόσβαση στις πληροφορίες και δημιουργεί νέες κοινωνικές επαφές.

Τα κοινωνικά δίκτυα συνήθως αποτελούνται από τα μέλη της οικογένειας, τους φίλους και τους γνωστούς και περιλαμβάνουν τρεις κρίσιμες έννοιες:

- το μέγεθος ή το εύρος, το οποίο αναφέρεται στον αριθμό των ατόμων που συμμετέχουν στο δίκτυο,
- τη σύνθεση, δηλαδή το ποσοστό συμμετοχής μελών της ευρύτερης οικογένειας ή φίλων στο δίκτυο, και
- τη συχνότητα, που δηλώνει το πόσο συχνά τα μέλη ενός κοινωνικού δικτύου αλληλεπιδρούν μεταξύ τους

Όταν πρόκειται για online κοινωνική δικτύωση, χρησιμοποιούνται ιστοσελίδες γνωστές ως ιστότοποι κοινωνικής δικτύωσης. Οι ιστότοποι κοινωνικής δικτύωσης λειτουργούν ως online κοινότητες των χρηστών του Διαδικτύου. Ανάλογα με τα χαρακτηριστικά κάθε ιστότοπου, τα μέλη της online κοινότητας έχουν κοινά ενδιαφέροντα όπως κάποιο χόμπι, τη θρησκεία ή την πολιτική. Η εγγραφή και πρόσβαση σε έναν ιστότοπο κοινωνικής δικτύωσης παρέχει στα μέλη του άμεση και συνεχή κοινωνικοποίηση. Αυτή η κοινωνικοποίηση μπορεί να περιλαμβάνει την ανάγνωση του προφίλ των άλλων μελών και συχνά και την επικοινωνία μαζί του.

Οι ιστοσελίδες κοινωνικής δικτύωσης ή ηλεκτρονικά κοινωνικά δίκτυα, έκαναν την εμφάνισή τους το 2002 με το Friendster. Αυτές οι ιστοσελίδες επιτρέπουν στα άτομα να παρουσιάσουν τους εαυτούς τους, να αναπτύξουν την κοινωνική τους δικτύωση, καθώς και να δημιουργήσουν ή να διατηρήσουν συνδέσεις με άλλους χρήστες. Η ιστοσελίδα κοινωνικής δικτύωσης όμως που ξεχώρισε ανάμεσα στις άλλες είναι το Facebook. Με 132 εκατομμύρια χρήστες, τον Ιούνιο του 2008, και 200% μέση ετήσια αύξηση εγγραφών το Facebook είναι αναμφισβήτητα η δημοφιλέστερη ιστοσελίδα κοινωνικής δικτύωσης.

1.2. Παραδείγματα δημοφιλών κοινωνικών δικτύων

Σήμερα, υπάρχουν πολλοί ιστότοποι κοινωνικής δικτύωσης. Οι ιστότοποι διακρίνονται σε κατηγορίες ανάλογα με:

- το αντικείμενό τους
- τον τρόπο εγγραφής και συμμετοχής μελών
- τον τρόπο επικοινωνίας μεταξύ των μελών τους
- το είδος του περιεχομένου

Το Facebook είναι ιστοχώρος κοινωνικής δικτύωσης που ξεκίνησε στις 5 Ιανουαρίου του 2005. Οι χρήστες μπορούν να επικοινωνούν μέσω μηνυμάτων με τις επαφές τους και να τους ειδοποιούν όταν ανανεώνουν τις προσωπικές πληροφορίες τους. Όλοι έχουν ελεύθερη πρόσβαση στο να συμμετάσχουν σε δίκτυα που σχετίζονται μέσω πανεπιστημίου, θέσεων απασχόλησης ή γεωγραφικών περιοχών.

Ο Μαρκ Ζάκερμπεργκ ίδρυσε το Facebook ως μέλος του Πανεπιστημίου Χάρβαρντ. Αρχικά δικαίωμα συμμετοχής είχαν μόνο οι φοιτητές του Χάρβαρντ ενώ αργότερα επεκτάθηκε για την Ivy League. Το όνομα της ιστοσελίδας προέρχεται από τα έγγραφα παρουσίασης των μελών πανεπιστημιακών κοινοτήτων μερικών Αμερικάνικων κολεγίων και προπαρασκευαστικών σχολείων που χρησιμοποιούσαν οι νεοεισερχόμενοι σπουδαστές για να γνωριστούν μεταξύ τους. Το 2005 το δικαίωμα πρόσβασης επεκτάθηκε σε μαθητές συγκεκριμένων λυκείων και μέλη ορισμένων μαθητικών κοινοτήτων, ενώ το 2006 η υπηρεσία έγινε προσβάσιμη σε κάθε

άνθρωπο του πλανήτη που η ηλικία του ξεπερνούσε τα 13 χρόνια. Το Facebook σήμερα έχει πάνω από 1 δισεκατομμύριο ενεργούς χρήστες, κατατάσσοντάς το έτσι στη λίστα ταξινόμησης του Alexa ως ένα από τα δημοφιλέστερα web sites του πλανήτη (2ο μετά το Google). Επίσης, το Facebook είναι ένα από τα δημοφιλέστερα sites για ανέβασμα φωτογραφιών με πάνω από 14 εκατομμύρια φωτογραφίες καθημερινά. Με αφορμή τη δημοτικότητά του, το Facebook έχει υποστεί κριτική και κατηγορηθεί σε θέματα που αφορούν τα προσωπικά δεδομένα και τις πολιτικές απόψεις των ιδρυτών του. Ωστόσο η συγκεκριμένη ιστοσελίδα παραμένει η πιο διάσημη κοινωνική περιοχή δικτύωσης σε πολλές αγγλόφωνες χώρες. Το Facebook είναι ένας καλός τρόπος δικτύωσης με φίλους και γνωστούς. Παρά το ότι ενέχει κινδύνους (κυρίως για παραβίαση προσωπικών δεδομένων), ο προσεκτικός χρήστης δεν έχει πρόβλημα. Το Facebook ακόμα παρέχει παιχνίδια και υπάρχει η δυνατότητα ανεβάσματος φωτογραφιών και βίντεο. Παρόλα αυτά όμως έχουν παρατηρηθεί κάποιες αρνητικές μεταβολές στην ψυχολογική κατάσταση των χρηστών του facebook. Σε μια έρευνα που έγινε με ερευνητές του πανεπιστημίου του Μίσιγκαν μελέτησαν για 2 εβδομάδες 82 νεαρούς ενήλικες χρήστες με ποιόν τρόπο σχετιζόταν η χρήση του κοινωνικού δικτύου με τα αισθήματα ψυχικής ικανοποίησης κάθε χρήστη. Τελικά οι ερευνητές κατέληξαν στο συμπέρασμα πως όσο περισσότερο οι νέοι χρησιμοποιούσαν το Facebook, τόσο χειρότερα ένιωθαν στη συνέχεια και τόσο περισσότερο το επίπεδο ικανοποίησής τους από τη ζωή ήταν πολύ χαμηλό. Επίσης, το Facebook είναι ένα από τα δημοφιλέστερα sites για ανέβασμα φωτογραφιών με πάνω από 14 εκατομμύρια φωτογραφίες καθημερινά. Με αφορμή τη δημοτικότητά του, το Facebook έχει υποστεί κριτική και έχει κατηγορηθεί σε θέματα που αφορούν τα προσωπικά δεδομένα και τις πολιτικές απόψεις των ιδρυτών του. Ωστόσο η συγκεκριμένη ιστοσελίδα παραμένει η πιο διάσημη κοινωνική περιοχή δικτύωσης σε πολλές αγγλόφωνες χώρες.

Ο τρόπος επικοινωνίας μεταξύ των μελών μπορεί να είναι ατομικός, δηλαδή ανάμεσα σε δύο άτομα, ή μαζικός, όπου τα μέλη δημοσιεύουν περιεχόμενο ορατό σε όλα τα υπόλοιπα μέλη. Επίσης, δίνεται η δυνατότητα δημιουργίας κλειστών ομάδων με μέλη που δέχονται πρόσκληση για να εισέλθουν, έχοντας κοινά ενδιαφέροντα. Αυτό εξυπηρετεί ομάδες όπως είναι μαθητές από διάφορα σχολεία, παλιούς ενήλικες πια συμμαθητές, ενήλικες με κοινό επαγγελματικό προσανατολισμό και ενδιαφέροντα. Μπορεί να περιλαμβάνει δημοσιεύσεις μηνυμάτων ή ακόμα και ζωντανές συνομιλίες, σχόλια, νήματα συζητήσεων κλπ. Το περιεχόμενο που δημοσιεύουν ή ανταλλάσσουν μεταξύ τους τα μέλη του ιστότοπου μπορεί να είναι κείμενο, φωτογραφίες, ή και βίντεο.

Στόχος τους είναι προσελκύσουν απλούς χρήστες του Διαδικτύου οι οποίοι θα γίνουν αποδέκτες των προσωπικών του απομεινών, ή θα εμπλακούν σε διάλογο μαζί τους. Μπορούν να χαρακτηριστούν ως “ερασιτέχνες” δημοσιογράφοι ή συγγραφείς.

Οι πιο δημοφιλείς ιστότοποι κοινωνικής δικτύωσης σήμερα είναι οι εξής:

- Facebook
- MySpace: Είναι μια υπηρεσία κοινωνικής δικτύωσης, η οποία ανήκει στην εταιρεία Specific Media LLC και στον ποπ σταρ Τζάστιν Τίμπερλεϊκ. Ξεκίνησε τον Αύγουστο του 2003 και εδρεύει στο Μπέμπερλι Χιλς, στην Καλιφόρνια.
- Youtube: Είναι ένας δημοφιλής διαδικτυακός τόπος, ο οποίος επιτρέπει αποθήκευση, αναζήτηση και αναπαραγωγή ψηφιακών ταινιών. Ξεκίνησε το Φεβρουάριο του 2005 ενώ τον Οκτώβριο του 2006, η εταιρεία αγοράστηκε από την Google.
- Blogger: Ο αγγλικός όρος blog προέρχεται από το weblog το οποίο έχει μεταφραστεί στα ελληνικά ως ιστολόγιο. Το weblog είναι επινόηση του Jorn Barger από τον Δεκέμβριο του 1997. Στις αρχές του 1999 όμως ο Peter Merholz ανακοίνωσε ότι θα το προφέρει «wee-blog» και έτσι καταλήξαμε στο σύντομο όρο blog αφού ο συντάκτης αναφερόταν ως blogger.
- Wordpress: είναι ελεύθερο και ανοικτού κώδικα λογισμικό ιστολογίου και πλατφόρμα δημοσιεύσεων, γραμμένο σε PHP και MySQL. Κυκλοφόρησε για πρώτη φορά στις 27 Μαΐου 2003, από τον Matt Mullenweg ως παραλλαγή του b2/cafelog. Από τον Φεβρουάριο του 2011, η έκδοση 3.0 έχει ληφθεί περισσότερες από 32,5 εκατομμύρια φορές.
- Flickr: Είναι μια ιστοσελίδα, η οποία δημιουργήθηκε για να φιλοξενεί φωτογραφίες και βίντεο. Δημιουργήθηκε αρχικά από την εταιρία Ludicorp και ύστερα εξαγοράστηκε από

την Yahoo!. Η υπηρεσία χρησιμοποιείται συχνά από bloggers για να ενσωματώσουν τις φωτογραφίες τους στα blogs τους. Τον Σεπτέμβριο του 2010, το Flickr έφτασε τα 5 δις. φωτογραφιών. Αρκετές από τις φωτογραφίες του Flickr κυκλοφορούν υπό την άδεια Creative Commons.

- Twitter: Είναι ένας ιστοχώρος κοινωνικής δικτύωσης που επιτρέπει στους χρήστες του να στέλνουν και να διαβάζουν σύντομα μηνύματα (μέχρι 140 χαρακτήρες), τα οποία ονομάζονται τουίτς (Tweets). Τα μηνύματα μπορούν να αναγνωστούν και από μη συνδεδεμένους χρήστες, αλλά μόνο οι συνδεδεμένοι μπορούν να δημοσιεύσουν κείμενα. Δημιουργήθηκε τον Μάρτιο του 2006 από τον Τζακ Ντόρσει και δημοσιεύθηκε τον Ιούλιο του ίδιου χρόνου. Η υπηρεσία έγινε γρήγορα δημοφιλής και σήμερα έχει 200 εκατομμύρια ενεργούς χρήστες.
- LinkedIn: Είναι ένας ιστοχώρος επαγγελματικής κοινωνικής δικτύωσης. Ιδρύθηκε τον Δεκέμβριο του 2002 από τον Ρέιντ Χόφμαν, αλλά ξεκίνησε επίσημα στις 5 Μαΐου του 2003. Η έδρα της εταιρίας είναι στη Σίλικον Βάλεϊ και έχει γραφεία σε όλο τον κόσμο. Τα εγγεγραμμένα μέλη του έχουν τη δυνατότητα να δημιουργήσουν το προσωπικό επαγγελματικό τους προφίλ, να συνδεθούν με άλλους χρήστες, να αναζητήσουν εργασία, αλλά και να δημιουργήσουν πελατολόγιο.

1.3. Πλεονεκτήματα κοινωνικών δικτύων

Η χρήση online κοινωνικών δικτύων προσφέρει στους χρήστες τους πολλά οφέλη και πλεονεκτήματα. Στην περίπτωση του social networking, τα οφέλη είναι πολλά, ωστόσο μόνο εάν οι οργανισμοί σκεφτούν δημιουργικά για το πλαίσιο χρήσης των social media θα καταφέρουν να αποκομίσουν αυτά τα οφέλη, αποκτώντας διαρκές ανταγωνιστικό πλεονέκτημα. Τα οφέλη τους παρουσιάζονται παρακάτω:

- η δυνατότητα δημιουργίας δεσμών με πολύ μεγάλο αριθμό ατόμων, εφόσον το διαδίκτυο συγκεντρώνει πλήθος άτομα από όλο τον κόσμο,
- η δυνατότητα δημιουργίας δεσμών με άτομα που μπορεί να βρίσκονται σε μεγάλη γεωγραφική απόσταση το ένα από το άλλο, εφόσον το Διαδίκτυο καταργεί τις αποστάσεις. Αυτή η δυνατότητα δεν υπάρχει στα στενά κοινωνικά δίκτυα του άμεσου πραγματικού περιβάλλοντος του ατόμου,
- η δυνατότητα δημιουργίας μεγάλης ποικιλίας κοινωνικών δεσμών εφόσον στο Διαδίκτυο συρρέουν άτομα από διαφορετικές χώρες, κοινωνίες, πολιτισμούς και με διαφορετικές συνήθειες και χαρακτηριστικά,
- η δυνατότητα επιλογής ανάμεσα σε μεγάλο πλήθος κοινωνικών ομάδων και η αναζήτηση της ομάδας που εκφράζει και ωφελεί το άτομο με τον καλύτερο δυνατό τρόπο,
- η δυνατότητα επαφής με πολλούς διαφορετικούς πολιτισμούς,
- η δυνατότητα διεύρυνσης των γνώσεων και των πνευματικών οριζόντων του ατόμου,
- η δυνατότητα αναζήτησης και ανεύρεσης περιεχομένου (φωτογραφιών, βίντεο κλπ) στο οποίο οι χρήστες δε θα μπορούσαν να έχουν πρόσβαση με διαφορετικό τρόπο,
- η δυνατότητα άμεσης και από πρώτο χέρι ενημέρωσης για οτιδήποτε συμβαίνει στον κόσμο, εφόσον οι ειδήσεις μεταδίδονται μεταξύ των χρηστών του Διαδικτύου από τη μια άκρη της γης στην άλλη, σε ελάχιστο χρόνο, και
- η δυνατότητα εύκολης πρόσβασης σε ψυχαγωγικό περιεχόμενο, από το σπίτι, μέσω π.χ. της παρακολούθησης βίντεο και μέσω της ενασχόλησης με εφαρμογές ψυχαγωγικού χαρακτήρα όπως τα παιχνίδια.

Η χρήση των εργαλείων κοινωνικής δικτύωσης, όπως το Facebook και το Twitter, ως πλατφόρμων συνεργασίας συνδέει τις επιχειρήσεις με τον υπόλοιπο κόσμο με δεκάδες διαφορετικούς τρόπους. Τα εν λόγω εργαλεία φέρνουν την τεχνολογία σε επαφή με τις επιχειρήσεις, συνδέουν τους ανθρώπους με τις πληροφορίες, καθιερώνουν πιθανούς νέους δρόμους προς την αγορά και βελτιώνουν την επικοινωνία με τους πελάτες και τη διάδοση του εμπορικού σήματος. Τα παραπάνω ευρήματα καταδεικνύουν ότι ο κόσμος των επιχειρήσεων

βρίσκεται στα πρόωρα στάδια της υιοθέτησης αυτών των εργαλείων καθώς και στη διαδικασία υιοθέτησης σημαντικών προκλήσεων - όπως είναι η ανάγκη για διαχείριση των εργαλείων αυτών και η συμμετοχή του IT, που μπορεί να επηρεάσουν την ολοκλήρωση και υιοθέτηση των νέων πλατφορμών και τεχνολογιών.

Όσο αφορά στην επιχειρηματική στρατηγική, τα social media χρησιμοποιούνται ως μέσα δημιουργίας εταιρικής εικόνας, ενημέρωσης, επικοινωνίας και ανάπτυξης σχέσεων με τους πελάτες και έχουν απίστευτη δυναμική.

Οι εταιρείες που χρησιμοποιούν σωστά τα social networks μπορούν να «χτίσουν» την εικόνα τους, να αναπτύξουν τις δημόσιες σχέσεις τους καθώς και να δημιουργήσουν και ή να επηρεάσουν θετικά τις συζητήσεις που γίνονται γύρω από το brand τους «εκτοξεύοντας» την αναγνωσιμότητα και αξιοπιστία τους.

Οι εφαρμογές κοινωνικής δικτύωσης δημιουργούν ένα σημαντικό αριθμό ευκαιριών και προκλήσεων στον επιχειρησιακό κόσμο. Τα τελευταία χρόνια, παρατηρείται μια έκρηξη τεχνολογικών εφαρμογών που βασίζονται στη λογική της Κοινωνικής Δικτύωσης στα πλαίσια του Παγκόσμιου Ιστού, γεγονός που δε μπορεί παρά να επηρεάζει την επιχειρηματική δραστηριότητα στο σύνολό της. Οι νέες τεχνολογίες, όπως τα ιστολόγια (blogs), τα wikis, η κοινωνική σήμανση (tagging), οι ιστοχώροι κοινωνικής δικτύωσης, δημιουργούν ευκαιρίες για νέους τρόπους ενδοεταιρικής συνεργασίας και διαχείρισης της γνώσης, αλλάζουν το τοπίο στην παροχή υπηρεσιών αλλά και στις διεπιχειρησιακές ανταλλαγές, ενώ παράλληλα αναδιαμορφώνουν τις υπάρχουσες επιχειρησιακές εφαρμογές.

Τα κύρια πλεονεκτήματα που προσφέρουν οι εφαρμογές κοινωνικής δικτύωσης σε μια επιχείρηση είναι ποικίλα. Αναλυτικότερα, κύριο πλεονέκτημα είναι η δυνατότητα που προσφέρεται στην επιχείρηση να απευθυνθεί σε αγορά τεράστιου μεγέθους, χωρίς γεωγραφικούς περιορισμούς. Επιπλέον η τμηματοποίηση της αγοράς με βάση διάφορα κριτήρια όπως γεωγραφικά, κοινωνικά, δημογραφικά, εθνικά, θρησκευτικά γίνεται πολύ πιο εύκολα, ενώ η «πληροφορία» μεταφέρεται εύκολα και με γρήγορους ρυθμούς.

Οι εφαρμογές κοινωνικής δικτύωσης δίνουν τη δυνατότητα σε μια επιχείρηση να προσφέρει συνεχώς κίνητρα στους καταναλωτές, γεγονός που αυξάνει την πιστότητα στα προϊόντα και υπηρεσίες της. Βασικό πλεονέκτημα είναι και η συνεχής και εύκολη ανατροφοδότηση σχετικά με τη συμπεριφορά και τον βαθμό ικανοποίησης των καταναλωτών, γεγονός που διευκολύνει την έρευνα και συμβάλλει στην ανάπτυξη της επιχείρησης.

Επίσης, τα δίκτυα κοινωνικής δικτύωσης κάνουν εφικτή την παρουσίαση της επιχείρησης 24 ώρες του εικοσιτετράωρου όλο τον χρόνο, μειώνουν το λειτουργικό κόστος, ενώ τέλος διευκολύνεται η εύρεση νέου προσωπικού.

Υπάρχουν πολλές μεγάλες εταιρίες που αυτή τη στιγμή απολαμβάνουν και εκμεταλλεύονται τα οφέλη του social media marketing. Χαρακτηριστικό είναι το παράδειγμα της Coca-Cola, που προσέγγισε το κοινό της μέσα από τα κοινωνικά δίκτυα, το photo και video sharing.

Είναι σημαντικό να σημειωθεί το 88% των εταιριών μέχρι το 2015 θα χρησιμοποιούν εργαλεία κοινωνικής δικτύωσης στο μάρκετινγκ, ενώ πολλές από αυτές χρησιμοποιούν τα κοινωνικά μέσα αποκλειστικά ως μάρκετινγκ εργαλεία.

Επίσης, η χρήση των εργαλείων κοινωνικής δικτύωσης, όπως το Facebook και το Twitter, ως πλατφόρμων συνεργασίας συνδέει τις επιχειρήσεις με τον υπόλοιπο κόσμο με δεκάδες διαφορετικούς τρόπους. Τα εν λόγω εργαλεία φέρνουν την τεχνολογία σε επαφή με τις επιχειρήσεις, συνδέουν τους ανθρώπους με τις πληροφορίες, καθιερώνουν πιθανούς νέους δρόμους προς την αγορά και βελτιώνουν την επικοινωνία με τους πελάτες και τη διάδοση του εμπορικού σήματος. Τα παραπάνω ευρήματα καταδεικνύουν ότι ο κόσμος των επιχειρήσεων βρίσκεται στα πρόωρα στάδια της υιοθέτησης αυτών των εργαλείων καθώς και στη διαδικασία υιοθέτησης σημαντικών προκλήσεων - όπως είναι η ανάγκη για διαχείριση των εργαλείων αυτών και η συμμετοχή του IT, που μπορεί να επηρεάσουν την ολοκλήρωση και υιοθέτηση των νέων πλατφορμών και τεχνολογιών.

1.4. Μειονεκτήματα κοινωνικών δικτύων

Η κοινωνική δικτύωση έχει ένα σοβαρό μειονέκτημα. Η υπερβολική χρήση, τις περισσότερες φορές οδηγεί στον εθισμό, πράγμα ανησυχητικό όχι μόνο για τους γονείς των παιδιών αλλά και για τους εργοδότες των εργαζομένων.

Οι επιπτώσεις από τη χρήση των κοινωνικών δικτύων είναι και αρνητικές και αφορούν ποικίλους τομείς της ζωής του ανθρώπου, όπως η εργασία, η ψυχολογία και η ιδιωτικότητα.

Καθένας που έχει ιστοσελίδα κοινωνικής δικτύωσης χάνει τον έλεγχο όλων των προσωπικών του στοιχείων που περιλαμβάνονται σ' αυτή. Η εταιρεία πάροχος έχει δικαίωμα χρήσης όλων αυτών των πληροφοριών και διατήρησής τους στο διηνεκές αυτά περιλαμβάνονται στη σύμβαση «άνοιγμα σελίδας» εδώ στηρίζεται και το γενικότερο εμπόριο των προσωπικών πληροφοριών των εγγεγραμμένων.

Εύκολη και διαδεδομένη είναι η αυθαίρετη κατασκευή από τρίτους λογαριασμών για γνωστά δημόσια πρόσωπα, με σκοπό την παραπλάνηση και λασπολογία. Επίσης, είναι δυνατή και η κλοπή γνήσιας διαδικτυακής ταυτότητας από κακοπροαίρετους χάκερς και χρήση της για διάπραξη παρενοχλήσεων ακόμη και εγκλημάτων. Εξίσου πολλές είναι τα τελευταία χρόνια οι περιπτώσεις αποκάλυψης ότι τα εικονικά κοινωνικά δίκτυα λειτούργησαν ως παγίδα για τα μέλη τους, στην επαγγελματική τους ζωή. Εργοδότης απολύει ή απορρίπτει αιτήσεις για πρόσληψη με βάση τα στοιχεία για τον υποψήφιο που αλίευσε από τις σελίδες κοινωνικής δικτύωσης. Το πάτημα του κουμπιού αποστολής προσωπικών δεδομένων στο Διαδίκτυο μπορεί να έχει σοβαρές συνέπειες.

Επιπλέον κίνδυνοι των κοινωνικών δικτύων είναι:

- η έκθεση σε πολύ μεγαλύτερο αριθμό κινδύνων κοινωνικού χαρακτήρα. Ενώ θα περίμενε κανείς ότι το Διαδίκτυο λόγω της απόστασης που εμπεριέχει, θα προφύλασσε τα άτομα από την επαφή με πραγματικούς κινδύνους, είναι πολλές οι περιπτώσεις όπου ιδιαίτερα οι νέοι βιώνουν έντονα καταστάσεις μέσα από την χρήση του κυβερνοχώρου, οι οποίες είναι τραυματικές για την ψυχική τους υγεία εξαιτίας της χρήσης online κοινωνικών δικτύων και της επαφής τους με πλήθος αγνώστων ατόμων,
- η επιβλαβής έκθεση της προσωπικής ζωής του ατόμου. Τα online κοινωνικά δίκτυα είναι χώροι όπου μπορεί να συγκεντρώνεται πλήθος διαφορετικών και άγνωστων ατόμων,
- η καταχώρηση και δημοσίευση προσωπικών στοιχείων σε αυτά, καθιστά τους χρήστες τους ευάλωτους σε πολύ μεγαλύτερο αριθμό ατόμων, συχνά με ανεξέλεγκτες συνέπειες,
- η επαφή και η διασύνδεση με παντελώς αγνώστους ενέχει περισσότερους κινδύνους απ' ότι η επαφή με αγνώστους στον πραγματικό κόσμο, γιατί δεν υπάρχει η αμεσότητα και η δυνατότητα ανάπτυξης κριτικής σκέψης και αντίδρασης,
- η παρενόχληση από άτομα εντελώς άγνωστα μέλη του κοινωνικού δικτύου χωρίς τη δυνατότητα προστασίας ή αντιμετώπισης τέτοιων ενεργειών,
- η κλοπή των προσωπικών δεδομένων και γενικά η απειλή της ασφάλειας των προσωπικών δεδομένων των χρηστών τους,
- η εύκολη μετάδοση και επαφή με ηλεκτρονικό περιεχόμενο που δεν είναι ασφαλές για τους χρήστες του, όπως κακόβουλο λογισμικό, πορνογραφικό υλικό, υλικό προσηλυτισμού, υλικό εθισμού σε επικίνδυνες ουσίες, υλικό με θέμα τη βία, ρατσιστικό περιεχόμενο κ.ά., και
- ο βομβαρδισμός με διαφημιστικά μηνύματα από τις σελίδες του ιστότοπου.

Ο διαρκής κίνδυνος που κρύβει η χρήση των social networking sites είναι οι εξωτερικές «εισβολές» στα εταιρικά IT δίκτυα. Τέτοιες επιθέσεις μπορεί να οδηγήσουν σε απώλεια απόρρητων πληροφοριών και να δημιουργήσουν προβλήματα στα συστήματα πληροφορικής της εταιρείας. Ωστόσο, δεν γνωρίζουμε με σαφήνεια πόσο συχνά ευθύνεται η χρήση μέσων κοινωνικής δικτύωσης για τέτοιες επιθέσεις και το είδος της ζημιάς που μπορεί να προκληθεί. Τα εταιρικά συστήματα ασφάλειας δικτύων μπορούν σίγουρα να συμβάλουν στον περιορισμό

τέτοιων κινδύνων, μέσω της φραγής πρόσβασης σε συγκεκριμένα sites. Και πάλι όμως, μεγάλο μέρος του κινδύνου οφείλεται στη μη ασφαλή χρήση του internet από τους εργαζομένους. Μόνο αλλάζοντας αυτές τις συμπεριφορές θα μπορέσουν οι εταιρείες να κάνουν ένα σημαντικό βήμα προς την ενίσχυση της ασφάλειας.

Είναι γεγονός ότι η κοινωνική δικτύωση εγείρει ζητήματα και ανησυχίες όσον αφορά τις επιπτώσεις της στην παραγωγικότητα, τη φήμη και την ασφάλεια των επιχειρήσεων. Ωστόσο, αυτό δεν σημαίνει ότι οι επιχειρήσεις δεν θα πρέπει να αναπτύξουν και να εφαρμόσουν συγκεκριμένες κατευθυντήριες γραμμές σχετικά με τη χρήση και την κατάχρηση της κοινωνικής δικτύωσης. Με τη διαφορά ότι αυτές οι κατευθυντήριες γραμμές δεν θα πρέπει να δίνουν έμφαση στον έλεγχο της συμπεριφοράς των εργαζομένων όσον αφορά την κοινωνική δικτύωση, αλλά στον προσανατολισμό της χρήσης προς θετικές, δημιουργικές κατευθύνσεις, υπέρ των οργανισμών και των εργαζομένων. Τρόπος να γυρίσουμε στην προ social media εποχή δεν υπάρχει. Μπορούμε όμως να αναδείξουμε την προστιθέμενη αξία τους για τους οργανισμούς και να προωθήσουμε την αποτελεσματική τους χρήση.

Τα κυριότερα μειονεκτήματα των εφαρμογών κοινωνικής δικτύωσης είναι:

- κόστος χρόνου που συνεπάγεται η διαδικασία ενημέρωσης των πελατών, η δημιουργία και επεξεργασία πληροφοριών,
- μη αποδοχή της νέας εφαρμογής από το προσωπικό της εταιρείας λόγω έλλειψη γνώσεων και δεξιοτήτων,
- μη ασφαλές περιβάλλον κατά την δημοσίευση πληροφοριών στο Διαδίκτυο, και
- ελεύθερη δημοσίευση κριτικών από πελάτες και ανταγωνιστές.
- η χρήση των κοινωνικών δικτύων σε ώρα εργασίας επηρεάζει την απόδοση των υπαλλήλων. Μάλιστα, το φαινόμενο αυτό οδηγεί όχι μόνο σε αντιπαραγωγικότητα, αλλά και σε μεγάλη χρηματική ζημία. Επίσης ένα μεγάλο ποσοστό μπαίνει στον λογαριασμό του έστω και για λίγο, αν και ελάχιστοι παραδέχονται ότι αυτό επηρεάζει την ποιότητα της εργασίας τους. Δύο στους τρεις υποστηρίζουν ότι δεν πρέπει να διακοπεί η επαφή με τα κοινωνικά δίκτυα εν ώρα εργασίας, ενώ οι υπόλοιποι το θεωρούν αναγκαίο.

1.5. Επιχειρηματικές ευκαιρίες και κοινωνικά δίκτυα

Τα κοινωνικά δίκτυα σήμερα πρέπει να θεωρηθούν από τις επιχειρήσεις ένα μέσο για την προσέλκυση νέων πελατών αλλά και προβολής. Οι ευκαιρίες για της επιχειρήσεις μπορεί να είναι οι παρακάτω:

- Ένα χρήσιμο μέσο για προβολή, προώθηση προϊόντων online και προσέλκυση νέων πελατών-συνεργατών:
- Μπορεί να προωθήσει προϊόντα η υπηρεσίες με επιλεκτική διαφήμιση και viral marketing,
- Δυνατότητα ανατροφοδότησης και αλληλοεπίδρασης για τα προϊόντα η υπηρεσίες ειδικά δε μέσω άλλων συνδέσμων η δικτυακών τόπων.
- Ανακάλυψη νέων συνεργατών, νέα στρατηγική επικοινωνίας, διαρκής ενημέρωση.
- Μία ευκαιρία να δημιουργηθεί ένα εσωτερικό δίκτυο για την εκμετάλλευση πληροφοριών και γνώσης:
- Διαχείριση του γνωστικού ενεργητικού της επιχείρησης,
- Αποτελεσματικός τρόπος «εξόρυξης» εσωτερικής εμπειρογνωμοσύνης,
- Μείωση του χρόνου αποστολής αλληλογραφίας και εσωτερικής επικοινωνίας, διοικητικά πλεονεκτήματα,
- Ενεργοποίηση της συμμετοχής υπαλλήλων, νέων στελεχών, συνταξιούχων, και άλλων συνεργατών,
- Αναγνωσιμότητα και status,
- Ανοιχτή επαφή με το κοινό, και
- Ανταγωνιστικότητα.

2. Συστήματα συστάσεων (Recommendation systems)

2.1. Γενικά

Η εισροή των δεδομένων σήμερα είναι όλο και μεγαλύτερη στον παγκόσμιο ιστό ενώ ο όγκος πληροφορίας την οποία δέχεται ο χρήστης αυξάνεται εκθετικά. Για τον λόγο αυτό τα συστήματα συστάσεων αποκτούν ακόμη μεγαλύτερη σημασία. Ο σκοπός της ανάπτυξης τέτοιων συστημάτων είναι η επιλεκτική προώθηση πληροφορίας στο χρήστη με κύριο γνώμονα την προσωπικότητα που παρουσιάζεται από το προφίλ του. Μέσα από τις επιλογές του χρήστη, το σύστημα αποκτάει γνώση για τις προτιμήσεις του και εκμεταλλευόμενο τη γνώση αυτή μπορεί να του κάνει «κατάλληλες» συστάσεις προσεγγίζοντας τις προτιμήσεις του, ειδικευόμενο στα ενδιαφέροντά του.

Αυτό πραγματοποιείται με την χρήση των κατάλληλων αλγορίθμων δηλαδή των κατάλληλων τεχνικών. Έμμεσα, υπονοείται έτσι η δημιουργία εικονικών κοινοτήτων, δηλαδή ομάδων χρηστών με παρόμοιο προφίλ. Οι αλγόριθμοι συστάσεων είναι γνωστοί για τη χρήση τους στις εφαρμογές ομαδοποίησης προτιμήσεων χρηστών, όπου χρησιμοποιούν την εισαγωγή δεδομένων ενός πελάτη για τα ενδιαφέροντά του, προκειμένου να παραχθεί ένας κατάλογος συνιστώμενων στοιχείων. Στην παρούσα διπλωματική εργασία υλοποιήθηκε η ιστοσελίδα με το όνομα exodus, όπου με τη χρήση των κατάλληλων συστημάτων συστάσεων καθορίζονται προβλέψεις στις προτιμήσεις χρηστών όσων αφορά τον τρόπο διασκέδασης τους.

Με αυτή την προσέγγιση η ιστοσελίδα αλλάζει ριζικά βασισμένη στα ενδιαφέροντα του χρήστη. Αλγόριθμοι συστάσεων στα Κοινωνικά Δίκτυα λειτουργούν συχνά σε περιβάλλοντα όπου:

- Οι νέοι χρήστες έχουν εξαιρετικά περιορισμένες πληροφορίες, βασισμένες μόνο σε μερικές εκτιμήσεις
- Οι παλαιότεροι χρήστες μπορούν να έχουν ένα πλεόνασμα πληροφοριών, με βάση τις προηγούμενες εκτιμήσεις που έχουν κάνει (rates)
- Τα δεδομένα των χρηστών είναι ευμετάβλητα: Κάθε αλληλεπίδραση παρέχει πολύτιμα δεδομένα και ο αλγόριθμος πρέπει να αποκριθεί αμέσως στην αξιολόγηση των νέων πληροφοριών.

Οι βασικές προσεγγίσεις στην επίλυση του προβλήματος παροχής συστάσεων είναι τρεις: Η παραδοσιακή προσέγγιση του συνεργατικού φιλτραρίσματος (collaborative filtering), τα πρότυπα ομαδοποίησης (cluster models) και οι μέθοδοι σύστασης βασισμένες στη γνώση (knowledge-based recommendation).

Η διαδικασία διάχυσης της γνώσης έχει ιδιαίτερη σημασία σε κάθε τομέα της κοινωνικής ζωής. Ο ρυθμός της κοινωνικής προόδου εξαρτάται σε μεγάλο βαθμό από την διαθεσιμότητα και την ποιότητα των μέσων επικοινωνίας, και της διακίνησης γνώσης. Επομένως, η ύπαρξη και διατήρηση αποτελεσματικών μέσων συγκέντρωσης και διάχυσης γνώσης και πληροφορίας είναι κρίσιμος παράγοντας για την πρόοδο της κοινωνίας.

Στο επίπεδο του μεμονωμένου ατόμου, η γενική συγκεντρωμένη γνώση, πολλές φορές δεν αρκεί. Για να μπορέσει ένα άτομο να λάβει σωστές αποφάσεις για ότι τον αφορά, χρειάζεται πρόσθετες και εξειδικευμένες πληροφορίες που αφορούν στο συγκεκριμένο ζήτημα. Στις προϊστορικές κοινωνίες, όπου το συνολικό ποσό της συσσωρευμένης γνώσης ήταν ελάχιστο, και οι περιπτώσεις όπου απαιτείτο η λήψη περίπλοκων αποφάσεων ήταν σπάνιες, η ανάκτηση των εξειδικευμένων πληροφοριών δεν παρουσίαζε ιδιαίτερες δυσκολίες. Στις σύγχρονες, εξελιγμένες κοινωνίες όμως, το άτομο βρίσκεται συνεχώς αντιμέτωπο με προκλήσεις όπου απαιτείται λήψη αποφάσεων για περίπλοκα ζητήματα. Επίσης, στην σύγχρονη κοινωνία υπάρχει μία τεράστια ποσότητα συσσωρευμένης γνώσης και ποικιλία πιθανών επιλογών και προτιμήσεων. Για το λόγο αυτό, ο σύγχρονος άνθρωπος στρέφεται στο στενό ή ευρύτερο κοινωνικό περίγυρο προκειμένου να προσλάβει τις απαραίτητες εξειδικευμένες πληροφορίες. Οι πληροφορίες αυτές έχουν συνήθως την μορφή συστάσεων – προτάσεων. Σε τέτοιες περιπτώσεις, ο κοινωνικός περίγυρος του ατόμου λειτουργεί ως «φίλτρο» που επιτρέπει στις σχετικές πληροφορίες να φτάσουν στον ενδιαφερόμενο, μπλοκάροντας ταυτόχρονα τον όγκο των άχρηστων – στην συγκεκριμένη περίπτωση – προτιμήσεων. Αυτή η μέθοδος απαιτεί από τον ενδιαφερόμενο να καταβάλλει προσπάθεια, ζητώντας ενεργά πληροφορίες από το περιβάλλον του.

Στο σύγχρονο περιβάλλον, τα ενδιαφέροντα ενός μέσου ανθρώπου και οι τομείς της γνώσης στους οποίους υποχρεώνεται να εντρυφήσει πολλαπλασιάζονται συνεχώς, δημιουργώντας μεγάλη πίεση χρόνου. Οι απαραίτητες πληροφορίες πρέπει να ανευρίσκονται σύντομα και μάλιστα χωρίς να υπάρχει η ανάγκη ο ενδιαφερόμενος να περιπλανηθεί σε έναν λαβύρινθο άχρηστων στοιχείων, προκειμένου να βρει αυτά που τον ενδιαφέρουν. Η πίεση χρόνου και η ανάγκη για μεγαλύτερη αποτελεσματικότητα γίνεται ιδιαίτερα εμφανής στον τομέα των αγορών. Ο σύγχρονος άνθρωπος μπορεί να διασκεδάζει χρησιμοποιώντας ολοένα και περισσότερο τα κοινωνικά δίκτυα και οι προτιμήσεις του να ομαδοποιούνται. Η έρευνα προτιμήσεων είναι ένα σημείο ολοένα και πιο κρίσιμο, λόγω του πλήθους των ομοειδών καταστημάτων διασκέδασης και συγκέντρωσης που συναντά πλέον ο σύγχρονος άνθρωπος.

Εξηγείται έτσι ως έναν βαθμό η τεράστια άνοδος και ανάπτυξη που παρουσιάζεται στη χρήση του Διαδικτύου στις σύγχρονες κοινωνίες. Μέσω του Διαδικτύου μπορεί κανείς να αποκτήσει πρόσβαση σε ένα πλήθος πληροφοριών κάθε είδους, καθώς και πρόσβαση σε μία τεράστια παγκόσμια ενημέρωση των καταστημάτων διασκέδασης και τις αντίστοιχης προσφοράς υπηρεσιών. Η τεράστια, όμως έκταση του Διαδικτύου και των πληροφοριών που αυτό περιέχει μπορεί να αποτελέσει μειονέκτημα όταν ο χρήστης πρέπει να βρει γρήγορα την πληροφορία ή το προϊόν που επιθυμεί. Όμως ακόμα και όταν υπάρχει επάρκεια χρόνου, η εξέταση και αξιολόγηση ενός τεράστιου αριθμού ομοειδών καταστημάτων είναι μία διαδικασία αυξημένης πολυπλοκότητας. Επιπλέον, η λύση αυτού του προβλήματος αξιολόγησης δεν είναι πάντα εφικτή λόγω των διαφορετικών προσωπικοτήτων και συνηθειών κάτι που αποτελεί περισσότερο τροχοπέδη στην λήψη αποφάσεων.

Προς βοήθεια των χρηστών του Διαδικτύου που αντιμετωπίζουν τέτοια προβλήματα έρχονται τα Πληροφοριακά Συστήματα Συστάσεων (Recommendation Systems). Τα συστήματα αυτά, λύνουν τέτοιου είδους προβλήματα παρέχοντας στους χρήστες τους αξιόπιστες συστάσεις για διάφορα προϊόντα, βιβλία, κείμενα, πληροφορίες, καταστήματα κ.α. Τα Συστήματα Συστάσεων καταλήγουν σε συστάσεις μέσω των τεχνικών που θα αναλύσουμε στις επόμενες παραγράφους και υποκαθιστούν ως έναν βαθμό τη διαδικασία διάχυσης γνώσης μέσω του κοινωνικού περιγύρου, σε συγκεκριμένες κατηγορίες προβλημάτων. Τα Συστήματα Συστάσεων είναι βασισμένα συνήθως στο Διαδίκτυο και παρέχουν συμβουλές σε διαδικτυακούς χρήστες σχετικά με προϊόντα ή υπηρεσίες, τα οποία επιθυμούν να αγοράσουν ή να χρησιμοποιήσουν. Οι συστάσεις που παρέχονται μέσω των συστημάτων αυτών μπορούν να κατευθύνουν τους χρήστες μέσα σε μεγάλο όγκο πληροφοριών. Στόχος της εισαγωγής και ενσωμάτωσης διαφόρων τύπων Συστημάτων Συστάσεων σε διάφορα Κοινωνικά Δίκτυα, είναι η επιτέλεση μιας πιο ακριβούς αξιολόγησης των διαφόρων υπάρχοντων εναλλακτικών επιλογών και επομένως ο εμπλουτισμός των προτάσεων που παρέχονται στους χρήστες.

Ιστορικά, το πρώτο σύστημα συστάσεων ήταν το «Tapestry». Οι δημιουργοί του εισήγαγαν για πρώτη φορά τον όρο Συνεργατικό Φιλτράρισμα (Collaborative Filtering) για να περιγράψουν την δυνατότητα σύγκρισης των αξιολογήσεων ενός χρήστη με τις αξιολογήσεις των άλλων, με στόχο την παραγωγή συστάσεων οι οποίες να ικανοποιούν τις προτιμήσεις του. Το σύστημα αυτό λειτουργούσε στα πλαίσια μίας σχετικά μικρής ομάδας ανθρώπων με ύπαρξη διαπροσωπικών σχέσεων ανάμεσα στα μέλη της, όπως οι εργαζόμενοι σε ένα γραφείο ή μία επιχείρηση. Γενικά, τα συστήματα συστάσεων χρησιμοποιούν διάφορες τεχνικές για να συλλέξουν πληροφορίες σχετικά με τους χρήστες (προτιμήσεις, προτεραιότητες κ.α.) και τις εναλλακτικές επιλογές – προϊόντα, τα οποία βρίσκονται στην διάθεση των χρηστών (τεχνικά χαρακτηριστικά, βαθμολογία προερχόμενη από άλλους χρήστες, κ.α.). Τα στοιχεία αυτά αξιοποιούνται με σκοπό να συσχετιστούν οι χρήστες με αυτά τα αντικείμενα – προτάσεις τα οποία είναι πιο πιθανό να καλύπτουν τα ενδιαφέροντά τους. Η αξιοποίηση των στοιχείων αυτών και η εξαγωγή των συστάσεων, γίνονται μέσω της χρήσης κατάλληλων αλγορίθμων.

2.2. Κατηγορίες συστημάτων συστάσεων

Υπάρχουν τρεις βασικές κατηγορίες που αφορούν τα Συστήματα Συστάσεων και ο διαχωρισμός τους γίνεται με βάση την τεχνική αξιοποίησης στοιχείων και εξαγωγής συστάσεων στην οποία βασίζονται. Οι κατηγορίες αυτές είναι οι εξής:

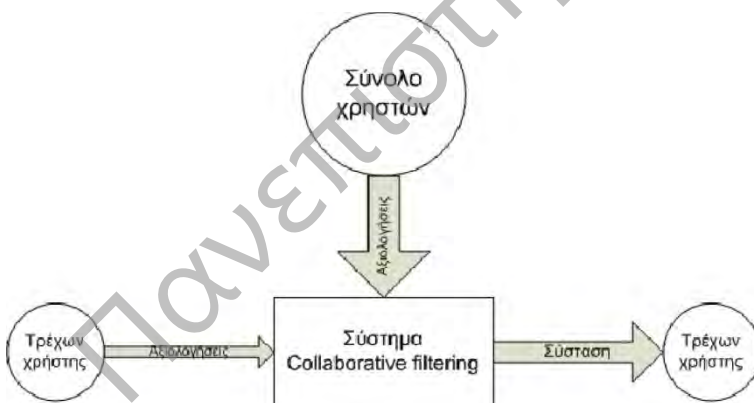
- Συστήματα collaborative – social filtering: Τα συστήματα αυτά βασίζονται στις αξιολογήσεις – βαθμολογήσεις των αντικειμένων από τους χρήστες τους. Τις αξιολογήσεις αυτές τις αξιοποιούν με βάση την εξής αρχή: οι χρήστες οι οποίοι δίνουν παρεμφερείς αξιολογήσεις για ορισμένα αντικείμενα, έχουν παρεμφερή ενδιαφέροντα και ανάγκες, και κατά συνέπεια

θα ενδιαφέρονται για τα ίδια αντικείμενα. Τα συστήματα αυτά δεν έχουν καμία γνώση για τα χαρακτηριστικά των αντικειμένων που προτείνουν, ούτε για τις ανάγκες των χρηστών τους.

- Συστήματα βασισμένα στην γνώση (Knowledge - based Recommendation Systems): Τα συστήματα αυτά δημιουργούν γνώση για να εξάγουν χρήσιμες συστάσεις. Πιο συγκεκριμένα βασίζονται:
 - α) Στην γνώση των χαρακτηριστικών των αντικειμένων που έχουν στην διάθεσή τους.
 - β) Στην γνώση των αναγκών και των απαιτήσεων των χρηστών τους.
 - γ) Στην γνώση σχετικά με το πώς ορισμένα χαρακτηριστικά των αντικειμένων καλύπτουν και ανταποκρίνονται σε συγκεκριμένες ανάγκες των χρηστών.
- Υβριδικά συστήματα: Τα συστήματα που αναφέρθηκαν πιο πάνω έχουν ορισμένα προβλήματα και αδυναμίες - διαφορετικά το καθένα - που θα αναλυθούν στην συνέχεια της εργασίας. Ένας τρόπος να αντιμετωπιστούν αυτά τα προβλήματα, είναι να δημιουργηθεί ένα σύστημα που να αξιοποιεί περισσότερες από μία από τις προαναφερθείσες μεθόδους. Με αυτό τον τρόπο η μία μέθοδος μπορεί να αναρρέσει τις αδυναμίες της άλλης και να προκύψει έτσι ένα πιο αξιόπιστο σύστημα, βέβαια εξαρτάται και των αποτελεσμάτων που θέλουμε να άρουμε.

2.2.1. Συστήματα collaborative filtering

Η τεχνική του Collaborative Filtering (CF) ήταν η πρώτη που χρησιμοποιήθηκε σαν βάση για τα συστήματα συστάσεων (Tapestry, Group Lens). Η τεχνική αυτή δεν βασίζεται σε γνώση των χαρακτηριστικών των αντικειμένων προκειμένου να παράγει συστάσεις για κάποιον χρήστη, αλλά αντίθετα βασίζεται στην ομοιότητα του χρήστη αυτού με άλλους (εικόνα 2-1). Η ομοιότητα προκύπτει από την σύγκριση των αξιολογήσεων του εν λόγω χρήστη με τις αξιολογήσεις άλλων χρηστών. Όταν βρεθούν χρήστες με μεγάλο βαθμό ομοιότητας, το σύστημα προβαίνει σε συστάσεις με βάση την λογική ότι οι χρήστες των οποίων οι αξιολογήσεις μοιάζουν, θα προτιμήσουν τα ίδια αντικείμενα. Στην πραγματικότητα, το σύστημα δεν χρειάζεται να έχει καμία γνώση για τα χαρακτηριστικά ή τις ιδιότητες των αντικειμένων που βρίσκονται στην διάθεσή του. Τα αντικείμενα απεικονίζονται αποκλειστικά μέσω των αξιολογήσεων που δέχονται από τους χρήστες. Η τεχνική αυτή είναι ευρέως χρησιμοποιούμενη στα συστήματα συστάσεων σήμερα. Στις επόμενες παραγράφους θα εξετάσουμε εκτενέστερα την διαδικασία που ακολουθείται.



Εικόνα 2-1. Αναπαράσταση συστήματος Collaborative Filtering

2.2.1.1. Λειτουργία συστημάτων collaborative filtering

Η τεχνική του Collaborative Filtering βασίζεται κυρίως στην συλλογή ενός μεγάλου αριθμού αξιολογήσεων από τους χρήστες για διάφορα αντικείμενα. Οι αξιολογήσεις αυτές μπορεί να έχουν διάφορες μορφές. Η συνηθέστερη είναι η μορφή της άμεσης αξιολόγησης σε μία διακριτή κλίμακα βαθμολογίας (π.χ. από το 1 έως το 10). Δημιουργείται έτσι ένας πίνακας Collaborative Filtering Matrix (CFM) που περιέχει όλες τις αξιολογήσεις που έχουν δοθεί από όλους τους χρήστες για όλα τα αντικείμενα. Ο πίνακας αυτός έχει μέγεθος $m \times n$ όπου m είναι ο

αριθμός των αντικειμένων του συστήματος και n ο αριθμός των χρηστών. Ο πίνακας αυτός είναι η βάση για την λειτουργία όλου του συστήματος. Όταν τα συστήματα αυτού του είδους δεχτούν ένα αίτημα για παροχή συστάσεων από έναν χρήστη, εφαρμόζουν σε γενικές γραμμές τα ακόλουθα βήματα:

1ο Βήμα:

Στην περίπτωση που ένας χρήστης ζητήσει σύσταση, το σύστημα συγκρίνει τις αξιολογήσεις που έχει δώσει ο χρήστης αυτός, με αυτές όλων των άλλων χρηστών που βρίσκονται στον πίνακα αξιολογήσεων. Η αξιολόγηση από τον συγκεκριμένο χρήστη καταχωρείται εκείνη την στιγμή, εάν δεν βρίσκεται ήδη στον πίνακα αξιολογήσεων. Επιπλέον, το σύστημα μπορεί να χρησιμοποιεί και έμμεσες αξιολογήσεις. Μέσω αυτής της σύγκρισης, και με βάση κατάλληλου αλγόριθμου, το σύστημα βρίσκει έναν αριθμό χρηστών που μοιάζουν περισσότερο στον συγκεκριμένο χρήστη. Οι χρήστες αυτοί ονομάζονται «γείτονες» (neighbours) του τρέχοντος χρήστη και το σύνολό τους «γειτονιά» του (neighbourhood). Ένας τέτοιος αλγόριθμος, ο οποίος χρησιμοποιείται πολύ συχνά για την εύρεση των γειτόνων ενός χρήστη, είναι ο αλγόριθμος Pearson.

Ο αλγόριθμος Pearson μας δίνει μία «ποσοτικοποίηση» του βαθμού ομοιότητας ανάμεσα σε δύο χρήστες a , b με βάση τις αξιολογήσεις που έχουν δώσει αυτοί στο παρελθόν. Πιο συγκεκριμένα, αν r_{mn} είναι ο «βαθμός» που έχει αποδοθεί στο αντικείμενο n από τον χρήστη m , τότε ο βαθμός συσχέτισης των δύο χρηστών με βάση τον αλγόριθμο Pearson δίνεται από τον τύπο:

$$cor_{a,b} = \frac{\sum_{i=1}^n (r_{ai} - \bar{r}_a)(r_{bi} - \bar{r}_b)}{\sqrt{\left[\sum_{i=1}^n (r_{ai} - \bar{r}_a)^2 \right] \left[\sum_{i=1}^n (r_{bi} - \bar{r}_b)^2 \right]}}$$

Μία άλλη μέθοδος για την εύρεση των κοντινότερων γειτόνων ενός χρήστη είναι η μοντελοποίηση όλων των χρηστών σαν διανύσματα σε έναν χώρο με m διαστάσεις (όπου m ο αριθμός των αντικειμένων στην διάθεση του συστήματος). Οι συνιστώσες του διανύσματος του κάθε χρήστη σε κάθε διάσταση προκύπτουν από την αξιολόγηση που έχει δώσει ο χρήστης στο αντικείμενο που αντιστοιχεί στην συγκεκριμένη διάσταση. Ο βαθμός ομοιότητας μεταξύ δύο χρηστών ισούται με το συνημίτονο της γωνίας που σχηματίζουν τα αντίστοιχα διανύσματά τους.

2ο Βήμα:

Αφού καθοριστεί το ποιοι είναι οι κοντινότεροι «γείτονες» του χρήστη, το σύστημα ελέγχει αν υπάρχουν κάποια αντικείμενα που να μην έχουν αξιολογηθεί από τον συγκεκριμένο χρήστη ενώ ταυτόχρονα να έχουν λάβει υψηλές βαθμολογίες από τους «γείτονές» του. Το σύστημα στην συνέχεια προτείνει αυτά τα αντικείμενα στον χρήστη υπό μορφή συστάσεων. Υπάρχουν δύο βασικοί τρόποι για να γίνει αυτό:

1ος τρόπος: Η μέθοδος των πιο συχνά προτιμώμενων αντικειμένων

Με αυτή τη μέθοδο ελέγχονται όλα τα αντικείμενα που έχουν προτιμηθεί από τους χρήστες που ανήκουν στην γειτονιά του τρέχοντος χρήστη. Βρίσκει τα αντικείμενα που έχουν προτιμηθεί συχνότερα από αυτούς και τα επιστρέφει ως σύσταση στον χρήστη.

2ος τρόπος: Η μέθοδος των κανόνων συσχέτισης (Association rules)

Η μέθοδος των κανόνων συσχέτισης θεωρεί το σύνολο των αντικειμένων που έχουν προτιμηθεί από κάθε μέλος της γειτονιάς του τρέχοντος χρήστη ως ξεχωριστό σύνολο αντικειμένων. Δημιουργείται έτσι ένα σύνολο αντικειμένων για κάθε έναν από τους γείτονες του τρέχοντος χρήστη. Ύστερα αξιοποιείται ένας αλγόριθμος εύρεσης κανόνων συσχέτισης ώστε να βρεθούν όλοι οι κανόνες συσχέτισης ανάμεσα στα συγκεκριμένα σύνολα αντικειμένων. Οι κανόνες αυτοί αξιολογούνται με βάση τον υπολογισμό του βαθμού εμπιστοσύνης (confidence) και υποστήριξης (support) του καθενός. Οι κανόνες με τους υψηλότερους βαθμούς εμπιστοσύνης και υποστήριξης χρησιμοποιούνται για την εύρεση των αντικειμένων που πρέπει να προταθούν στον τρέχοντα χρήστη. Αν οι χρήστες που απαρτίζουν την «γειτονιά» του τρέχοντος χρήστη είναι λίγοι, υπάρχει η πιθανότητα το σύστημα να μην μπορέσει να βρει αρκετά ισχυρούς κανόνες συσχέτισης ώστε να μπορέσει να προτείνει αρκετά αντικείμενα. Στην περίπτωση αυτή η

μέθοδος μπορεί να «ενισχυθεί» με την μέθοδο των πιο συχνά προτιμώμενων αντικειμένων, ώστε να συμπληρωθεί η λίστα των αντικειμένων που θα συσταθούν.

2.2.1.2. Item-based collaborative filtering

Η αρχή λειτουργίας της μεθόδου collaborative filtering, όπως παρουσιάστηκε στην προηγούμενη παράγραφο, επικεντρώνεται στον συσχετισμό των χρηστών μεταξύ τους. Η μέθοδος αυτή ονομάζεται user-based collaborative filtering. Είναι φανερό από την περιγραφή της λειτουργίας αυτής της μεθόδου, ότι ο πίνακας αξιολογήσεων θα πρέπει να επαναπροσδιορίζεται σε πραγματικό χρόνο (real time) κάθε φορά που εισάγεται μία αξιολόγηση από έναν χρήστη, καθιστώντας έτσι τη σχέση μεταξύ των χρηστών δυναμική. Αυτό κάνει την εν λόγω μέθοδο να έχει αρκετά μεγάλο υπολογιστικό κόστος. Το πρόβλημα αυτό μπορεί να αντιμετωπιστεί με την αξιοποίηση μιας εναλλακτικής μεθόδου collaborative filtering που ονομάζεται item-based. Η μέθοδος αυτή, αντί να ερευνά τις ομοιότητες μεταξύ των χρηστών, ανακαλύπτει τις ομοιότητες που υπάρχουν ανάμεσα στα διάφορα αντικείμενα του συστήματος.

Η βάση του συστήματος είναι ένας πίνακας, ο οποίος περιλαμβάνει τα διάφορα αντικείμενα και τις ιδιότητές τους. Όταν ένας χρήστης ζητήσει σύσταση από το σύστημα, αυτό εξετάζει τα διάφορα αντικείμενα που έχει κατά καιρούς προτιμήσει ο χρήστης, όπως προκύπτει από στοιχεία αγοράς και άλλες πηγές. Στην συνέχεια το σύστημα, μέσω ενός αλγορίθμου συσχέτισης και αξιοποιώντας τα στοιχεία με τις ιδιότητες των αντικειμένων που υπάρχουν στον πίνακα, βρίσκει τα αντικείμενα που έχουν την μεγαλύτερη ομοιότητα με αυτά που έχει προτιμήσει ο χρήστης, και του τα προτείνει.

Σε αντίθεση με την μέθοδο user based, ο πίνακας συσχετισμών μεταξύ των αντικειμένων, που χρησιμοποιεί αυτή η μέθοδος, είναι σχετικά στατικός. Δεν υπάρχει επομένως η ανάγκη για online ενημέρωσή του κάθε φορά που κάποιος χρήστης αλληλεπιδρά με το σύστημα, και έτσι οι ανάγκες του σε υπολογιστική ισχύ είναι πολύ περιορισμένες. Η μέθοδος αυτή χρησιμοποιείται σήμερα σε πολλούς δικτυακούς τόπους.

Η μέθοδος όμως αυτή παρουσιάζει άλλου είδους προβλήματα. Το σημαντικότερο από αυτά είναι ότι αδυνατεί να αξιοποιήσει τα στοιχεία που δίνονται στο σύστημα την στιγμή που γίνεται το αίτημα για σύσταση. Δεν μπορεί, δηλαδή, να εκτιμήσει το ζήτημα που απασχολεί τον χρήστη την στιγμή που αυτός χρησιμοποιεί το σύστημα, παρά μόνο τα αντικείμενα που τον ενδιέφεραν στο παρελθόν. Οι επιστήμονες προτείνουν μια λύση σε αυτό το πρόβλημα. Σύμφωνα με την λύση αυτή το σύστημα διερευνά τις προθέσεις και τα ενδιαφέροντα του χρήστη την στιγμή που αυτός ζητεί σύσταση, πέρα από τα ενδιαφέροντα και τις προτιμήσεις του στο παρελθόν. Οι προθέσεις του εκείνη την στιγμή μπορούν να προκύψουν με διάφορους τρόπους: είτε άμεσα με την παροχή συγκεκριμένων στοιχείων από την μεριά του που προσδιορίζουν τις προθέσεις του, είτε έμμεσα με παρατήρηση της συμπεριφοράς του.

2.2.1.3. Αξιολόγηση της μεθόδου collaborative filtering

Στο σημείο αυτό είναι σκόπιμο να αναφερθούμε στα πλεονεκτήματα και τα μειονεκτήματα που παρουσιάζουν τα συστήματα συστάσεων που βασίζονται στην μέθοδο collaborative filtering και να τα αναλύσουμε εν συντομία.

A. Πλεονεκτήματα της μεθόδου:

- Δεν υπάρχει ανάγκη διατήρησης βάσης γνώσης για τα χαρακτηριστικά των αντικειμένων. Τα συστήματα collaborative filtering βασίζονται αποκλειστικά στην συσχέτιση χρηστών μεταξύ τους, μέσω των αξιολογήσεων που έχουν δώσει για τα διάφορα αντικείμενα. Έτσι, δεν έχουν ανάγκη υποστήριξης από μια μεγάλη, πολύπλοκη και δύσκολη στην ανάπτυξη βάση γνώσης. Στην πραγματικότητα, τα συστήματα αυτά δεν χρειάζεται να έχουν καμία γνώση για τα αντικείμενα που χειρίζονται. Αυτό είναι αρκετά χρήσιμο, αφού βοηθά στην γρήγορη ανάπτυξη ενός τέτοιου συστήματος.
- Η ποιότητα των συστάσεων της δεν είναι στατική αλλά βελτιώνεται με τον χρόνο. Ο πίνακας αξιολογήσεων των αντικειμένων από τους χρήστες, στον οποίο βασίζονται τα συστήματα collaborative filtering προκειμένου να βρουν τους «γείτονες» κάθε τρέχοντος χρήστη, εμπλουτίζεται διαρκώς με καινούριες αξιολογήσεις. Έτσι η ποιότητα των συστάσεων που μπορούν να παράσχουν αυτά τα συστήματα βελτιώνεται συνεχώς με την χρήση τους.

B. Μειονεκτήματα της μεθόδου:

- Η ποιότητα των συστάσεων βασίζεται σε ένα μεγάλο σύνολο δεδομένων παρουσιάζοντας το πρόβλημα «ramp-up». Πιο συγκεκριμένα, προκειμένου να μπορέσει το σύστημα να δώσει αξιόπιστες συστάσεις, πρέπει να υπάρχει ήδη ένας αρκετά μεγάλος αριθμός αξιολογήσεων από χρήστες. Είδαμε όμως, ότι η βάση αξιολογήσεων στα συστήματα collaborative filtering δημιουργείται και εμπλουτίζεται προοδευτικά. Στην περίπτωση λοιπόν που δεν υπάρχει ικανός αριθμός αξιολογήσεων από τους χρήστες, τα συστήματα αυτά αδυνατούν να δώσουν αξιόπιστες συστάσεις, ειδικά κατά τα πρώτα στάδια της λειτουργίας τους.
- Εμφανίζει το πρόβλημα “sparsity”, όταν υπάρχουν πάρα πολλά αντικείμενα και λίγοι σχετικά χρήστες. Ένας μεγάλος αριθμός συστημάτων collaborative filtering δραστηριοποιείται σε αγορές τέτοιες, ώστε χρειάζεται να περιλαμβάνει έναν μεγάλο αριθμό αντικειμένων. Όταν κάθε χρήστης του συστήματος δεν αξιολογεί μεγάλο αριθμό αντικειμένων, τότε το σύστημα θα αντιμετωπίσει το πρόβλημα «sparsity». Ο πίνακας αντικειμένων – χρηστών, ο οποίος περιέχει τις αξιολογήσεις θα παραμείνει «αραιός» καθιστώντας έτσι εξαιρετικά δύσκολη την εύρεση κατάλληλων «γειτόνων» των χρηστών. Χαρακτηριστικά παραδείγματα συστημάτων που μπορούν να αντιμετωπίσουν το πρόβλημα sparsity είναι το σύστημα Amazon και το CDnow.com.

Και στις δύο αυτές περιπτώσεις υπάρχει ένας πολύ μεγάλος αριθμός αντικειμένων (ηλεκτρονικές συσκευές, βιβλία, άλμπουμ μουσικής, ταινίες), ενώ ο κάθε χρήστης είναι απίθανο να μπορέσει να αξιολογήσει περισσότερο από το 1% των συνολικών διαθέσιμων αντικειμένων.

- Παρατηρείται το πρόβλημα «first rater». Τα συστήματα Collaborative Filtering συστήνουν αντικείμενα στους χρήστες ανάλογα με τις αξιολογήσεις που έχουν δεχθεί τα αντικείμενα αυτά από άλλους. Ένα νεοεισελθέν αντικείμενο δεν μπορεί να συσταθεί, τουλάχιστον μέχρι να δεχθεί ικανό αριθμό αξιολογήσεων από τους χρήστες του συστήματος. Αφού όμως ένα νεοεισελθέν αντικείμενο δεν συστήνεται στους χρήστες προκειμένου αυτοί να το χρησιμοποιήσουν και να το αξιολογήσουν, το πρόβλημα μοιάζει να ανατροφοδοτείται. Τα νεοεισελθέντα αντικείμενα χρειάζονται έτσι αρκετό χρόνο μέχρι να μπορέσουν να συσταθούν από το σύστημα.

Αυτό το πρόβλημα είναι σημαντικό, αν σκεφτεί κανείς ότι στις περισσότερες αγορές όπου δραστηριοποιούνται συστήματα συστάσεων (μουσικά CD, βιβλία, αυτοκίνητα, κ.α.) η είσοδος νέων προϊόντων είναι συνεχής και το ενδιαφέρον του αγοραστικού κοινού γι’ αυτά τα προϊόντα είναι μεγάλο.

- Δυσκολία στις συστάσεις προς κάποιον με ιδιαίτερες απαιτήσεις (popularity bias). Είναι φανερό από την περιγραφή της λειτουργίας του συστήματος, ότι όσο πιο συνηθισμένες προτιμήσεις έχει ένας χρήστης, τόσο πιο εύκολη θα είναι η εύρεση «γειτόνων». Επομένως, η παροχή συστάσεων προς έναν τέτοιο χρήστη θα είναι εύκολη. Για έναν χρήστη όμως, με ιδιαίτερες και ασυνήθιστες απαιτήσεις και ενδιαφέροντα, θα είναι πολύ δύσκολο να βρεθεί μία γειτονιά με αρκετά μέλη ώστε να λειτουργήσουν σωστά οι αλγόριθμοι που περιγράφονται στο 2ο βήμα της μεθόδου.
- Ευαισθησία σε στατιστικές ανωμαλίες. Τυχόν στατιστικές ανωμαλίες στο σύνολο των χρηστών μπορούν να έχουν σημαντικές επιπτώσεις στη λειτουργία του συστήματος.
- Ανάγκη μεγάλης υπολογιστικής ισχύος. Οι αλγόριθμοι που εκτελούν τα συστήματα collaborative filtering κάθε φορά που χρειάζεται να βρουν τους κοντινότερους «γείτονες» ενός χρήστη, έχουν ανάγκη σημαντικής υπολογιστικής ισχύος. Αυτή η ανάγκη μεγαλώνει όσο περισσότερους χρήστες και αντικείμενα καλείται να χειριστεί το σύστημα. Τα συστήματα που δραστηριοποιούνται στο διαδίκτυο συχνά καλούνται να χειριστούν εκατομμύρια αντικείμενα και χρήστες και έτσι έχουν ανάγκη πολύ μεγάλης υπολογιστικής ισχύος.
- Μικρή ευαισθησία σε αλλαγές προτιμήσεων των χρηστών. Ο πίνακας αντικειμένων – χρηστών περιέχει αξιολογήσεις των αντικειμένων από τους χρήστες. Οι αξιολογήσεις αυτές μπορεί να είναι, όπως θα δούμε αργότερα, άμεσες ή έμμεσες. Οι αξιολογήσεις αποτελούν ουσιαστικά ένα «ιστορικό» των προτιμήσεων του χρήστη. Το ιστορικό αυτό μπορεί να είναι μακροχρόνιο και έτσι ακόμα και μία ριζική και ξαφνική αλλαγή προτιμήσεων του χρήστη θα χρειαστεί κάποιο χρόνο ώστε να επηρεάσει σημαντικά το «ιστορικό» του ώστε το σύστημα να αλλάξει τις συστάσεις του προς αυτόν. Επομένως, η ευαισθησία του συστήματος στις αλλαγές προτιμήσεων των χρηστών του είναι αρκετά μικρή, και μάλιστα

τόσο μικρότερη, όσο πιο εκτεταμένο χρονικά είναι το αρχείο των αξιολογήσεων που διατηρείται.

- Το πρόβλημα της σύστασης συχνά προτιμώμενων αντικειμένων (πρόβλημα της «μπανάνας»). Ένα πρόβλημα που συνδέεται στενά με τον τρόπο λειτουργίας των συστημάτων collaborative filtering είναι το ότι τα συστήματα αυτού του είδους έχουν την τάση να συστήνουν κατά κόρον κάποια αντικείμενα που προτιμώνται συχνά από πολλούς χρήστες. Το πρόβλημα αναφέρεται από τον Burke με τον όρο «πρόβλημα της μπανάνας» (banana problem). Ας φανταστούμε ένα αντικείμενο που προτιμάται συχνά από τους χρήστες ενός συστήματος (όπως οι μπανάνες από τους πελάτες των οπωροπωλείων). Ένα τέτοιο αντικείμενο θα προκύπτει σχεδόν πάντα κατά την εκτέλεση των αλγορίθμων του 2ου βήματος της διαδικασίας και, επομένως, θα προτείνεται σχεδόν πάντα στους χρήστες.

2.2.1.4. Άμεση – έμμεση αξιολόγηση

Όπως έχει ήδη γίνει σαφές, τα συστήματα συστάσεων βασίζονται σε μεγάλο βαθμό στην αξιολόγηση από τους χρήστες των διαφόρων αντικειμένων που μπορούν να προτείνουν. Αυτή η εξάρτηση είναι φανερό σε πολύ μεγαλύτερο βαθμό στα συστήματα συστάσεων που βασίζονται στην τεχνική του Collaborative Filtering. Χωρίς την ύπαρξη ενός πολύ μεγάλου όγκου δεδομένων αξιολόγησης, είναι το σύστημα αδύνατο να εφαρμόσει τον αλγόριθμο και να υπολογίσει τους βαθμούς συσχέτισης μεταξύ των χρηστών, ώστε να παρέχει συστάσεις. Τα δεδομένα αυτά πρέπει να είναι σε μορφή τέτοια ώστε να μπορούν να τα επεξεργαστούν οι διαδικασίες του συστήματος. Προκύπτει έτσι το ζήτημα του πώς θα αντληθούν αυτές οι αξιολογήσεις από τους χρήστες. Υπάρχουν βασικά δύο τρόποι για να γίνει αυτό, άμεσα ή έμμεσα.

Ο άμεσος τρόπος αξιολόγησης των αντικειμένων από τους χρήστες είναι αυτός με την πιο ευρεία εφαρμογή. Χρησιμοποιείται όχι μόνο στα συστήματα συστάσεων, αλλά και σε πολλές άλλες περιπτώσεις όπως π.χ. σε βαθμολογία γραπτών από καθηγητές, σε βαθμολογία ταινιών από κριτικούς, κλπ. Η άμεση αξιολόγηση μπορεί να παρέχεται σε ορισμένες περιπτώσεις με μορφή ελεύθερου κειμένου – σχολίων κ.τ.λ., συνήθως όμως βρίσκεται στην μορφή μίας διακριτής κλίμακας βαθμολογίας (π.χ. από το 1 μέχρι το 10). Αυτή η μορφή είναι η επιβαλλόμενη στα πληροφορικά συστήματα, προκειμένου να μπορούν μετέπειτα οι αξιολογήσεις αυτές να υποστούν στατιστική επεξεργασία, και να αξιοποιηθούν με κατάλληλο τρόπο.

Όπως είναι φανερό, η άμεση αξιολόγηση παρέχεται από τους χρήστες, αφού τους ζητηθεί από το σύστημα συστάσεων. Ο χρήστης πρέπει να έχει εξετάσει το προς αξιολόγηση αντικείμενο πριν το βαθμολογήσει με βάση την κλίμακα που του προτείνεται από το σύστημα. Προκύπτει λοιπόν ότι η άμεση μορφή αξιολόγησης απαιτεί προσπάθεια και συμμετοχή εκ μέρους του χρήστη. Επομένως, είναι αναμενόμενη η ύπαρξη μιας τάσης για αποφυγή της αξιολόγησης από την πλευρά του χρήστη, όπου αυτό είναι δυνατόν. Αυτό είναι συχνά πηγή προβλημάτων για πολλά συστήματα συστάσεων που χρησιμοποιούν την μέθοδο Collaborative Filtering, όταν δεν γίνεται κατανοητό στον χρήστη το ότι «αμείβεται» με κάποια ανταλλάγματα για την προσπάθεια που καταβάλλει αξιολογώντας αντικείμενα.

Ο έμμεσος τρόπος απόκτησης στοιχείων αξιολόγησης λειτουργεί εντελώς διαφορετικά. Αντί το σύστημα να ζητά από τους χρήστες να βαθμολογήσουν αντικείμενα, αντλεί χρήσιμα στοιχεία αξιολόγησης από την συμπεριφορά των χρηστών. Έτσι αποφεύγονται πολλά από τα προβλήματα που προκαλούνται όταν ένα σύστημα βασίζεται αποκλειστικά στη μέθοδο άμεσης αξιολόγησης. Η μέθοδος αυτή χρησιμοποιεί σαν πηγή δεδομένων κάποια αρχεία καταγραφής ενεργειών, στα οποία αποθηκεύονται στοιχεία που περιγράφουν την συμπεριφορά των χρηστών. Τέτοια στοιχεία συμπεριφοράς μπορεί να είναι τα εξής:

- Ποιους δικτυακούς τόπους επισκέπτεται συχνότερα ο χρήστης και για πόσο χρόνο παραμένει σ' αυτούς.
- Ποια προϊόντα αγοράζει μέσω διαδικτύου.
- Ποια αντικείμενα εξετάζει συχνότερα και για πόσο χρόνο.

Τα στοιχεία αυτά μπορούν να συλλεχθούν κατά την διάρκεια της χρήσης του συστήματος χωρίς την εμπλοκή του χρήστη (όπως αυτή που απαιτεί ένα αίτημα για αξιολόγηση), ειδικά στην

περίπτωση που τα αντικείμενα που τον ενδιαφέρουν μπορεί να τα αποκτήσει – αγοράσει κανείς μέσω του Διαδικτύου. Τα στοιχεία που αναφέραμε μπορούν να συλλεχθούν σε πολύ μεγάλες ποσότητες στο διαδίκτυο, ιδιαίτερα λόγω της ολοένα και μεγαλύτερης εξάπλωσης του ηλεκτρονικού εμπορίου. Και από τεχνολογικής πλευράς η διαδικασία δεν είναι δύσκολη. Τα στοιχεία που συλλέγονται όμως, βρίσκονται σε μορφή ακατάλληλη για να αξιοποιηθούν άμεσα από ένα σύστημα συστάσεων. Πρέπει, επομένως, να υποστούν την κατάλληλη επεξεργασία ώστε να ποσοτικοποιηθούν. Αυτή η διαδικασία, καθώς και η διαδικασία συλλογής και αποθήκευσης των στοιχείων, έχει υπολογιστικό κόστος, που όμως δεν επιβαρύνει τον χρήστη του συστήματος. Η όλη διαδικασία μπορεί μάλιστα να εκτελείται εν αγνοία του χρήστη, εκτός αν άλλοι λόγοι επιβάλλουν την ενημέρωσή του. Αν και η μέθοδος αυτή λύνει πολλά από τα προβλήματα που προκύπτουν από την μέθοδο άμεσης αξιολόγησης, προκύπτουν άλλα προβλήματα που οφείλονται κυρίως στην φύση των στοιχείων που χρησιμοποιούνται. Τα στοιχεία αυτά παράγονται σε τεράστιες ποσότητες, κάθε φορά που κάποιος χρησιμοποιεί το διαδίκτυο, σε βαθμό που μπορεί να δημιουργηθεί ζήτημα ύπαρξης πολύ μεγάλου όγκου δεδομένων, με όλα τα συνεπακόλουθα σε κόστος συλλογής, αποθήκευσης και επεξεργασίας. Η σχετική χρησιμότητα των στοιχείων αυτών είναι πολύ μικρή, αν συγκριθεί με αυτή των στοιχείων άμεσης αξιολόγησης.

Ένα πρόβλημα που παρατηρείται συχνά σε συστήματα συστάσεων που χρησιμοποιούν έμμεσες αξιολογήσεις είναι η λανθασμένη ποσοτικοποίηση ή εξαγωγή λανθασμένων συμπερασμάτων από τα δεδομένα που υπάρχουν για τους χρήστες. Αυτό μπορεί να είναι ιδιαίτερα σημαντικό στην περίπτωση που χρησιμοποιείται σαν στοιχείο ο χρόνος παραμονής ενός χρήστη σε μία ηλεκτρονική σελίδα. Δεν υπάρχει κανένας τρόπος να καταλάβουμε αν ο χρήστης πραγματικά εξέταζε την σελίδα ή είχε απομακρυνθεί από τον υπολογιστή του, αφήνοντάς τον με την σελίδα ανοιχτή. Σημαντική δυσχέρεια επίσης παρουσιάζει το γεγονός ότι οι πίνακες αντικειμένων – χρηστών που περιλαμβάνουν τις αξιολογήσεις είναι πολύ πιο πιθανό να αντιμετωπίζουν το πρόβλημα sparsity, όταν αυτές οι αξιολογήσεις έχουν προέλθει από έμμεσα στοιχεία.

Ένα άλλο πολύ σημαντικό πρόβλημα που προκύπτει είναι αυτό της εμπιστευτικότητας. Τα στοιχεία που αφορούν στην συμπεριφορά των χρηστών στο διαδίκτυο μπορούν να θεωρηθούν σε ορισμένες περιπτώσεις ευαίσθητα στον βαθμό που μπορούν να χρησιμοποιηθούν για να δημιουργηθεί ένα προφίλ της γενικότερης συμπεριφοράς ή και της προσωπικότητας του χρήστη. Είναι, επομένως, αναμενόμενο οι χρήστες να είναι ιδιαίτερα επιφυλακτικοί και να μην δίνουν εύκολα την συγκατάθεσή τους στην συλλογή και χρήση τέτοιων δεδομένων. Είναι λοιπόν απαραίτητο να παρέχονται στους χρήστες οι απαραίτητες εγγυήσεις για την χρήση των δεδομένων αυτών. Στο άμεσο μέλλον, με την συνεχώς ευρύτερη χρήση των πιστωτικών καρτών καθώς και του Ηλεκτρονικού Εμπορίου, θα γίνει ίσως δυνατή η ανακατασκευή ενός πλήρους καταναλωτικού προφίλ για όλους τους πολίτες. Οι δυνατότητες, αλλά και τα ηθικά ζητήματα, που προκύπτουν από μια τέτοια εξέλιξη είναι προφανή.

2.2.2. Συστήματα βασισμένα στην γνώση

Η δεύτερη κατηγορία συστημάτων που θα εξεταστεί είναι η κατηγορία των συστημάτων που βασίζονται στην Γνώση. Τα συστήματα που ανήκουν σε αυτή την κατηγορία αξιοποιούν την γνώση σχετικά με τα χαρακτηριστικά των διαθέσιμων προϊόντων ώστε να καταλήξουν σε μια λογική υπόθεση – σύσταση για το ποια προϊόντα καλύπτουν τις ανάγκες ενός συγκεκριμένου χρήστη. Το σύνολο της γνώσης που χρησιμοποιεί ένα τέτοιο σύστημα μπορεί να χωριστεί σε τρεις κατηγορίες. Όπου κάθε μία από αυτές απαιτεί ιδιαίτερο χειρισμό για την απόκτησή της:

- Γνώση για τα αντικείμενα: Το σύστημα πρέπει να έχει γνώση σχετικά με τα αντικείμενα που μπορεί να συστήσει, για τα χαρακτηριστικά αυτών, καθώς και για το πώς αυτά σχετίζονται μεταξύ τους.
- Γνώση για τους χρήστες: Το σύστημα πρέπει, επίσης, να αποκτήσει γνώση σχετικά με τους χρήστες του. Πρέπει να γνωρίζει τις ανάγκες, τα ενδιαφέροντα και, ενδεχομένως τις συνήθειες του χρήστη στον οποίο πρόκειται να κάνουν σύσταση. Από το είδος αυτό της γνώσης συντίθεται ουσιαστικά ένα «προφίλ» για τον κάθε χρήστη.
- Γνώση για τον τρόπο κάλυψης των αναγκών: Το σύστημα πρέπει να γνωρίζει το πώς συγκεκριμένες ανάγκες των χρηστών μπορούν να καλυφθούν από συγκεκριμένα

χαρακτηριστικά των αντικειμένων. Πρέπει, δηλαδή, να γνωρίζει ποια αντικείμενα ταιριάζουν στο «προφίλ» ενός χρήστη.

Ο τρόπος απόκτησης αυτών των ειδών γνώσης καθώς και ο τρόπος χρήσης τους, αποτελούν χαρακτηριστικά των διαφόρων ειδών αυτής της κατηγορίας συστημάτων. Το πρώτο είδος γνώσης, δηλαδή η γνώση σχετικά με τα αντικείμενα - προϊόντα που βρίσκονται στην διάθεση του, μπορεί να αντλείται από το σύστημα μέσω τεχνικών εξόρυξης γνώσης (data mining) που εφαρμόζονται σε μεγάλους όγκους δεδομένων. Το σύστημα αποφασίζει ποια αντικείμενα θα συμπεριλάβει στην βάση δεδομένων και με ποια χαρακτηριστικά. Αυτός ο τρόπος δημιουργίας γνώσης θα μπορούσε να χαρακτηριστεί ως έμμεσος. Άλλος τρόπος δημιουργίας αυτού του είδους γνώσης είναι ο άμεσος, όπου τα αντικείμενα και τα χαρακτηριστικά του καθενός μπορούν να εισάγονται στο σύστημα απευθείας από τους διαχειριστές του. Και στις δύο αυτές περιπτώσεις, το αποτέλεσμα είναι μια βάση δεδομένων «γνώσης». Η βάση αυτή περιλαμβάνει στοιχεία για τα αντικείμενα και τα χαρακτηριστικά τους, δομημένα με τρόπο κατάλληλο ώστε να χρησιμοποιηθούν στην διεργασία της παραγωγής της σύστασης.



Εικόνα 2-2. Λειτουργία συστήματος γνώσης

2.2.2.1. Critique-based recommendation

Το επόμενο στάδιο στην διαδικασία, είναι η κατασκευή ενός προφίλ για τον χρήστη, η ενσωμάτωση δηλαδή του δευτέρου είδους γνώσης. Ο τρόπος κατασκευής του προφίλ καθώς και το πώς αυτό χρησιμοποιείται, αποτελεί την βάση για τον περαιτέρω διαχωρισμό των συστημάτων που βασίζονται στην γνώση, σε υποκατηγορίες. Η πρώτη υποκατηγορία είναι αυτή της case – based reasoning (ή αλλιώς critique - based recommendation). Τα συστήματα που ανήκουν σε αυτήν την υποκατηγορία προσπαθούν να εντοπίσουν αντικείμενα με χαρακτηριστικά τέτοια που να ικανοποιούν τις απαιτήσεις του χρήστη, όπως αυτές καθορίζονται από τον ίδιο. Η διαδικασία αυτή περιλαμβάνει τα ακόλουθα βήματα:

- Το πρώτο βήμα αφορά στην καταγραφή των προτιμήσεων του χρήστη. Αυτό μπορεί να γίνει ερωτώντας τον χρήστη κατά πόσο ενδιαφέρεται για ορισμένα χαρακτηριστικά των προϊόντων. Το ενδιαφέρον αυτό ποσοτικοποιείται εύκολα, αφού συνήθως είναι σε μορφή σειράς διακριτών επιπέδων, όπως: πολύ, αρκετά, λίγο, καθόλου κ.τ.λ. Οι προτιμήσεις – απαιτήσεις του χρήστη, όπως καθορίζονται στο πρώτο βήμα, αποτελούν το προφίλ του χρήστη.
- Στο δεύτερο βήμα της διαδικασίας, το σύστημα εξετάζει τα χαρακτηριστικά των αντικειμένων που έχει στην διάθεσή του, και γνωρίζοντας το πώς τα χαρακτηριστικά αυτά καλύπτουν τις απαιτήσεις των χρηστών, εντοπίζει τα προϊόντα που ταιριάζουν περισσότερο στον συγκεκριμένο χρήστη και τα παρουσιάζει σε αυτόν ως συστάσεις. Η διαδικασία τελειώνει εδώ, αν ο χρήστης δηλώσει ικανοποιημένος με κάποιες από τις συστάσεις που του παρουσιάστηκαν, ή αν δεν υπάρχει πια ενδιαφέρον από την πλευρά του.
- Στο τρίτο βήμα της διαδικασίας ο χρήστης, έχοντας ήδη εξετάσει τα αντικείμενα που του συστάθηκαν στο δεύτερο βήμα, μεταβάλλει σε κάποιο βαθμό τις προτιμήσεις του. Το σύστημα δέχεται αυτές τις καινούριες προτιμήσεις και επαναλαμβάνει το δεύτερο βήμα.

2.2.2.2. Goal-based recommendation

Η δεύτερη υποκατηγορία των συστημάτων που βασίζονται στην γνώση είναι αυτή που χρησιμοποιεί τα μέτρα ομοιότητας σαν μέσο εύρεσης αντικειμένων που εξυπηρετούν τον ίδιο σκοπό με αυτά που έχουν επιλεγεί από τον χρήστη στο παρελθόν. Η υποκατηγορία αυτή ονομάζεται «Συστήματα βασιζόμενα στον σκοπό» ή «Goal-based systems». Αναφέρεται, επίσης, ο όρος «Συστήματα εύρεσης» για αυτήν την κατηγορία συστημάτων, ειδικότερα όταν αυτά χρησιμοποιούν παραδείγματα για να υποβοηθήσουν την επικοινωνία τους με τον χρήστη. Η λειτουργία των συστημάτων αυτών περιλαμβάνει τρία στάδια ολοκλήρωσης:

- Το πρώτο στάδιο της διαδικασίας είναι ο καθορισμός των προτιμήσεων του χρήστη, μέσω των αντικειμένων που αυτός έχει επιλέξει στο παρελθόν. Τα αντικείμενα αυτά έχουν ορισμένα χαρακτηριστικά, τα οποία υποδηλώνουν τους λόγους για τους οποίους ο χρήστης τα προτίμησε και επομένως και τον σκοπό του.
- Το δεύτερο βήμα είναι η εύρεση αντικειμένων που είναι παρόμοια με αυτά που είχε προτιμήσει ο χρήστης. Λέγοντας παρόμοια, εννοούμε ότι εξυπηρετούν παρόμοιους σκοπούς. Χαρακτηριστικό παράδειγμα που αναφέρει ο Burke είναι το εξής: μία ομπρέλα είναι παρόμοια με ένα αδιάβροχο όταν ο σκοπός είναι η προστασία από την βροχή. Είναι όμως παρόμοια με ένα μπαστούνι, όταν ο σκοπός είναι η προστασία από έναν σκύλο. Στο βήμα αυτό χρησιμοποιείται εκτεταμένα η μέθοδος των κριτηρίων ομοιότητας.
- Το τρίτο βήμα της διαδικασίας είναι και εδώ, όπως και στην προηγούμενη υποκατηγορία συστημάτων, ανατροφοδότηση από τον χρήστη, ο οποίος μπορεί να μην μείνει ικανοποιημένος από την πρώτη σύσταση που θα λάβει.

Η ανατροφοδότηση αυτή μπορεί να είναι μέσω της μεθόδου “tweak” που στρέφει το σύστημα σε μία διαφορετική κατεύθυνση από την αρχική. Αυτό μπορεί να γίνει με άμεση βαθμολόγηση των προτεινόμενων αντικειμένων από τον χρήστη. Το σύστημα μπορεί, επίσης, να παρακολουθήσει την συμπεριφορά του χρήστη μετά από την σύσταση και να εξάγει έμμεσα συμπεράσματα. Έτσι, την πρώτη ομάδα συστάσεων μπορεί να ακολουθήσει και δεύτερη ομάδα που να εξυπηρετεί καλύτερα τις ανάγκες του.

Το πρώτο σύστημα Συστάσεων που χαρακτηρίστηκε με τον όρο “Findme” ήταν το σύστημα “Car Navigator”. Το σύστημα αυτό βοήθησε τους χρήστες του να βρουν πληροφορίες για καινούρια μοντέλα αυτοκινήτων των οποίων είχε αποθηκεύσει τα χαρακτηριστικά. Τα χαρακτηριστικά αυτά αποτελούν και τα κριτήρια μέσω των οποίων το σύστημα παρουσιάζει μία πρώτη ομάδα μοντέλων σαν αρχική σύσταση προς τον χρήστη. Τα κριτήρια μπορούν να διαφοροποιηθούν απευθείας από τον χρήστη για να προκύψει μια δεύτερη σύσταση, πλησιέστερη στις ανάγκες του. Παρόλα αυτά, το σύστημα εμπεριέχει την δυνατότητα αλλαγής πολλών κριτηρίων ταυτόχρονα με το πάτημα ενός πλήκτρου (τεχνική tweak). Αν για παράδειγμα, ο χρήστης ήθελε ένα αυτοκίνητο πιο οικονομικό από αυτά που του είχαν ήδη προταθεί, μέσω του αντίστοιχου κουμπιού μπορούσε να αλλάξει τα κριτήρια εύρεσης προς αυτήν την κατεύθυνση. (π.χ. μικρότερος κινητήρας, χαμηλότερη κατανάλωση, μικρότερη τιμή αγοράς, κ.τ.λ.).

Χαρακτηριστικό παράδειγμα εφαρμογής της τεχνικής Συστάσεων βασισμένης στην γνώση είναι το σύστημα “Entree”, που παρέχει στους χρήστες του συστάσεις για εστιατόρια. Το βασικό του πλεονέκτημα είναι η εκτεταμένη βάση δεδομένων του, η οποία είναι δομημένη με αποτελεσματικό τρόπο. Το σύστημα Entree μαζί με το σύστημα personal logic, ανήκουν στην υποκατηγορία των συστημάτων εκείνων που βασίζονται στον σκοπό.

2.2.2.3. Αξιολόγηση των συστημάτων γνώσης

Στο σημείο αυτό θα αναφερθούμε στα πλεονεκτήματα και τα μειονεκτήματα που παρουσιάζουν τα συστήματα που βασίζονται στην γνώση και θα αναλύσουμε εν συντομία το κάθε ένα από αυτά.

Πλεονεκτήματα των συστημάτων γνώσης:

- Δεν αντιμετωπίζουν το πρόβλημα Ramp-up ή cold start, ένα πρόβλημα πολύ κοινό στη μέθοδο Collaborative filtering. Λόγω του ότι το σύστημα δεν βασίζεται σε μία σειρά

αξιολογήσεων των αντικειμένων από την μεριά των χρηστών, η οποία χρειάζεται αρκετό χρόνο, δεν αντιμετωπίζει το πρόβλημα Rump-up, που είναι ένα από τα πιο σημαντικά προβλήματα που παρουσιάζουν τα συστήματα Collaborative filtering. Ένα σύστημα βασισμένο στην γνώση μπορεί να δώσει αξιόπιστες συστάσεις από την αρχή της λειτουργίας του και αρκετά μεγάλη ευαισθησία στις αλλαγές των προτιμήσεων των χρηστών.

- Τα συστήματα που βασίζονται στην γνώση έχουν την δυνατότητα να δέχονται από τους χρήστες πληροφορίες για τις προτιμήσεις τους με άμεσο τρόπο. Χάρη σ' αυτή την δυνατότητα τα μπορούν να αναγνωρίζουν άμεσα, ακόμη και τις πιο ξαφνικές και απότομες αλλαγές στις προτιμήσεις ενός χρήστη. Οι συστάσεις τους μπορούν, λοιπόν, να προσαρμοστούν αμέσως στα νέα ενδιαφέροντα των χρηστών.
- Δεν χρειάζεται μεγάλη υπολογιστική ισχύ για να υπολογίζει σε πραγματικό χρόνο τις επιδράσεις των αξιολογήσεων, σε αντίθεση με τα συστήματα collaborative filtering, τα οποία απαιτούν σημαντική υπολογιστική ισχύ για να εντοπίσουν τους κοντινότερους «γείτονες» ενός χρήστη, κάθε φορά που οποιοσδήποτε χρήστης υποβάλλει μία αξιολόγηση.
- Δεν επηρεάζεται από στατιστικές ανωμαλίες στο δείγμα χρηστών.
- Μπορεί να κάνει προτάσεις σε χρήστες με ιδιαίτερες απαιτήσεις. Ένας χρήστης που έχει προτιμήσεις ή απαιτήσεις που διαφέρουν πολύ από αυτές των περισσότερων άλλων χρηστών θα αντιμετωπίσει πρόβλημα αν χρησιμοποιήσει ένα σύστημα Collaborative filtering. Ένα σύστημα που βασίζεται στην γνώση, όμως θα μπορέσει να του παράσχει αξιόπιστες συστάσεις. Δεν έχει καμία σημασία πόσο ιδιαίτερες είναι οι ανάγκες του, αρκεί αυτές να μπορούν να εισαχθούν στο σύστημα και βάση γνώσης του συστήματος να «γνωρίζει» το πώς μπορούν να καλυφθούν.
- Μπορεί να προτείνει καινούρια και μη αξιολογημένα αντικείμενα. Τα συστήματα αυτού του είδους δεν βασίζονται σε αξιολογήσεις των αντικειμένων. Ένα καινούριο αντικείμενο μπορεί να ενσωματωθεί αμέσως μαζί με τα χαρακτηριστικά του στην βάση γνώσης και να προτείνεται κανονικά, όταν χρειάζεται, χωρίς η έλλειψη αξιολογήσεων γι' αυτό να αποτελεί πρόβλημα, όπως θα αποτελούσε για ένα σύστημα collaborative filtering. Αυτό το πλεονέκτημα είναι ιδιαίτερα σημαντικό σε αγορές όπου ο ρυθμός ανανέωσης των προϊόντων είναι μεγάλος.
- Μπορεί να παράσχει αιτιολόγηση των συστάσεών του μέσω των χαρακτηριστικών των αντικειμένων που προτάθηκαν. Εφόσον απαιτείται από την φύση των συστάσεων ή ζητείται από τον χρήστη, το σύστημα μπορεί να παρουσιάσει συνοπτικά και με κατάλληλη μορφή και διατύπωση την σειρά των διαδικασιών που οδήγησαν στις συστάσεις που έγιναν προς τον συγκεκριμένο χρήστη, παρέχοντάς του έτσι μια «αιτιολόγηση» των συστάσεων αυτών. Το γεγονός αυτό δίνει την δυνατότητα στο σύστημα να είναι πιο «διαφανές» ως προς τις διαδικασίες που ακολουθεί αυξάνοντας έτσι σημαντικά τη «φιλικότητα» και λειτουργικότητα του συστήματος.

Μειονεκτήματα των συστημάτων γνώσης:

- Η κατασκευή μιας κατάλληλης βάσης γνώσης για το σύστημα είναι μια δύσκολη και απαιτητική διαδικασία. Η βάση γνώσης που χρησιμοποιούν τα συστήματα αυτού του είδους είναι αρκετά δύσκολο να κατασκευαστεί και απαιτείται σημαντική εργασία. Η βάση αυτή πρέπει να περιλαμβάνει γνώσεις σχετικά με το ποια χαρακτηριστικά των αντικειμένων είναι σημαντικά, ποιες ανάγκες καλύπτουν και με ποιο τρόπο. Πρέπει επίσης να υπάρχει πρόσβαση σε μία βάση δεδομένων που να περιέχει τα χαρακτηριστικά αυτά των αντικειμένων. Βεβαίως, σε ορισμένες περιπτώσεις, το μειονέκτημα αυτό είναι λιγότερο έντονο. Τα συστήματα αυτά δεν απαιτούν ιδιαίτερα μεγάλες προσπάθειες για την κατασκευή βάσης γνώσης. Χαρακτηριστικό παράδειγμα αποτελεί το σύστημα «Vintage Exchange», ένα σύστημα συστάσεων για κρασιά, όπου για την κατασκευή της βάσης γνώσης του συστήματος απαιτήθηκε μόλις ένας μήνας εργασίας.
- Η δυνατότητα σύστασης του συστήματος είναι στατική. Η βάση γνώσης του συστήματος έχει κατασκευαστεί πριν ακόμη το σύστημα τεθεί σε λειτουργία. Η δυνατότητα σύστασης του συστήματος εξαρτάται αποκλειστικά από την βάση αυτή και για να βελτιωθεί πρέπει να ανανεωθεί ή να επανωξηθεί. Σε αντίθετη περίπτωση, η ποιότητα των συστάσεων που παρέχονται από το σύστημα παραμένει η στατική.

- Χρειάζεται ποσοτικοποίηση των χαρακτηριστικών των αντικειμένων. Προκειμένου να λειτουργήσουν αποτελεσματικά τα συστήματα του τύπου αυτού, πρέπει να έχουν πολύ καλή γνώση των χαρακτηριστικών των αντικειμένων που συστήνουν. Τα χαρακτηριστικά αυτά πρέπει να ποσοτικοποιηθούν ώστε να μπορεί η βάση γνώσης του συστήματος να τα χειριστεί. Συχνά είναι απαραίτητη η επεξεργασία κειμένου γραμμένου σε φυσική γλώσσα με σκοπό την εξαγωγή συμπερασμάτων που να μπορούν να αξιοποιηθούν. Χαρακτηριστικό παράδειγμα είναι το σύστημα Entrée που αξιοποιεί κείμενα που περιλαμβάνουν κριτικές διαφόρων εστιατορίων για να υποβοηθήσει την διαδικασία της εύρεσης των χαρακτηριστικών των αντικειμένων.

2.2.3. Υβριδικά Συστήματα Συστάσεων

Στις προηγούμενες παραγράφους αναφερθήκαμε σε κάποιες από τις κύριες τεχνικές συστάσεων. Προκύπτει λοιπόν το συμπέρασμα ότι αν ένα σύστημα συστάσεων συνδυάσει τις τεχνικές αυτές, υπάρχει η δυνατότητα αναιρέσης πολλών μειονεκτημάτων τους. Επίσης, ένα σύστημα που χρησιμοποιεί μόνο μία από τις προαναφερθείσες μεθόδους είναι πολύ πιθανό να μην ανταποκρίνεται πλήρως στις απαιτήσεις του περιβάλλοντος στο οποίο θα χρησιμοποιηθεί.

Για τους λόγους αυτούς, υπάρχει ισχυρός προσανατολισμός στην ανάπτυξη και χρήση συστημάτων συστάσεων που να ενσωματώνουν στην λειτουργία τους περισσότερες μεθόδους, έτσι ώστε να συνδυάζονται οι δυνατότητές τους και να αλληλοεξουδετερώνονται οι περισσότερες από τις αδυναμίες τους. Προκύπτουν έτσι τα λεγόμενα υβριδικά συστήματα συστάσεων, τα οποία έχουν πολλές δυνατότητες καθώς υπάρχουν περιθώρια βελτίωσης, μέσω της αξιοποίησης τεχνικών με τις οποίες μπορεί να επιτευχθεί ο συνδυασμός – υβριδισμός των δύο βασικών τεχνικών.

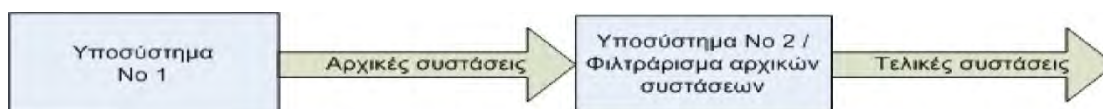
2.2.3.1 Υβριδικά συστήματα τύπου «καταρράκτη»

Ένας τρόπος συνδυασμού των δύο βασικών τεχνικών συστάσεων αποτελεί τα συστήματα τύπου «καταρράκτη». Τα συστήματα αυτά χρησιμοποιούν την μία τεχνική για να ταξινομήσουν καλύτερα τις συστάσεις στις οποίες έχει καταλήξει η άλλη. Σε αντίθεση με τις τεχνικές που είδαμε προηγουμένως, στην τεχνική καταρράκτη τα υποσυστήματα συστάσεων δεν λειτουργούν ανεξάρτητα, αλλά επεμβαίνουν το ένα στην λειτουργία του άλλου. Τα συστήματα τύπου καταρράκτη αποτελούνται, όπως και τα προηγούμενα συστήματα που εξετάσαμε, από δύο υποσυστήματα – συνιστώσες: ένα υποσύστημα γνώσης και ένα υποσύστημα collaborative filtering. Η λειτουργία τους εκτελείται σε δύο στάδια (Εικόνα 2-3):

A) Το ένα από τα δύο υποσυστήματα, χρησιμοποιώντας τα διαθέσιμα στοιχεία, καταλήγει σε μία αρχική ομάδα συστάσεων. Η ομάδα αυτή μπορεί και πρέπει να περιλαμβάνει μεγάλο αριθμό συστάσεων. Συνήθως το υποσύστημα που αναλαμβάνει τον σχηματισμό της πρώτης ομάδας συστάσεων είναι το υποσύστημα γνώσης.

B) Η ομάδα των συστάσεων που αποτελεί την έξοδο του πρώτου σταδίου, γίνεται είσοδος για το δεύτερο υποσύστημα. Οι συστάσεις της ομάδας αυτής κατατάσσονται με βάση τα κριτήρια του δεύτερου υποσυστήματος και διακριτοποιούνται περισσότερο. Γίνεται έτσι δυνατή η περαιτέρω βελτίωση των συστάσεων που παρέχει το πρώτο βήμα.

Σημαντικό πλεονέκτημα της μεθόδου αυτής είναι η υψηλή της αποδοτικότητα. Η μέθοδος αποδίδει καλύτερα διότι η διαδικασία του δεύτερου βήματος εφαρμόζεται μόνο σε αντικείμενα που έχουν ήδη αξιολογηθεί θετικά από το πρώτο βήμα και όχι σε όλα τα διαθέσιμα αντικείμενα.



Εικόνα 2-3. Λειτουργία συστήματος τύπου «Καταρράκτη»

Βασικό σημείο στην κατανόηση της λειτουργίας τέτοιων συστημάτων είναι το ότι το δεύτερο βήμα μπορεί μόνο να βελτιώσει τις συστάσεις που παρέχει το πρώτο βήμα, δεν μπορεί

όμως να τις ανατρέψει. Τα δύο υποσυστήματα λειτουργούν σαν «φίλτρα» τοποθετημένα εν σειρά. Αυτό το στοιχείο μπορεί να αποβεί θετικό ή αρνητικό ανάλογα με τις συνθήκες λειτουργίας του συστήματος.

2.2.4 Παράμετροι Επιτυχημένων τεχνικών συστάσεων

Είναι δεδομένο ότι η επιθυμία των σχεδιαστών ενός συστήματος συστάσεων είναι η παροχή αξιόπιστων συστάσεων στους χρήστες. Σε ορισμένες περιπτώσεις όμως, η ανάγκη αξιοπιστίας από την αρχή κίχλας της λειτουργίας του συστήματος είναι επιτακτική. Τέτοιες περιπτώσεις είναι συνηθισμένες σε ένα ανταγωνιστικό περιβάλλον, όταν λειτουργούν ήδη παρόμοια συστήματα με αυτό που σχεδιάζεται. Στις περιπτώσεις αυτές η τεχνική collaborative filtering υστερεί, λόγω της ύπαρξης του προβλήματος ramp-up. Η τεχνική αυτή χρειάζεται έναν σχετικά μεγάλο αριθμό αξιολογήσεων για να λειτουργήσει σωστά και γι' αυτό είναι συνηθισμένο να δίνει ανακριβείς συστάσεις κατά τα πρώτα στάδια της λειτουργίας της. Αντίθετα, η τεχνική γνώσης βασίζεται σε μεγάλο βαθμό στην βάση γνώσης που διαθέτει από την αρχή της λειτουργίας της, για αυτό και μπορεί να δίνει ακριβείς συστάσεις από τα πρώτα στάδια λειτουργίας.

2.2.4.1 Συνθήκες έλλειψης επαρκούς αριθμού αξιολογήσεων

Η επιτυχής λειτουργία ενός συστήματος collaborative filtering βασίζεται κατά κύριο λόγο σε έναν πίνακα αξιολογήσεων. Ο πίνακας αυτός περιέχει τις αξιολογήσεις που έχει κάνει ο κάθε χρήστης για κάθε αντικείμενο. Με βάση αυτόν τον πίνακα συσχετίζονται μεταξύ τους οι χρήστες. Στην περίπτωση που ο πίνακας αυτός είναι «αραιός» (sparse), περιέχει δηλαδή μικρό αριθμό αξιολογήσεων, οι συστάσεις που θα παράγει το σύστημα συστάσεων θα είναι αναξιόπιστες. Οι περιπτώσεις έλλειψης επαρκούς αριθμού αξιολογήσεων είναι αρκετά συχνές. Οι χρήστες μπορεί να δείχνουν απροθυμία να υποβάλλουν αξιολογήσεις στο σύστημα για διάφορους λόγους. Είναι επίσης πιθανό, το σύστημα να σχεδιάζεται με σκοπό να λειτουργήσει για να εξυπηρετεί έναν σχετικά μικρό αριθμό χρηστών, οι οποίοι όμως να μπορούν να επιλέξουν ανάμεσα σε μεγάλο αριθμό αντικειμένων. Σε αυτήν την περίπτωση μπορεί να παρουσιαστεί αυτό το πρόβλημα. Στις περιπτώσεις αυτές ένα σύστημα συστάσεων βασισμένο στην τεχνική γνώσης θα έχει σαφές πλεονέκτημα.

Ένα ακόμη περιβάλλον λειτουργίας στο οποίο παρουσιάζεται έλλειψη αξιολογήσεων είναι το περιβάλλον στο οποίο εμφανίζονται καινούρια αντικείμενα με αρκετά μεγάλη συχνότητα. Τέτοιες συνθήκες λειτουργίας επικρατούν στην μουσική βιομηχανία ή στην αγορά βιβλίων, όπου εμφανίζονται συνεχώς καινούρια προϊόντα. Τα καινούρια αντικείμενα δεν θα προτείνονται στους χρήστες, έως ότου δεχθούν ικανό αριθμό αξιολογήσεων, καθιστώντας έτσι εγκυρότερες τις συστάσεις (πρόβλημα του νεοεισελθέντος αντικειμένου). Είναι όμως φανερό, και αυτό ισχύει εμφανέστατα στην περίπτωση της μουσικής βιομηχανίας, ότι τα καινούρια προϊόντα πρέπει να προτείνονται συχνά, ώστε το αγοραστικό κοινό να αντιληφθεί την άφιξή τους. Αυτή η ανάγκη δεν μπορεί να καλυφθεί από ένα σύστημα συστάσεων βασισμένο αποκλειστικά στην τεχνική collaborative filtering.

2.2.4.2 Αιτιολόγηση συστάσεων

Η επιτυχία ενός συστήματος συστάσεων δεν κρίνεται αποκλειστικά από την ικανότητά του να δίνει καλές συστάσεις. Κρίνεται και από την ικανότητά του να πείθει τους χρήστες του να ακολουθήσουν τις συστάσεις που τους παρέχει. Με άλλα λόγια το σύστημα συστάσεων πρέπει να εμπνέει εμπιστοσύνη στους χρήστες του. Προφανώς, ο βασικός τρόπος για να γίνει αυτό, είναι οι συστάσεις να είναι ακριβείς. Ένας άλλος τρόπος, με τον οποίο μπορεί ένα σύστημα συστάσεων να εμπνεύσει εμπιστοσύνη είναι να αιτιολογήσει την σύστασή του, να εξηγήσει δηλαδή στον χρήστη τους λόγους για τους οποίους θα πρέπει να προτιμήσει το αντικείμενο που του προτείνεται. Αυτή η δυνατότητα αυξάνει την αποτελεσματικότητα του συστήματος σχεδόν σε κάθε περίπτωση, και υπάρχει μόνο στα συστήματα που βασίζονται σε κάποιο βαθμό στην τεχνική γνώσης. Δεδομένης της εμπορικής – οικονομικής διάστασης της λειτουργίας των συστημάτων συστάσεων, το πλεονέκτημα αυτό της τεχνικής γνώσης είναι σημαντικό.

2.2.4.3 Χρήστες με ασυνήθιστες προτιμήσεις

Σε αρκετές συνθήκες λειτουργίας αναμένεται να προκύπτουν περιπτώσεις χρηστών, οι οποίες να βρίσκονται σε μεγάλη απόσταση από τους συνηθισμένους χρήστες όσον αφορά στις προτιμήσεις τους. Οι χρήστες αυτοί δεν μπορούν να λάβουν ικανοποιητικές συστάσεις από ένα σύστημα collaborative filtering. Αντίθετά, ένα σύστημα γνώσης θα τους εξυπηρετήσει καλύτερα, αφού μπορεί να εκτιμήσει τον καλύτερο τρόπο κάλυψης οποιασδήποτε ανάγκης, αρκεί βέβαια να περιλαμβάνεται στην βάση γνώσης του συστήματος η ανάγκη προσαρμογής σε μεταβαλλόμενες προτιμήσεις. Μία κατάσταση στην οποία θα κληθεί να ανταποκριθεί το οποιοδήποτε σύστημα συστάσεων είναι αυτή όπου οι χρήστες του μεταβάλλουν τις προτιμήσεις τους, είτε γρήγορα, είτε σε βάθος χρόνου.

Κάτι τέτοιο αναμένεται να συμβαίνει συχνότερα σε συστήματα που αντιμετωπίζουν έναν μεγάλο όγκο αρκετά διαφοροποιημένων μεταξύ τους αντικειμένων. Σε τέτοιες περιπτώσεις, η τεχνική collaborative filtering παρουσιάζει σημαντική υστέρηση στην καταγραφή και ανταπόκριση της μεταβολής των προτιμήσεων ή των ενδιαφερόντων ενός χρήστη. Αντίθετα, τα συστήματα γνώσης ανταποκρίνονται σε αυτές τις αλλαγές σχεδόν αμέσως. Αυτό ισχύει ιδιαίτερα για τα «critique – based» συστήματα.

2.2.4.4. Έλλειψη επαρκών στοιχείων γνώσης

Η βάση γνώσης είναι το σημείο – κλειδί για να μπορέσει η τεχνική γνώσης να δώσει αξιόπιστες συστάσεις. Είναι δυνατόν να υπάρχουν περιπτώσεις, κατά τις οποίες οι σχεδιαστές ενός συστήματος συστάσεων να μην γνωρίζουν τα χαρακτηριστικά των αντικειμένων που αυτό θα χειρίζεται. Είναι πιθανόν τα εν λόγω αντικείμενα να μην είναι ακόμη έτοιμα. Μία άλλη κατάσταση που συναντάται συχνά είναι να μην υπάρχουν διαθέσιμα στοιχεία σχετικά με τον τρόπο ικανοποίησης των αναγκών των χρηστών από τα αντικείμενα (το τρίτο είδος γνώσης) και να είναι έτσι δύσκολη η κατασκευή ενός μοντέλου γνώσης. Σε αυτές τις περιπτώσεις η χρησιμοποίηση ενός συστήματος βασισμένου στην τεχνική collaborative filtering είναι προτιμότερη με την μεγαλύτερη αξιοπιστία δεδομένων.

2.2.4.5. Ανάγκη γρήγορης ανάπτυξης

Ο κόσμος του ηλεκτρονικού εμπορίου είναι πολύ ανταγωνιστικός και οι εξελίξεις είναι ταχύτατες. Η ταχύτητα είναι συνήθως κρίσιμος παράγοντας για την επιτυχία ενός εγχειρήματος στον τομέα των συστημάτων συστάσεων. Η έλλειψη ταχύτητας ανάπτυξης σαφώς μειώνει την ωφέλιμη διάρκεια «ζωής» ενός συστήματος. Στον τομέα αυτό η τεχνική collaborative filtering έχει σαφές πλεονέκτημα. Ο χρόνος που χρειάζεται για την ανάπτυξη ενός συστήματος βασισμένου σε αυτήν την τεχνική είναι πραγματικά πολύ μικρός, μια και το βασικό σημείο που χρειάζεται προσοχή είναι ο αλγόριθμος που θα χρησιμοποιηθεί για την συσχέτιση των χρηστών μεταξύ τους και την δημιουργία «γειτονιών» χρηστών. Σε αντίθεση με αυτά, η ανάπτυξη συστήματος βασισμένου στην τεχνική γνώσης είναι χρονοβόρα διαδικασία. Ιδιαίτερα η βάση γνώσης είναι ένα στοιχείο του συστήματος πολύ απαιτητικό σε χρόνο εξειδικευμένης εργασίας.

Πρέπει να σημειώσουμε ότι τα συστήματα γνώσης, αντιμετωπίζουν αυτό το πρόβλημα σε μικρότερο βαθμό και επομένως μπορούν να θεωρηθούν ικανοποιητικός συμβιβασμός σε περιπτώσεις όπου η ταχύτητα ανάπτυξης είναι σημαντική.

3. Εξόρυξη γνώσης από δεδομένα (Knowledge Mining)

3.1 Εισαγωγή

Η εξόρυξη γνώσης από βάσεις δεδομένων (Knowledge Mining in Databases - KMD) αναφέρεται στη διεργασία ανακάλυψης εμφωλευμένης γνώσης σε μεγάλες βάσεις δεδομένων. Ο όρος εξόρυξη δεδομένων (Data Mining) χρησιμοποιείται ως συνώνυμο της ανακάλυψης γνώσης, καθώς επίσης και για αναφορά στις τεχνικές που χρησιμοποιούνται για την ανάλυση και την εξαγωγή γνώσης από διάφορα σύνολα δεδομένων. Κατά γενική παραδοχή, ο όρος KMD αναφέρεται στο σύνολο των διαδικασιών από την επεξεργασία των δεδομένων έως τα τελικά

στάδια αξιοποίησης της γνώσης, ενώ ο όρος Data Mining αποτελεί μια από τις διαδικασίες του KMD και αναφέρεται στην εξόρυξη γνώσης από τα δεδομένα.

Εν ολίγοις, η εξόρυξη δεδομένων (Data Mining) αναφέρεται στην βιβλιογραφία ως η σύνθετη διαδικασία εξαγωγής συγκεκριμένης, προηγούμενης άγνωστης και δυνητικά ωφέλιμης, γνώσης από δεδομένα (Frawley, 1992). Εναλλακτικά, συναντάται και ως η επιστήμη της εξόρυξης χρήσιμης πληροφορίας από σύνολα ή βάσεις δεδομένων μεγάλου μεγέθους (Handetal, 2001). Αν και η εξόρυξη δεδομένων αποτελεί σχετικά νέο όρο, η τεχνολογία δεν είναι καινούργια. Εδώ και αρκετό καιρό, οι εταιρείες χρησιμοποιούν ακριβούς και ισχυρούς υπολογιστές που αναλύουν όγκους δεδομένων και συντάσσουν εκθέσεις έρευνας αγοράς. Οι τεχνικές της εξόρυξης δεδομένων είναι αποτέλεσμα μιας μακράς διαδικασίας έρευνας και ανάπτυξης προϊόντων. Την πρόοδο της τεχνολογίας των υπολογιστών, που είχε ως αποτέλεσμα το χαμηλό κόστος υπολογιστών με πολύ-επεξεργαστές και τεράστιες δυνατότητες αποθήκευσης, ακολούθησαν βελτιώσεις στις τεχνολογίες που προέκυψαν από την πρόσβαση σε δεδομένα. Οι τεχνολογίες αυτές επιτρέπουν στους χρήστες να πλοηγηθούν στα δεδομένα τους σε πραγματικό χρόνο. Η εξόρυξη δεδομένων χρησιμοποιεί κάθε διαθέσιμη τεχνολογία ώστε η εξελικτική αυτή διαδικασία να ξεπεράσει την πρόσβαση και πλοήγηση στα δεδομένα μέσω της παράδοσης προοπτικών και προ-δραστικών πληροφοριών.

Για να διαφοροποιηθούμε μεταξύ της διαδικασίας και των εργαλείων, θα χρησιμοποιήσουμε τον πρώτο όρο, δηλαδή τον όρο KMD, για να περιγράψουμε ολόκληρη τη διαδικασία ανάλυσης ενός συνόλου δεδομένων, και το δεύτερο όρο, την εξόρυξη δεδομένων, για να αναφερθούμε κυρίως στις μεθόδους και τις τεχνικές που χρησιμοποιούνται στη διαδικασία ανάλυσης. Πολλοί ερευνητές, θεωρούν τον όρο εξόρυξη δεδομένων μη αντιπροσωπευτικό της διαδικασίας που περιγράφει, υποστηρίζοντας ότι ο όρος εξόρυξη γνώσης θα ήταν μια πιο κατάλληλη περιγραφή. Εντούτοις, ένας τέτοιος όρος μπορεί να μη δίνει έμφαση στην ανάλυση και την εξαγωγή των προτύπων από μεγάλα σύνολα δεδομένων.

Ο όρος εξόρυξη δεδομένων είναι αυτός που έχει επικρατήσει και χαρακτηρίζει τη διαδικασία της εύρεσης δομών γνώσης, οι οποίες περιγράφουν με ακρίβεια μεγάλα σύνολα πρωτογενών δεδομένων. Οι δομές αυτές αναδεικνύουν γνώση (συσχετίσεις ή κανόνες) που είναι κρυμμένη μέσα στα δεδομένα και δεν μπορεί να εξαχθεί από τον άνθρωπο-χρήστη με «γυμνό» μάτι. Οι προκύπτουσες δομές είναι πλούσιες σε σημασιολογία και εκμεταλλεύονται πιθανές κοινές ιδιότητες των πρωτογενών δεδομένων. Ο Fayyad, επικεντρώθηκε στην ανακάλυψη γνώσης από βάσεις δεδομένων και θεώρησε την προσπάθεια αυτή ως: «τη διεργασία της χρήσης των βάσεων δεδομένων μαζί με οποιαδήποτε απαίτηση επίλογής, προεπεξεργασίας, υποδειγματοποίησης και μετασχηματισμού, προκειμένου να εφαρμοστούν οι διάφοροι μέθοδοι εξόρυξης δεδομένων για να απარიθμηθούν τα μοτίβα που ενυπάρχουν στη βάση, και να αξιολογηθούν τα προϊόντα της εξόρυξης γνώσης για να αναγνωριστούν τα υποσύνολα των απარიθμηθέντων μοτίβων που θεωρούνται γνώση». Έτσι, παρατηρούμε ότι παρ' όλη την διεργασία και τους υπολογισμούς για την εξαγωγή συμπερασμάτων με ντετερμινιστικό τρόπο, τελικά ο χρήστης – ειδικός, είναι αυτός που θα κρίνει την χρησιμότητα ή την καταλληλότητα της εξαγόμενης γνώσης.

Η ανακάλυψη γνώσης από μία βάση δεδομένων αναφέρεται σε ολόκληρη τη διαδικασία ανακάλυψης χρήσιμης πληροφορίας από μεγάλα σύνολα δεδομένων. Ένας γενικός ορισμός, που παρουσιάζει με περισσότερη σαφήνεια την έννοια του όρου Knowledge Discovery in Databases (KDD) είναι: «Η ανακάλυψη γνώσης στις βάσεις δεδομένων είναι η ντετερμινιστική και σημαντική διαδικασία αναγνώρισης έγκυρων, καινοτόμων, ενδεχομένως χρήσιμων και εν τέλει κατανοητών μοτίβων στα δεδομένα». Ο όρος μοτίβο ή πρότυπο (pattern) ενσωματώνει έννοιες όπως: συσχετίσεις (correlations), συνάψεις ή σχέσεις (relationships), τάσεις (trends), περιγραφές σπάνιων γεγονότων, κλπ. Βέβαια υπάρχουν και άλλοι τρόποι περιγραφής του στόχου της ανακάλυψης γνώσης και της εξόρυξης δεδομένων, ανάλογα πάντα με τη γενικότητα του προβλήματος και τις προσδοκίες των επιστημόνων.

Η διαδικασία KDD είναι μια διαλογική και επαναληπτική διαδικασία που αποτελείται από τα ακόλουθα βήματα

1. Την ανάπτυξη και κατανόηση της περιοχής της εφαρμογής, της σχετικά προγενέστερης γνώσης του προς εξέταση τομέα, και τους στόχους του τελικού χρήστη.
2. Την ολοκλήρωση των δεδομένων. Υπάρχουν διαφορετικά είδη αποθηκών πληροφοριών που μπορούν να χρησιμοποιηθούν στη διαδικασία εξόρυξης γνώσης. Κατά συνέπεια, οι πολλαπλές πηγές δεδομένων μπορούν να συνδυαστούν καθορίζοντας το σύνολο στο οποίο

πρόκειται τελικά να εφαρμοστεί η διαδικασία εξόρυξης.

3. Τη δημιουργία του στόχου-συνόλου δεδομένων. Αφορά στην επιλογή του συνόλου δεδομένων (δηλαδή μεταβλητές, δείγματα δεδομένων) στο οποίο πρόκειται να εκτελεστεί η διαδικασία εξόρυξης.
4. Τον καθαρισμό και την προεπεξεργασία δεδομένων. Αυτό το βήμα περιλαμβάνει βασικές διαδικασίες όπως η αφαίρεση του θορύβου, η συλλογή των απαραίτητων πληροφοριών για τη διαμόρφωση ή τη μέτρηση του θορύβου, και η απόφαση σχετικά με τις στρατηγικές διαχείρισης των ελλειπόντων πεδίων δεδομένων.
5. Τον μετασχηματισμό των δεδομένων. Τα δεδομένα μετασχηματίζονται ή παγιώνονται σε μορφές κατάλληλες για εξόρυξη. Γίνεται χρήση μεθόδων μείωσης διαστάσεων ή μετασχηματισμού για τη μείωση του αριθμού των υπό εξέταση μεταβλητών ή την εύρεση κατάλληλης αντιπροσώπευσης των δεδομένων.
6. Την επιλογή των στόχων και των αλγορίθμων εξόρυξης δεδομένων. Σε αυτό το βήμα αποφασίζεται ο στόχος της διαδικασίας KMD, επιλέγοντας τους στόχους εξόρυξης δεδομένων που θέλουμε να επιτύχουμε. Επίσης, επιλέγονται οι μέθοδοι που θα χρησιμοποιηθούν. Αυτό περιλαμβάνει την επιλογή του κατάλληλου μοντέλου και παραμέτρων. Επίσης η μέθοδος εξόρυξης δεδομένων πρέπει να αντιστοιχηθεί με τις απαιτήσεις και τα γενικά κριτήρια της διαδικασίας KMD.
7. Την εξόρυξη δεδομένων. Εφαρμόζοντας ευφείς μεθόδους, αναζητούνται ενδιαφέροντα μοτίβα γνώσης. Τα μοτίβα αυτά θα μπορούσαν να είναι μιας συγκεκριμένης αντιπροσωπευτικής μορφής ή ενός συνόλου αντιπροσωπεύσεων όπως κανόνες κατηγοριοποίησης (classification rules), δέντρα, παλινδρόμηση, ομαδοποίηση (clustering) κ.λπ. Η απόδοση και τα αποτελέσματα της μεθόδου εξόρυξης δεδομένων εξαρτώνται από τα προηγούμενα βήματα της συνολικής διαδικασίας ανακάλυψης γνώσης.
8. Την αξιολόγηση των μοτίβων. Τα εξαγόμενα πρότυπα ή μοτίβα αξιολογούνται με κάποια μέτρα, προκειμένου να προσδιοριστούν εκείνα τα οποία αντιπροσωπεύουν καλύτερα τη γνώση, δηλαδή τα μοτίβα για τα οποία ενδιαφερόμαστε περισσότερο.
9. Την σταθεροποίηση και παρουσίαση της γνώσης. Σε αυτό το βήμα, η εξορυγμένη γνώση ενσωματώνεται στο σύστημα και χρησιμοποιούνται κάποιες τεχνικές αντιπροσώπευσης αυτής προκειμένου να παρουσιαστεί ευκρινώς στο χρήστη.

Όπως προαναφέραμε, η εξόρυξη δεδομένων ως βήμα της διαδικασίας KMD ενδιαφέρεται κυρίως για τις μεθοδολογίες και τις τεχνικές εξαγωγής προτύπων δεδομένων ή τις περιγραφές δεδομένων από μεγάλες αποθήκες δεδομένων. Αφ' ετέρου, η διαδικασία KMD περιλαμβάνει την αξιολόγηση και την ερμηνεία των προτύπων, την επιλογή της κωδικοποίησης των προτύπων, της προ-επεξεργασίας, της δειγματοληψίας και του μετασχηματισμού των δεδομένων πριν από το βήμα της εξόρυξης των δεδομένων. Η διαδικασία KMD είναι επαναληπτική και θα μπορούσε να περιέχει βρόχους μεταξύ οποιωνδήποτε από τα παραπάνω βήματα.

3.2. Τεχνικές εξόρυξης δεδομένων

Η εξόρυξη δεδομένων διεξάγεται επί δεδομένων που παρουσιάζονται σε μορφή ποσοτική, κειμενική ή μορφή πολυμέσων. Οι εφαρμογές της περιλαμβάνουν τη διασύνδεση (Association: σχήματα σύμφωνα με τα οποία ένα συμβάν διασυνδέεται με άλλο συμβάν, όπως οι επικοινωνία χρηστών), η αλληλουχία ή ανάλυση οδού (sequence or path analysis: σχήματα όπου το ένα συμβάν οδηγεί σε άλλο συμβάν, όπως η μεγάλη χρήση φωνής και η υψηλότερη κερδοφορία του πελάτη), η ταξινόμηση (classification: αναγνώριση νέων σχημάτων), η συσταδοποίηση (clustering: η ανεύρεση και καταγραφή ομάδων προηγούμενων γνωστών στοιχείων, όπως η προτιμήσεις) και η πρόβλεψη (forecasting: η ανακάλυψη σχημάτων από τα οποία είναι δυνατόν να γίνουν λογικές προβλέψεις σχετικά με μελλοντικές δραστηριότητες.

Όταν τα εργαλεία εξόρυξης δεδομένων εφαρμόζονται σε συστήματα παράλληλης επεξεργασίας υψηλής απόδοσης, είναι σε θέση να αναλύσουν τεράστιες βάσεις δεδομένων μέσα σε λίγα λεπτά. Η ταχύτερη επεξεργασία δίνει στους χρήστες τη δυνατότητα να πειραματιστούν αυτομάτως με περισσότερα πρότυπα, ώστε να κατανοήσουν σύνθετα δεδομένα. Η υψηλή ταχύτητα καθιστά πρακτική την ανάλυση τεράστιων όγκων δεδομένων των χρηστών και, με τη σειρά τους, οι μεγαλύτερες βάσεις δεδομένων οδηγούν σε βελτιωμένες προβλέψεις.

Η πράξη της οικοδόμησης προτύπων είναι κάτι που οι άνθρωποι κάνουν εδώ και πολύ καιρό, σαφώς πριν την εμφάνιση των υπολογιστών ή της τεχνολογίας εξόρυξης δεδομένων. Αυτό, όμως, που συμβαίνει με τους υπολογιστές δεν διαφέρει και πολύ από τον τρόπο που οικοδομούν πρότυπα οι άνθρωποι. Στους υπολογιστές παρέχονται πολλές πληροφορίες σχετικά με διάφορες καταστάσεις, όπου είναι γνωστή μια απάντηση και στη συνέχεια το λογισμικό της εξόρυξης δεδομένων του υπολογιστή πρέπει να διέλθει από τα δεδομένα και να «διυλίσσει» τα χαρακτηριστικά των δεδομένων που πρέπει να περιληφθούν στο πρότυπο. Αφού οικοδομηθεί το πρότυπο, στη συνέχεια μπορεί να χρησιμοποιηθεί σε παρόμοιες καταστάσεις, όπου η απάντηση δεν είναι γνωστή.

Ενώ η εξόρυξη δεδομένων αποτελεί σημαντική πρόοδο στον τύπο των εργαλείων ανάλυσης που έχουμε σήμερα στη διάθεσή μας, οι δυνατότητές της έχουν τα όριά τους. Ένας περιορισμός έγκειται στο γεγονός ότι, παρ' όλο που η εξόρυξη δεδομένων μπορεί να αποκαλύψει σχήματα και σχέσεις, δεν λέει στον χρήστη την αξία ή τη σημασία αυτών των δεδομένων. Ένας δεύτερος περιορισμός είναι ότι, ενώ η εξόρυξη δεδομένων μπορεί να αναγνωρίσει διασυνδέσεις μεταξύ συμπεριφορών και/ή μεταβλητών, δεν αναγνωρίζει οπωσδήποτε την αιτιολογική σχέση. Για να είναι επιτυχής, η εξόρυξη δεδομένων εξακολουθεί να απαιτεί ειδικούς στην τεχνολογία και την ανάλυση οι οποίοι να μπορούν να συνθέσουν την ανάλυση και να ερμηνεύσουν τα εξαγόμενα αποτελέσματα.

3.3. Παράγοντες επιτυχίας εξόρυξης δεδομένων

Υπάρχουν δύο παράγοντες επιτυχίας στην εξόρυξη δεδομένων. Ο πρώτος είναι η ακριβής διατύπωση του προβλήματος. Μια επικεντρωμένη πρόταση έχει συνήθως τα καλύτερα αποτελέσματα. Ο δεύτερος παράγοντας - κλειδί είναι τα ορθά δεδομένα. Η εξόρυξη δεδομένων δεν παρέχει αυτομάτως λύσεις χωρίς καθοδήγηση. Εκτός αυτού, αν και ένα καλό εργαλείο εξόρυξης δεδομένων προστατεύει από περίπλοκες στατιστικές τεχνικές, είναι απαραίτητη η κατανόηση του τρόπου λειτουργίας των εργαλείων που επιλέγονται και των αλγορίθμων επί των οποίων βασίζονται. Όπως και με όλες τις τεχνικές διαχείρισης γνώσης, η χρήση τόσο των ορθών δεδομένων (ρητή γνώση) όσο και η καλή τεχνογνωσία που αφορά την επιχειρηματική λειτουργία (άρρητη γνώση) έχουν μεγάλη σημασία.

Μεγάλος αριθμός εταιρειών έχουν αναπτύξει επιτυχείς εφαρμογές εξόρυξης δεδομένων. Ενώ οι πρώτοι που υιοθέτησαν την τεχνολογία αυτή ανήκαν κυρίως στον τομέα έντασης πληροφοριών, όπως οι χρηματο-οικονομικές υπηρεσίες και το μάρκετινγκ άμεσης ταχυδρόμησης, η τεχνολογία είναι εφαρμόσιμη σε οποιαδήποτε εταιρεία αναζητά να χρησιμοποιήσει αποδοτικά μια μεγάλη αποθήκη δεδομένων ώστε να διαχειριστεί με καλύτερο τρόπο τις πελατειακές της σχέσεις.

Το μέγιστο επιχειρηματικό όφελος επιτυγχάνεται από την διάθεση των μοντέλων που προκύπτουν μέσω της διαδικασίας της εξόρυξης των δεδομένων (Data Mining) στα σημεία επαφής της επιχείρησης με τον πελάτη (διαδίκτυο, καταστήματα πώλησης, τηλεφωνικό κέντρο, γραπτή επικοινωνία κ.λπ.), οπότε μπορούμε να απαντάμε ερωτήματα της μορφής "τι θα μπορούσαμε να προσφέρουμε στον συγκεκριμένο πελάτη σήμερα για να τον διατηρήσουμε ενεργό στην επιχείρηση".

Αρκετό ενδιαφέρον παρουσιάζουν επίσης και τα εργαλεία εξόρυξης δεδομένων τα οποία εξελίσσονται συνεχώς βασισμένα σε ιδέες από τις τελευταίες επιστημονικές έρευνες. Πολλά από τα εργαλεία αυτά ενσωματώνουν τους πιο πρόσφατους αλγορίθμους από την τεχνητή νοημοσύνη, τη στατιστική και τη βελτιστοποίηση. Προς το παρόν η γρήγορη επεξεργασία επιτυγχάνεται με χρήση σύγχρονων τεχνικών βάσεων δεδομένων, όπως κατακεντρωμένη επεξεργασία με αρχιτεκτονικές client/server με παράλληλες βάσεις δεδομένων και με αποθήκες δεδομένων. Η μελλοντική τάση είναι προς την ανάπτυξη πιο πλήρων διαδικασιών διαδικτύου. Η επεξεργασία πρέπει να εκτελείται με χρήση όλων των διαθέσιμων πηγών. Η ύπαρξη κατακεντρωμένων περιβαλλόντων, παρέχοντας τη δυνατότητα κατανομής των πόρων όλων των συστημάτων θα ωφελήσει τη διαδικασία της επεξεργασίας των δεδομένων ως προς το χρόνο και τη μνήμη.

Τομείς που χρησιμοποιούν την εξόρυξη δεδομένων είναι γνωστό πως ένα μεγάλο μέρος της ερευνητικής βάσης της εξόρυξης δεδομένων βασίζεται στη στατιστική. Αυτό είναι λογικό μιας και η στατιστική έχει ανάλογους σκοπούς με την εξόρυξη δεδομένων αφού αποσκοπούν στην αναγνώριση χρήσιμων πληροφοριών και προτύπων στα δεδομένα. Μέρος των διαδικασιών σε ένα μοντέλο εξόρυξης δεδομένων μπορεί να αποτελεί η αναζήτηση των δεδομένων και η εξαγωγή συμπερασμάτων από τα αποτελέσματα μιας αναζήτησης. Μια συχνά

χρησιμοποιούμενη τεχνική στην εξόρυξη δεδομένων είναι αυτή της δειγματοληψίας. Αυτός ο τρόπος στη στατιστική λέγεται «στατιστική εξαγωγή συμπεράσματος». Ακόμα και σήμερα, ένα σημαντικό τμήμα των νέων υλοποιημένων αλγόριθμων εξόρυξης δεδομένων αποτελούν στην ουσία στατιστικές τεχνικές που έχουν προσαρμοστεί στις απαιτήσεις των αλγορίθμων και των υπολογισμών. Όπως και με τις κλασικές τεχνικές στατιστικής στην εξόρυξη δεδομένων ακολουθούμε τεχνικές όπως η ανάλυση παλινδρόμησης (regression analysis), ανάλυση συστάδων (cluster analysis) κ.α.. Ακόμα και όταν οι αλγόριθμοι εξόρυξης δεδομένων δεν χρησιμοποιούν άμεσα τεχνικές στατιστικής, πολλές φορές οι βασικές τους ιδέες έχουν ως αρχική επιρροή την στατιστική.

Δύο άλλοι τομείς που σχετίζονται με αυτόν της εξόρυξης δεδομένων είναι η τεχνητή νοημοσύνη και η μηχανική μάθηση. Σκοπός της τεχνητής νοημοσύνης είναι να βγάξει λογικά συμπεράσματα από ανεπεξέργαστα δεδομένα, κάτι που κάνει και ο τομέας της εξόρυξης δεδομένων. Επίσης ο τομέας της εξόρυξης δεδομένων κάνει εκτεταμένη χρήση εργαλείων τεχνητής νοημοσύνης και μηχανικής μάθησης. Μερικά παραδείγματα είναι τα νευρωνικά δίκτυα, τα δέντρα απόφασης και

Οι μηχανές διανυσμάτων (vector machines). Γενικά ο τομέας της τεχνητής νοημοσύνης είναι πιο γενικός και εμπεριέχει περιοχές εκτός των κλασικών μεθόδων εξόρυξης δεδομένων. Επίσης εφαρμογές τεχνητής νοημοσύνης μπορεί να μην σχετίζονται με κλιμάκωση δεδομένων μιας και ο όγκος τους μπορεί να είναι αρκετά μικρός.

Η μηχανική μάθηση είναι μια περιοχή της τεχνητής νοημοσύνης η οποία εξετάζει πώς μπορούμε να δημιουργούμε προγράμματα τα οποία μπορούν να μαθαίνουν. Στην εξόρυξη δεδομένων, η μηχανική μάθηση χρησιμοποιείται για τεχνικές πρόβλεψης ή κατηγοριοποίησης.

Με τη μηχανική μάθηση, ο υπολογιστής κάνει κάποιες προβλέψεις και μετά, βασιζόμενος στην ανατροφοδότηση (feedback), όποτε αυτό είναι ορθό, μαθαίνει από αυτό. Μαθαίνει από τα παραδείγματα, την αποθηκευμένη γνώση, και την ανατροφοδότηση. Όταν συμβεί μελλοντικά ανάλογη περίπτωση, η ανατροφοδότηση χρησιμοποιείται για να κάνει την ίδια πρόβλεψη ή για να κάνει μια εντελώς διαφορετική πρόβλεψη. Η στατιστική είναι πολύ σημαντική σε προγράμματα μηχανικής μάθησης γιατί τα αποτελέσματα των προβλέψεων πρέπει να είναι στατιστικά σημαντικά.

Τέλος μια βάση δεδομένων είναι μια συλλογή από δεδομένα. Αντίθετα με ένα απλό σύνολο, τα δεδομένα σε μια βάση έχουν μια ορισμένη δομή ή σχήμα με το οποίο είναι σχετιζόμενα. Έτσι τα δεδομένα σε μια βάση αναπαρίστανται με ένα πιο θεωρητικό τρόπο ή με ένα μοντέλο δεδομένων. Αυτό το μοντέλο χρησιμοποιείται για να περιγράψει τα δεδομένα, τα χαρακτηριστικά τους, και τις σχέσεις μεταξύ τους.

Ένα μεγάλο μέρος των σημερινών ερευνητών στην εξόρυξη δεδομένων είναι άτομα προερχόμενα από τον τομέα των βάσεων δεδομένων. Η σχέση των δύο αυτών τομέων είναι εμφανής μιας και πριν επεξεργαστούμε τα δεδομένα μας πρέπει πρώτα να μπορούμε να τα διαχειριστούμε ορθά. Έτσι χωρίς καλά συστήματα διαχείρισης δεδομένων δεν μπορούμε να εφαρμόσουμε αλγόριθμους εξόρυξης δεδομένων. Οι δύο τομείς ακόμη μοιράζονται πολλά, όπως διαδικτυακές βάσεις δεδομένων (Web databases), προσωρινές ή χωρικές βάσεις δεδομένων κ.α. Ένα αξιοσημείωτο παράδειγμα ενός πετυχημένου συνδυασμού εξόρυξης δεδομένων και βάσεων δεδομένων είναι η μηχανή αναζήτησης Google η οποία εκτελεί εργασίες πολύ γρήγορα, αποδοτικά και με ακριβή αποτελέσματα σε οποιοδήποτε ερώτημα.

3.4. Εφαρμογές εξόρυξης δεδομένων

Ο τομέας της εξόρυξης δεδομένων είχε άμεση εφαρμογή με επιτυχία στο Διαδίκτυο. Το πιο δημοφιλές παράδειγμα εξόρυξης δεδομένων στο διαδίκτυο είναι η Google. Για να γίνει πιο κατανοητή η σημαντικότητα της συνεισφοράς αυτής θα πρέπει να αντιληφθούμε πως ο όγκος της πληροφορίας που υπάρχει μέχρι τώρα στο διαδίκτυο είναι αδύνατο να μετρηθεί με ακρίβεια. Οι σελίδες που κάθε φορά ερευνά η Google δηλώνεται πως είναι περίπου 4.285.199.774. Κάθε ερώτημα στην μηχανή αναζήτησης δεν ξεπερνά σε χρόνο τα δυο δευτερόλεπτα.

Η Google και γενικά ο τομέας της εξόρυξης δεδομένων στο Διαδίκτυο έχουν σήμερα τεράστια επιτυχία γιατί έχουν εκπληρώσει δυο σημαντικούς στόχους. Πρώτα, μπορούν να κάνουν αναζήτηση (με κάθε ερώτημα) σε τόσα πολλά δεδομένα σε πολύ σύντομο χρόνο. Δεύτερον, μπορούν να επιστρέψουν σε κάθε ερώτημα τα πρώτα αποτελέσματα που είναι πιο χρήσιμα. Έτσι τελικά ο χρήστης λαμβάνει γρήγορα και εύκολα μόνο της ουσιώδης πληροφορίας που θέλει. Ο παγκόσμιος ιστός αποτελείται από δισεκατομμύρια σελίδες και τα αποτελέσματα από ένα ερώτημα σε μια μηχανή αναζήτησης μπορεί να επιστρέψει χιλιάδες σελίδες σαν

αποτελέσματα. Η συσταδοποίηση μπορεί να χρησιμοποιηθεί για να ομαδοποιήσει αυτά τα αποτελέσματα σε ένα μικρό αριθμό από ομάδες, καθεμιά από τις οποίες σκιαγραφεί μια διάσταση του ερωτήματος. Για παράδειγμα ένα ερώτημα για μια ταινία μπορεί να επιστρέψει σελίδες ομαδοποιημένες σε κατηγορίες όπως ανασκοπήσεις, τρέιλερ, ηθοποιούς και θέατρα. Κάθε κατηγορία (cluster) μπορεί να διασπαστεί σε υποκατηγορίες (subclusters), παράγοντες ιεραρχικής δομής που στη συνέχεια βοηθούν περισσότερο το χρήστη κατά την αναζήτηση που κάνει. Πολύ σημαντική είναι και η εξόρυξη δεδομένων στο ηλεκτρονικό εμπόριο. Η εξόρυξη γνώσης σε εμπορικές σελίδες είναι πολύ σημαντική στη βελτίωση των παρεχόμενων υπηρεσιών, την ικανοποίηση του κάθε πελάτη αλλά και στην αύξηση των κερδών της επιχείρησης. Η διαχείριση της σχέσης πελάτη και επιχείρησης (customer relationship management) μπορεί να βοηθηθεί από τη χρήση τεχνικών web usage mining. Με αυτό το σκεπτικό δίνεται έμφαση στην προσέλκυση πελατών, στη διατήρηση πελατών, την ανταλλαγή πωλήσεων και τέλος την ενεργή παρουσία πελατών.

Μια κατηγορία πολύ γνωστών εφαρμογών εξόρυξης δεδομένων είναι αυτές του μάρκετινγκ. Αυτό είναι αναμενόμενο μιας και μεγάλες εταιρίες χρησιμοποιούν μεγάλα συστήματα διαχείρισης δεδομένων για να διαχειρίζονται μεγάλο αριθμό πελατών και οικονομικών στοιχείων. Οι επιχειρήσεις συλλέγουν σημαντικό αριθμό πληροφοριών από πελάτες που ήδη έχουν ή ακόμα και από μελλοντικούς. Η συσταδοποίηση μπορεί να χρησιμοποιηθεί για να διαχωρίσει τους πελάτες σε μικρό αριθμό από ομάδες για περαιτέρω ανάλυση και λήψη στρατηγικών αποφάσεων. Σε μια επιχείρηση, η συσταδοποίηση μπορεί να βοηθήσει τους εμπόρους να ανακαλύψουν σημαντικές συστάδες στη βάση δεδομένων των πελατών τους και να τις χαρακτηρίσουν με βάση τα αγοραστικά πρότυπα. Αναζητούν απαντήσεις σε ερωτήματα όπως, τι είναι αυτό που θέλουν οι πελάτες, ποιες είναι οι ανάγκες τους κ.α. Ο τομέας της εξόρυξης δεδομένων έχει συνεισφέρει σημαντικά σε αυτή την κατεύθυνση της ανάλυσης δεδομένων μιας επιχείρησης και της εξαγωγής χρήσιμων συμπερασμάτων για την συμπεριφορά των πελατών.

ΠΡΑΚΤΙΚΟ ΜΕΡΟΣ

Το πρακτικό μέρος αποτελείται από την περιγραφή των τεχνολογιών που χρησιμοποιήθηκαν, την περιγραφή της Βάσης Δεδομένων και την εξήγηση των εκάστοτε πινάκων του, καθώς και από την λεπτομερή περιγραφή της υλοποίησης των αλγορίθμων που κατασκευάστηκαν. Τέλος θα αναλύσουμε την δομή της ιστοσελίδας της εργασίας Exodus και των κύριων λειτουργιών του μέσα από την παρουσίαση των κομματιών λογικής και διεπαφής που το απαρτίζουν.

4. Ανάλυση Τεχνολογιών

4.1. Ανάλυση τεχνολογιών που χρησιμοποιήθηκαν

Η δημιουργία της ιστοσελίδας Exodus βασίστηκε στις τελευταίες web τεχνολογίες που υπάρχουν αυτήν την στιγμή στον τομέα του web developing και web designing. Αυτές οι τεχνολογίες είναι οι παρακάτω:

Γλώσσες προγραμματισμού

- Γλώσσα προγραμματισμού C# .NET 4.0
- Γλώσσες προγραμματισμού Javascript & JQuery
- Γλώσσα διαμόρφωσης περιεχομένου CSS
- Γλώσσα ερωτημάτων βάσης δεδομένων MS SQL

Εργαλεία υλοποίησης

- Microsoft Visual Studio 2012
- Microsoft Sql Server 2008 R2

Στις ενότητες που ακολουθούν θα παρουσιάσουμε συνοπτικά τις παραπάνω γλώσσες προγραμματισμού και θα αναφερθούμε εν συντομία στα εργαλεία υλοποίησης με σκοπό την κατατόπιση του αναγνώστη για το σύνολο της διαδικασίας παραγωγής του προγράμματος(ιστοσελίδας Exodus) της εργασίας αυτής.

4.1.1. Η γλώσσα προγραμματισμού C# & .Net 4.5 Framework

Το **.NET** είναι μια νέα πλατφόρμα ανάπτυξης εφαρμογών σε περιβάλλοντα Windows. Σύμφωνα με την Microsoft, έχει ως σκοπό την απλοποίηση της ανάπτυξης εφαρμογών «κρύβοντας» τις τεχνικές λεπτομέρειες υλοποίησης πολλών λειτουργιών, όπως διαχείριση μνήμης, επικοινωνία μέσω δικτύου, είσοδο/έξοδο από συσκευές και αφήνοντας το προγραμματιστή ελεύθερο να επικεντρωθεί στην «λογική» του προγράμματος. **.NET** χαρακτηρίζεται ως managed πλατφόρμα με την έννοια ότι δημιουργεί ένα ελεγχόμενο και ασφαλές περιβάλλον μέσα στο οποίο μπορεί να τρέξει μια εφαρμογή. Η ασφάλεια έγκειται για παράδειγμα στον έλεγχο στην δέσμευση και προσπέλαση της μνήμης (δεν υπάρχουν pointers, δεν μπορείς να προσπελάσεις μια θέση μνήμης εκτός πίνακα), στο τύπο των μεταβλητών και δεδομένων (δεν μπορείς να θέσεις μια float τιμή σε μια ακέραια μεταβλητή) ή στην αυτόματη υλοποίηση δικλείδων ασφαλείας. Το **.NET** υποστηρίζει πληθώρα γλωσσών προγραμματισμού οι οποίες είναι ειδικά σχεδιασμένες για αυτό, όπως C#, Visual Basic.NET, J++ και managed C++. Στην πραγματικότητα, το **.NET** καταλαβαίνει μόνο μια γλώσσα προγραμματισμού την Microsoft Intermediate Language (MSIL). Συνεπώς, οποιαδήποτε γλώσσα προγραμματισμού μπορεί να μεταγλωττιστεί σε MSIL μπορεί να τρέξει στην πλατφόρμα **.NET**. Ο χρήστης μπορεί ακόμα να γράψει απευθείας ένα πρόγραμμα σε MSIL στο Notepad να το κάνει compile και να το τρέξει στο **.NET**.

Η **C#** είναι κατά κύριο λόγο μια αντικειμενοστρεφής γλώσσα προγραμματισμού που ενσωματώνει ωστόσο μερικά χαρακτηριστικά διαφόρων προγραμματιστικών προτύπων. Αναπτύχθηκε στη Microsoft, από μια ομάδα κάτω από την ηγεσία του Anders Hejlsberg, σαν μέρος του **.NET Framework**. Η **C#** είναι μια απλή γλώσσα που μπορεί να χρησιμοποιηθεί χωρίς εντατική εκμάθηση, ενώ ταυτόχρονα είναι εναρμονισμένη με σύγχρονες προγραμματιστικές πρακτικές. Οι θεμελιώδεις αρχές της γλώσσας μπορούν να κατανοηθούν γρήγορα κάτι που σημαίνει ότι οι προγραμματιστές θα είναι παραγωγικοί σε σύντομο χρονικό διάστημα.

Το **.NET** έχει πολλά πλεονεκτήματα για την ανάπτυξη εφαρμογών:

- Είναι εγγενώς αντικειμενοστραφής πλατφόρμα.
- Είναι ανεξάρτητο από γλώσσα προγραμματισμού. Σε μια εφαρμογή ένας προγραμματιστής μπορεί να γράφει κώδικα σε C#, άλλος σε VB.NET και άλλος σε managed C++ και τα τμήματα που αναπτύσσει ο καθένας να συνεργάζονται μεταξύ τους χωρίς προβλήματα.
- Η χρήση βιβλιοθηκών (assemblies) κάνει πολύ εύκολη την επαναχρησιμοποίηση κώδικα.
- Παρέχει πολύ εύκολη εγκατάσταση. Αρκεί να αντιγράψουμε το κατάλογο της εφαρμογής σε ένα άλλο υπολογιστή και αυτή θα τρέξει άμεσα. Δεν υπάρχει installation, δεν πειράζει το registry.
- Παρέχει πληθώρα έτοιμων λειτουργιών που κάνουν την ανάπτυξη κώδικα πολύ εύκολη.
- Αυτοματοποιημένη διαχείριση μνήμης, ο χρήστης δεν χρειάζεται να ασχοληθεί με αποδέσμευση μνήμης.

4.1.2. Η γλώσσες Javascript & JQuery

Η γλώσσα προγραμματισμού **Javascript** αναπτύχθηκε από την εταιρεία Netscape, σε συνεργασία με την SunMicrosystems και η πρώτη της έκδοση δημοσιεύτηκε το 1995. Ακολούθησε η αντίστοιχη γλώσσα της Microsoft η οποία ονομάστηκε Javascript και η επόμενη έκδοση της Javascript που είχε το όνομα ECMAScript που αργότερα όμως καθιερώθηκε με το όνομα που είναι γνωστό μέχρι σήμερα. Η Javascript είναι μία διερμηνευμένη (interpreted) γλώσσα προγραμματισμού με ιδιότητες αντικειμενοστραφούς γλώσσας προγραμματισμού, χωρίς όμως να μπορεί να χαρακτηριστεί ως πλήρης αντικειμενοστραφής.

Η γλώσσα αυτή, κτίστηκε ουσιαστικά πάνω στο πρότυπο των γλωσσών C, C++ και Java. Από την άλλη όμως έχει μία πολύ σημαντική διαφορά στο ότι διαχειρίζεται τους τύπους δεδομένων πιο χαλαρά (loosely typed) σε σχέση με τη σφικτή διαχείριση τύπων δεδομένων (strongly typed) που γίνεται στις προαναφερόμενες γλώσσες. Στην Javascript οι μεταβλητές δεν είναι απαραίτητο να έχουν ένα συγκεκριμένο τύπο ή ακόμη είναι δυνατόν να αλλάζουν τύπο κατά τη διάρκεια της ζωής τους. Επίσης, δεν πρέπει να συγχέεται η Javascript με την Java της SunMicrosystems. Η μία δεν έχει καμία σχέση με την άλλη. Η χρήση του ονόματος Javascript έγινε για λόγους προώθησης της γλώσσας σε μία εποχή που η εξάπλωση της Java ήταν πολύ μεγάλη. Η γλώσσα JavaScript χρησιμοποιείται κυρίως για την εξυπηρέτηση των παρακάτω σκοπών:

- *Λιγότερος φόρτος των server.* Ο έλεγχος και η επικύρωση των δεδομένων που εισάγονται από τους χρήστες γίνεται από τη μεριά του browser κι έτσι δεδομένα τα οποία δεν είναι σε κατάλληλη μορφή δεν αποστέλλονται στον server. Αυτό όμως δεν σημαίνει ότι ο έλεγχος δεν πρέπει να γίνεται και στη μεριά των εξυπηρετητών καθώς κάποιος χρήστης μπορεί να μην έχει ενσωματωμένη την Javascript στον browser του ή υπάρχει πιθανότητα να την έχει απενεργοποιήσει.
- *Άμεση αλληλεπίδραση με τους χρήστες.* Με την χρήση της Javascript για τον έλεγχο των δεδομένων μειώνονται οι χρόνοι αναμονής του χρηστών αφού αυτοί δεν χρειάζεται να περιμένουν μεγάλα χρονικά διαστήματα επαναφόρτωσης της σελίδας σε περίπτωση που έχουν ξεχάσει να εισάγουν κάποιο δεδομένο ή έχουν εισάγει κάτι λάθος.
- *Αυτόματη διόρθωση λαθών.* Ένα παράδειγμα που μπορεί να κάνει περισσότερο κατανοητό το πώς μπορεί να χρησιμοποιηθεί η JavaScript με αυτόν τον τρόπο είναι αυτό της ημερομηνίας. Πολλά συστήματα βάσεων δεδομένων αποθηκεύουν δεδομένα ημερομηνιών σε μορφή dd-mm-yyyy . Αν κάποιος χρήστης εισάγει κάποια ημερομηνία σε μορφή dd/mm/yyyy τότε κάτι τέτοιο θα μπορούσε να ανιχνευτεί αυτόματα από τον browser και να μετατραπεί στην σωστή μορφή πριν τα δεδομένα αποσταλούν στον server.
- *Αυξημένη χρηστικότητα.* Αυτό επιτυγχάνεται επιτρέποντας στον χρήστη την αλλαγή και αλληλεπίδραση με το γραφικό περιβάλλον χωρίς την επαναφόρτωση της σελίδας. Ένα τέτοιο παράδειγμα είναι τα πτυσσόμενα μενού.
- *Αυξημένη δυνατότητα αλληλεπίδρασης.* Ένα τέτοιο παράδειγμα όπου κάτι τέτοιο επιτυγχάνεται είναι τα μενού τα οποία αλληλεπιδρούν όταν ο χρήστης περάσει το mouse πάνω από αυτά – η λειτουργία hover – κάτι το οποίο έχει ως αποτέλεσμα να

δημιουργηθεί μία σειρά από γεγονότα τα οποία έχουν προγραμματιστεί να λειτουργούν με έναν συγκεκριμένο τρόπο.

Η **jQuery** είναι μια ελαφριά βιβλιοθήκη Javascript, συμβατή με όλους τους φυλλομετρητές (browser) που κυκλοφορούν, η οποία απλοποιεί την εκμάθηση και την χρήση της γλώσσας Javascript που χρησιμοποιείται στην δημιουργία ιστοσελίδων και web εφαρμογών. Με την χρήση της μπορούμε να προσθέσουμε κίνηση (animation), να αυξήσουμε την διαδραστικότητα του χρήστη (user interaction), να αλλάξουμε το περιεχόμενο της σελίδας χωρίς ο χρήστης να πρέπει να μεταφερθεί σε νέα σελίδα, να δημιουργήσουμε διάφορα εφέ και πολλά περισσότερα. Η jQuery δεν κάνει μόνο την χρήση της Javascript πιο εύκολη και λιγότερη χρονοβόρα, αλλά εξαλείφει και τα προβλήματα που συνεπάγονται την χρήση της Javascript, όπως το πρόβλημα συμβατότητας της με τους διάφορους browsers της αγοράς. Παρακάτω αναφέρουμε τα κύρια πλεονεκτήματα της γλώσσας jquery:

- *Ακολουθεί την αρχή KISS (Keep It Simple).* Η βιβλιοθήκη JQuery προσπαθεί να υπεραπλουστεύσει τον προγραμματισμό σε Javascript προσφέροντας πραγματικά απλούς μηχανισμούς και εντολές μέσω του framework της.
- *Παρέχει πλήρη και αναλυτικότερη τεκμηρίωση.* Η τεκμηρίωση αυτή συμπληρώνεται από την εκτεταμένη παρουσία ηλεκτρονικών βοηθημάτων. Εκτός από την πολύ καλοδομημένη τεκμηρίωσή του JQuery, οι ενδιαφερόμενοι μπορούν να ανατρέξουν και στην σελίδα του [Visual JQuery](#) όπου μπορούν να βρουν μία εναλλακτική αλλά πολύ βολική, από άποψη δομής, τεκμηρίωση.
- *Υποστηρίζεται από μία πάρα πολύ ενεργή κοινότητα.* Όπως για τα περισσότερα open source έργα λογισμικού, έτσι και για το JQuery η ύπαρξη μιας κατά το μέγιστο δυνατο ενεργής κοινότητας αποτελεί τον ακρογωνιαίο λίθο για την ανάπτυξη και ευημερία του.
- *Μικρό μέγεθος.* Το γεγονός ότι το βασικό πακέτο της JQuery είναι μόλις 20Kb αφενός επιβεβαιώνει την πρώτη παρατήρηση, ότι δηλαδή η φιλοσοφία της έγκειται στην απλότητα και αφετέρου κάνει πολύ εύκολη την κατανόηση της αρχιτεκτονικής της.
- *Ποικιλία χαρακτηριστικών.* Η JQuery δίνει τη δυνατότητα στον χρήστη να χρησιμοποιήσει σχεδόν το σύνολο των δυνατοτήτων που προσφέρει η γλώσσα JavaScript. Από απλά χαρακτηριστικά που σχετίζονται με βασικές λειτουργίες εμφάνισης / απόκρυψης, ως Ajax κλήσεις και σύνθετα εφέ.
- *Επεκτασιμότητα.* Η λογική με την οποία είναι φτιαγμένη η JQuery είναι απλή πράγμα που αντικατοπτρίζεται και στον ίδιο της τον κώδικα. Αυτό κάνει πολύ εύκολη την επέκτασή - τροποποίησή της.

4.1.3. Γλώσσα διαμόρφωσης περιεχομένου CSS

Η **CSS (Cascading Site Sheets-Διαδοχικά Φύλλα Στυλ)** ή (αλληλουχία φύλλων στυλ) είναι μια γλώσσα υπολογιστή που ανήκει στην κατηγορία των γλωσσών φύλλων στυλ που χρησιμοποιείται για τον έλεγχο της εμφάνισης ενός εγγράφου που έχει γραφτεί με μια γλώσσα σήμανσης. Χρησιμοποιείται δηλαδή για τον έλεγχο της εμφάνισης ενός εγγράφου που γράφτηκε στις γλώσσες HTML και XHTML, δηλαδή για τον έλεγχο της εμφάνισης μιας ιστοσελίδας και γενικότερα ενός ιστοτόπου. Η CSS είναι μια γλώσσα υπολογιστή προορισμένη να αναπτύσσει στιλιστικά μια ιστοσελίδα δηλαδή να διαμορφώνει περισσότερα χαρακτηριστικά, χρώματα, στοίχιση και δίνει περισσότερες δυνατότητες σε σχέση με την html. Για μια όμορφη και καλοσχεδιασμένη ιστοσελίδα η χρήση της CSS κρίνεται ως απαραίτητη.

Τα διαδοχικά Φύλλα Στυλ (CSS, Cascading Style Sheets) αποτελούν ένα πολύ καλό εργαλείο για να μπορούμε να αλλάζουμε την εμφάνιση και τη διάταξη (layout) των ιστοσελίδων μας. Μπορούν να μας γλυτώσουν από πολύ χρόνο και κόπο και μας δίνουν τη δυνατότητα να σχεδιάζουμε τις ιστοσελίδες μας με μια εντελώς καινούργια φιλοσοφία. Η κατανόηση των CSS απαιτεί να υπάρχει κάποια βασική εμπειρία με την HTML. Μπορούμε να χρησιμοποιήσουμε όποιον φυλλομετρητή (browser) επιθυμούμε για να βλέπουμε πώς θα εμφανίζονται οι ιστοσελίδες που θα δημιουργούμε. Συνεπώς αυτό που χρειαζόμαστε είναι ένας φυλλομετρητής (browser) και ένας απλός επεξεργαστής κειμένου (text editor). Παρακάτω παραθέτουμε τα κυριότερα πλεονεκτήματα της css.

- *Πολύ μεγαλύτερη ευελιξία.* Το CSS κατέστησε εφικτές μορφοποιήσεις οι οποίες ήταν αδύνατες ή πολύ δύσκολες με την κλασική HTML.

- *Ευκολότερη συντήρηση των ιστοσελίδων.* Η εμφάνιση ενός ολόκληρου στην ιστοσελίδα μπορεί να ελέγχεται από ένα μόνο εξωτερικό αρχείο CSS. Έτσι, κάθε αλλαγή στο στυλ της ιστοσελίδας μπορεί να γίνεται με μια μοναδική αλλαγή σε αυτό το αρχείο, αντί για την επεξεργασία πολλών σημείων σε κάθε σελίδα που υπάρχει στην ιστοσελίδα.
- *Μικρότερο μέγεθος αρχείου.* Δεδομένου ότι ο κάθε κανόνας μορφοποίησης γράφεται μόνο μια φορά και όχι σε κάθε σημείο που εφαρμόζεται.
- *Καλύτερο SEO (Search engine optimization).* Οι μηχανές αναζήτησης δεν «μπερδεύονται» ανάμεσα σε περιεχόμενο και τη μορφοποίηση του, αλλά έχουν πρόσβαση στο περιεχόμενο σκέτο, οπότε είναι πολύ ευκολότερο να το καταγράψουν και να το αρχειοθετήσουν (indexing).
- *Γρηγορότερες σελίδες.* Όταν χρησιμοποιούμε εξωτερικό αρχείο CSS, ο browser την πρώτη φορά που θα φορτώσει κάποια σελίδα του website μας το αποθηκεύει στην cache, οπότε δεν χρειάζεται να το κατεβάσει ξανά κάθε φορά που κατεβάζει ο χρήστης του κάποια άλλη σελίδα του website μας.

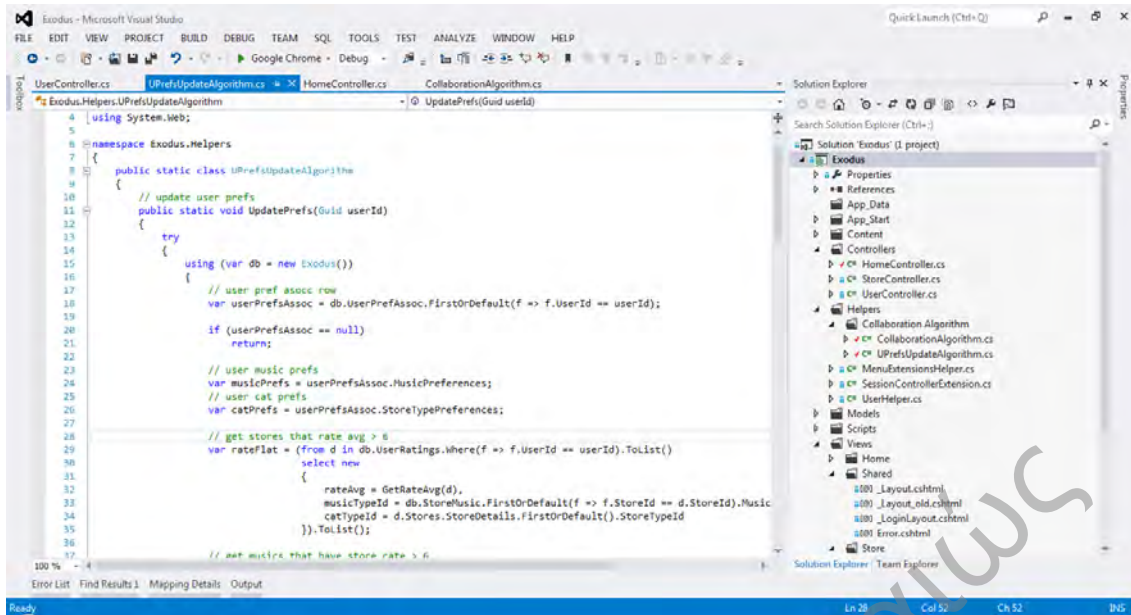
4.1.4. Γλώσσα ερωτημάτων βάσης δεδομένων SQL

Η **SQL** μία γλώσσα υπολογιστών στις βάσεις δεδομένων, που σχεδιάστηκε για τη διαχείριση δεδομένων, σε ένα σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων (Relational Database Management System, RDBMS) και η οποία, αρχικά, βασίστηκε στη σχεσιακή άλγεβρα. Η γλώσσα περιλαμβάνει δυνατότητες ανάκτησης και ενημέρωσης δεδομένων, δημιουργίας και τροποποίησης σχημάτων και σχεσιακών πινάκων, αλλά και ελέγχου πρόσβασης στα δεδομένα. Η SQL ήταν μία από τις πρώτες γλώσσες για το σχεσιακό μοντέλο του *Edgar F. Codd*, στο σημαντικό άρθρο του το 1970, και έγινε η πιο ευρέως χρησιμοποιούμενη γλώσσα για τις σχεσιακές βάσεις δεδομένων.

4.1.5. Το Microsoft Visual Studio 2012

Το Microsoft Visual Studio είναι ένα ολοκληρωμένο περιβάλλον ανάπτυξης (IDE) από τη Microsoft. Χρησιμοποιείται για την ανάπτυξη προγραμμάτων ηλεκτρονικών υπολογιστών για τα Microsoft Windows, καθώς και websites, εφαρμογών και υπηρεσιών web. Το Visual Studio χρησιμοποιεί τις πλατφόρμες ανάπτυξης λογισμικού της Microsoft, όπως τα Windows API, Windows Forms, Windows Presentation Foundation, Windows Store και Microsoft Silverlight. Μπορεί να παράγει τόσο εγγενή κώδικα και διαχειριζόμενο κώδικα.

Το Visual Studio περιλαμβάνει την λειτουργία IntelliSense η οποία βοηθά τον προγραμματιστή στην συγγραφή κώδικα καθώς και την λειτουργία scaffolding η οποία επιτρέπει την γρήγορη συγγραφή τυποποιημένων. Το ολοκληρωμένο πρόγραμμα εντοπισμού σφαλμάτων λειτουργεί τόσο ως ένα πρόγραμμα εντοπισμού σφαλμάτων πηγαίου κώδικα και ένα πρόγραμμα εντοπισμού σφαλμάτων κατά την εκτέλεση. Περιέχονται και κάποια ενσωματωμένα εργαλεία για τη δημιουργία εφαρμογών GUI, web designer, σχεδιαστής κλάσεων, και σχεδιαστής βάσεων δεδομένων. Το Visual Studio υποστηρίζει διαφορετικές γλώσσες προγραμματισμού και επιτρέπει τον επεξεργαστή κώδικα και εντοπισμού σφαλμάτων την υποστήριξη (σε διάφορους βαθμούς) σχεδόν σε οποιαδήποτε γλώσσα προγραμματισμού, εφόσον υπάρχει μια συγκεκριμένη γλώσσα υπηρεσία. Built-in γλώσσες περιλαμβάνουν C, C++ και C++ / CLI (μέσω του Visual C++), VB.NET (μέσω της Visual Basic NET), C# (μέσω του Visual C#), και F# (όπως του Visual Studio 2010). Επίσης παρέχει υποστήριξη για άλλες γλώσσες, όπως η M, Python και Ruby. Υποστηρίζει, επίσης, XML/XSLT, HTML/ XHTML, JavaScript και CSS. Το περιβάλλον το VS 2012 φαίνεται στην παρακάτω εικόνα.

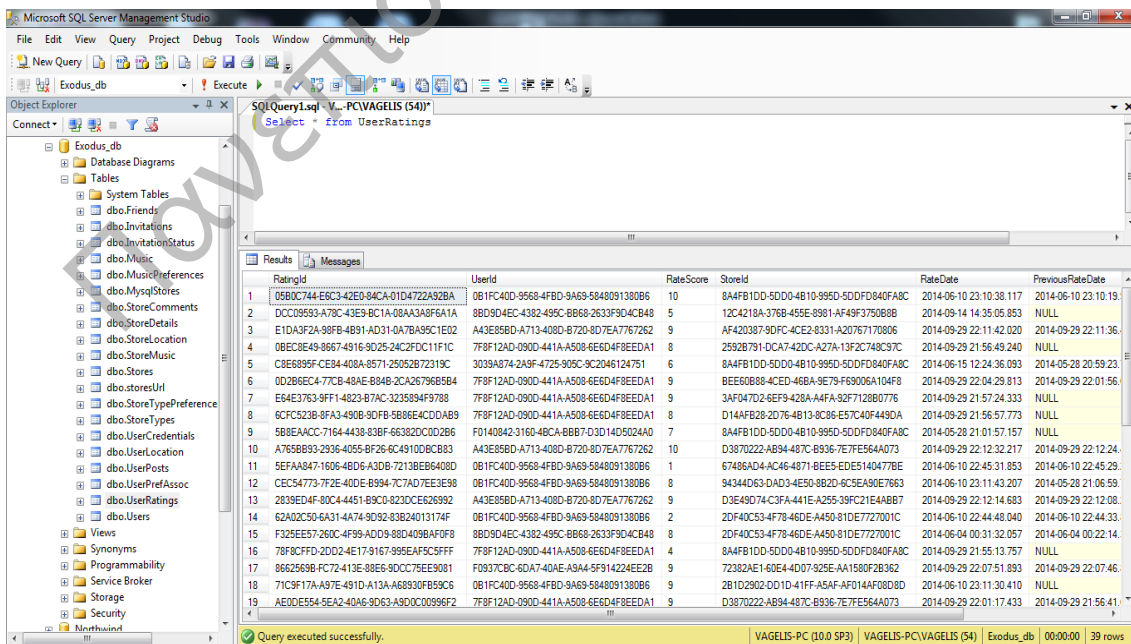


Εικόνα 4-1. Visual studio 2012 IDE

4.1.6. Το MS Sql studio 2008

Το SQL Server Management Studio(SSMS) είναι ένα ολοκληρωμένο περιβάλλον διαχείρισης βάσεων δεδομένων που προσφέρει ένα σύστημα διεπαφής με τον χρήστη για την επικοινωνία με τον MS Sql server. Η Microsoft προσέφερε το σύστημα αυτό με τον sql server 2005 και συνέχισε να είναι κομμάτι αυτού και στις εκδόσεις 2008 και 2012. Το SSMS παρέχει εργαλεία για την διαμόρφωση και την διαχείριση του sql server για την δημιουργία των βάσεων δεδομένων είτε μέσω του περιβάλλοντος του SSMS είτε με την βοήθεια των εντολών της SQL.

Το SSMS περιέχει διαχείριση πολλών οντοτήτων σχεσιακών βάσεων δεδομένων και παρέχει αυτοματοποιημένες διαδικασίες μέσω έτοιμων μπλοκ κώδικα για την δημιουργία, εργώτηση καθώς και την δημιουργία αντιγράφων ασφαλείας μιας σχεσιακής βάσης δεδομένων. Στην παρακάτω εικόνα φαίνεται το γραφικό περιβάλλον της εφαρμογής SSMS.



Εικόνα 4-2. MS Sql Management studio

4.2. Βάση Δεδομένων

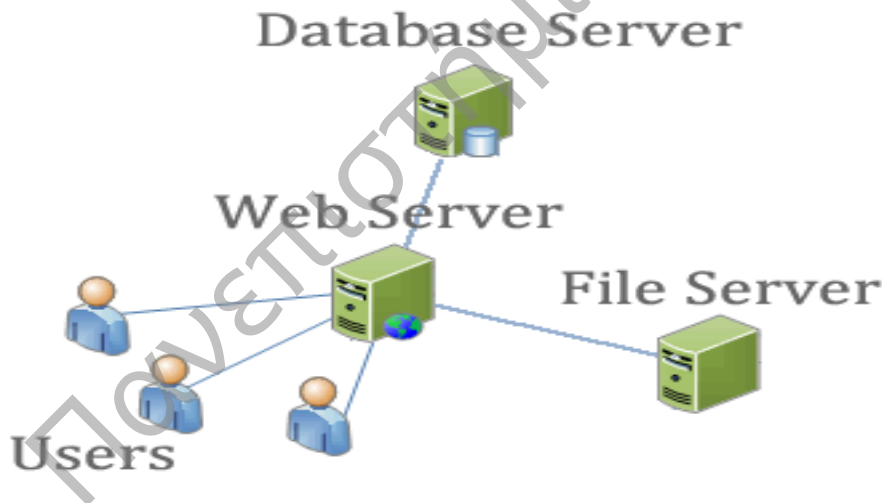
Για την υλοποίηση της εργασίας χρειάστηκε να αποθηκεύσουμε κάποια δεδομένα που αφορούν τους χρήστες και τις σχέσεις μεταξύ τους αλλά και την πληροφορία των προτιμήσεων τους. Όπως βλέπουμε στην παρακάτω εικόνα (εικόνα 4-3) για την λειτουργία μιας ιστοσελίδας, χρειάζεται εκτός από τον file server στον οποίο αποθηκεύονται τα δομικά αρχεία της και ένας database server δηλαδή μια βάση δεδομένων όπου αποθηκεύονται όλα τα απαραίτητα δεδομένα. Η δομή αυτή επιτρέπει στην ιστοσελίδα να περιλαμβάνει δυναμικό περιεχόμενο και να καλύπτει τις ανάγκες για αποθήκευση και επεξεργασία δεδομένων σε μεταγενέστερο χρόνο από το χρόνο δημιουργίας τους.

Δημιουργήσαμε για αυτόν το σκοπό μια σχεσιακή βάση δεδομένων MS SQL SERVER. Ο MS SQL SERVER, είναι ένα σύστημα διαχείρισης σχεσιακής βάσης δεδομένων (relational database management system, RDBMS) που επιτρέπει την ταυτόχρονη πρόσβαση στην βάση δεδομένων από πολλούς χρήστες. Επίσης επιτρέπει την επεξεργασία των δεδομένων αυτών μέσω της εκτέλεση ερωτημάτων πάνω στα δεδομένα με την βοήθεια της γλώσσας MSSQL.

Για τις ανάγκες αυτής της εργασίας επιλέξαμε την έκδοση MSSQL SERVER 8 η οποία συνοδεύεται από το γραφικό περιβάλλον MSSQL Management studio. Το MS SQL Management studio είναι ένα πρόγραμμα με σχετικά φιλικό για τον χρήστη περιβάλλον που επιτρέπει την εύκολη αλληλεπίδραση με τις βάσεις δεδομένων. Το περιβάλλον παρέχει δύο τρόπου αλληλεπίδρασης, την εκτέλεση SQL ερωτημάτων και την δημιουργία των ερωτημάτων αυτών μέσω του γραφικού περιβάλλοντος.

Η γλώσσα με την οποία συνεργάζεται το παραπάνω rdbms ονομάζεται T-SQL και είναι πανομοιότυπη με την SQL με μερικές διαφορές στην σύνταξη. Η γλώσσα αυτή επιτρέπει την συγγραφή Sql ερωτημάτων (sql queries) που εκτελούν τις CRUD (Create, Read, Update, Delete) οι οποίες αποτελούν την βάση αλληλεπίδρασης με μία σχεσιακή βάση δεδομένων όπως αυτές που επιλέξαμε.

Η συνεργασία του rdbms με την γλώσσα την οποία αυτό «καταλαβαίνει» μας βοηθά να δημιουργήσουμε πολύ σύνθετα ερωτήματα η εκτέλεση των οποίων μας επιστρέφει τα επιθυμητά δεδομένα για την εξυπηρέτηση της λειτουργίας της ιστοσελίδας μας.



Εικόνα 4-3. Αρχιτεκτονική server

4.2.1. Αναλυτική περιγραφή πινάκων

Σε αυτή την ενότητα θα αναφέρουμε και θα περιγράψουμε τους σχεσιακούς πίνακες που δημιουργήσαμε για την εκπόνηση της παρούσας εργασίας. Θα αναφέρουμε την χρήση τους στην λειτουργία του προγράμματος με σκοπό την κατανόηση του ρόλου τους στην διαδικασία του λειτουργίας της ιστοσελίδας μας και την χρήση τους από τον αλγόριθμο που έχουμε υλοποιήσει.

Η ιστοσελίδα μας, όπως είδαμε στην προηγούμενη ενότητα χρησιμοποιεί μία MSSQL βάση δεδομένων. Η σχεσιακή βάση ονομάζεται «Exodus_db». Αποτελείται από δεκαοκτώ πίνακες στους οποίους γίνεται χρήση πρωτευόντων κλειδιών(primary keys) καθώς και ξένων κλειδιών(foreign keys) για να επιτευχθούν οι σχέσεις κάθε πίνακα με έναν ή περισσότερους πίνακες.

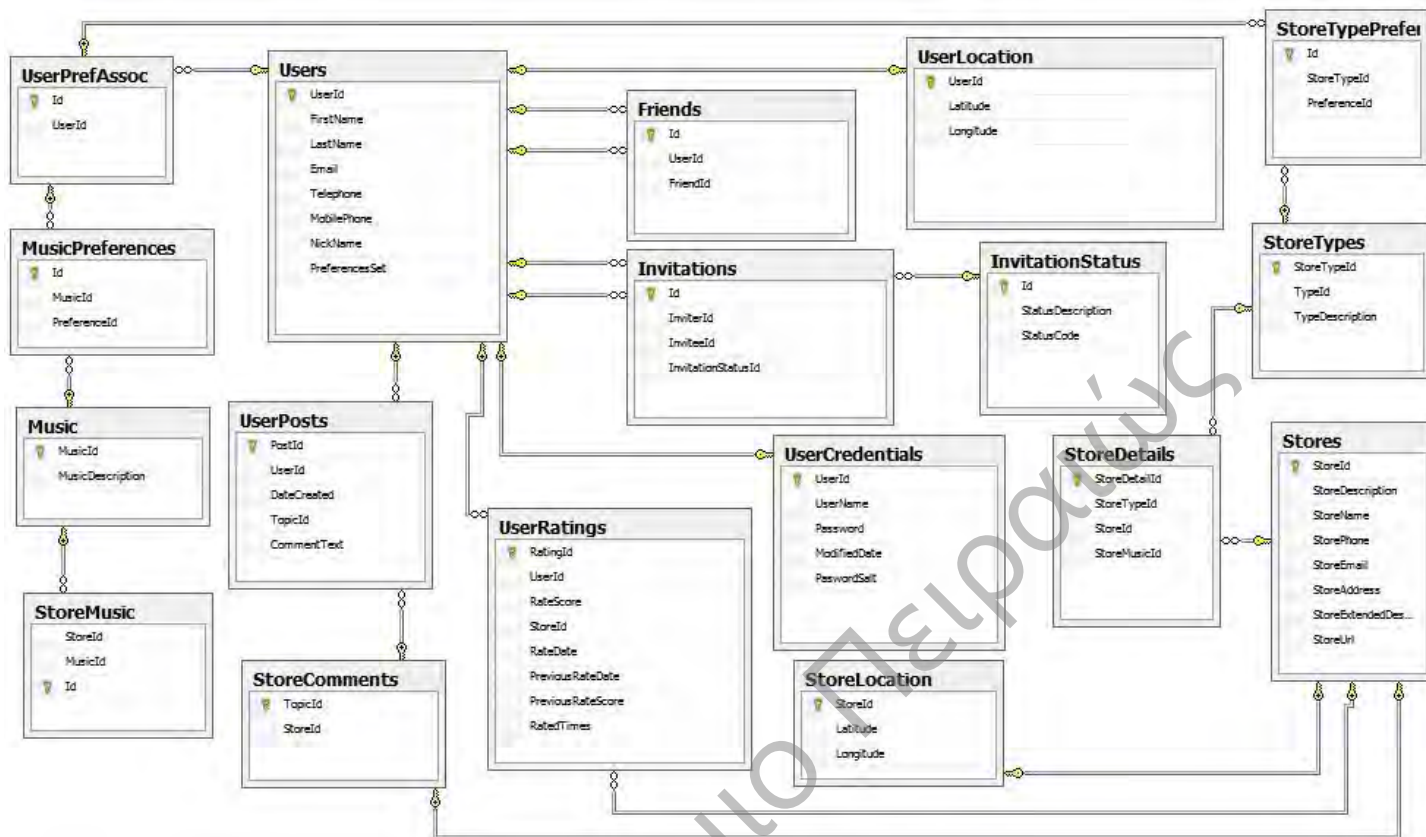
Η χρήση των πρωτευόντων κλειδιών(primary keys) εξυπηρετεί στην ταχύτερη αναζήτηση στους πίνακες αυτούς. Η χρήση τους μειώνει κατά πολύ τον χρόνο εύρεσης των επιθυμητών εγγραφών και άρα καθιστά και την ιστοσελίδα μας πιο γρήγορη. Για αυτό τον λόγο χρησιμοποιήσαμε και την λειτουργία των ευρετηρίων(dictionaries) τα οποία εξυπηρετούν την ταχύτητα των queries σε παραπάνω από μία στήλη.

Οι πίνακες που απαρτίζουν την βάση Exodus_db φαίνονται στον παρακάτω πίνακα με μία σύντομη περιγραφή.

| Όνομα πίνακα | Περιγραφή |
|----------------------|--|
| Friends | Σχέσεις των χρηστών |
| Invitations | Προσκλήσεις φιλίας χρηστών |
| InvitationStatus | Κατάσταση αιτήματος φιλίας |
| Music | Είδη μουσικής καταστημάτων |
| MusicPreferences | Προτιμήσεις χρήστη στην μουσική |
| StoreComments | Σχόλια χρηστών |
| StoreDetails | Λεπτομέρειες καταστημάτων |
| StoreLocation | Τοποθεσία καταστημάτων |
| StoreMusic | Συσχετισμός μουσικής-καταστημάτων |
| Stores | Καταστήματα |
| StoreTypePreferences | Προτιμήσεις χρήστη στο είδος |
| StoreTypes | Είδη καταστημάτων |
| UserCredentials | Στοιχεία σύνδεσης χρήστη |
| UserLocation | Στοιχεία τοποθεσίας χρηστών |
| UserPosts | Θεματικές ενότητες σχολίων |
| UserPrefAssoc | Συσχετιστικός πίνακας χρήστη-προτιμήσεων |
| Users | Χρήστες |

Πίνακας 4-1. Πίνακες Exodus_db

Το διάγραμμα πινάκων της βάσης Exodus_db καθώς και οι σχέσεις τους φαίνονται στο παρακάτω διάγραμμα (εικόνα 4-4).



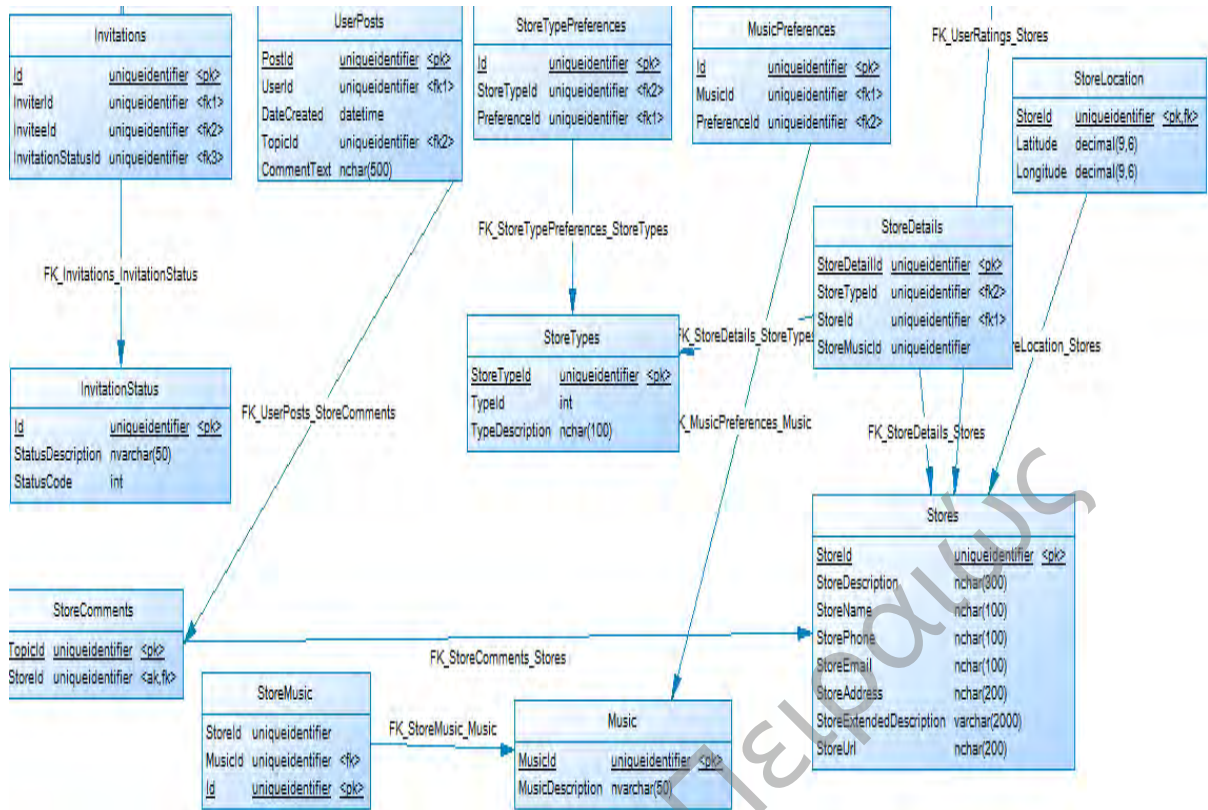
Εικόνα 4-4. Διάγραμμα σχέσεων Exodus_db

Αναλυτική περιγραφή πινάκων:

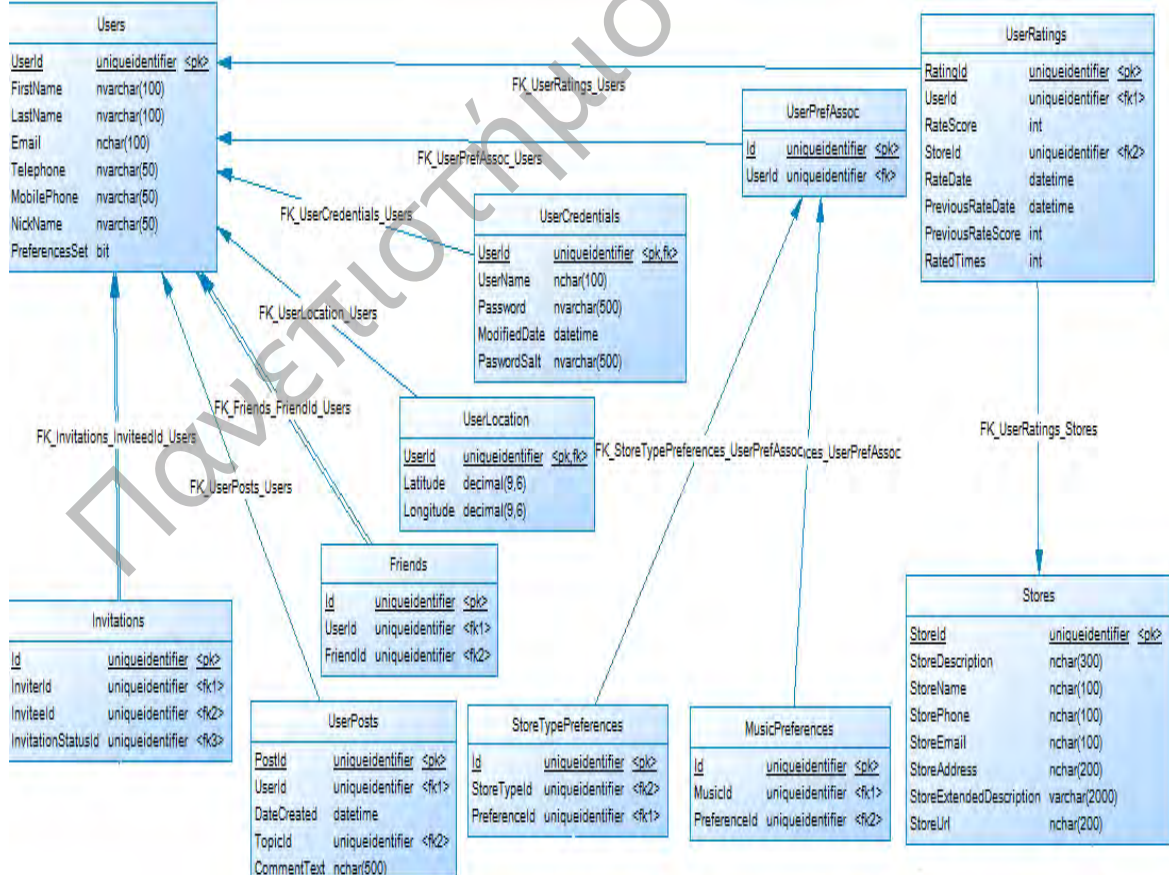
- **Users**: Αυτός ο πίνακας περιέχει τα προσωπικά στοιχεία που αφορούν τον χρήστη όπως το όνομα, το επώνυμο κ.λπ. Σημαντικό εδώ είναι να αναφέρουμε το πεδίο PreferencesSet. Το πεδίο PreferencesSet είναι ένα πεδίο τύπου bit και παίρνει τιμές 0,1. Σε αυτό το πεδίο αποθηκεύουμε την πληροφορία η οποία ορίζει αν ο χρήστης μας έχει αρχικοποιήσει τον λογαριασμό του και έχει συμπληρώσει την ανάλογη φόρμα ενεργοποίησης που ζητείται κατά την είσοδο του στην ιστοσελίδα. Αυτή η φόρμα όπως θα δούμε σε επόμενη ενότητα αφορά τις αρχικές προτιμήσεις του χρήστη μας.
- **UserCredentials**: Αυτός ο πίνακας περιέχει τα στοιχεία σύνδεσης των χρηστών δηλαδή το Όνομα χρήστη(username) και τον Κωδικό(password). Κατά την εγγραφή ενός νέου χρήστη δημιουργείται μία καινούργια εγγραφή σε αυτόν τον πίνακα. Ο κωδικός του χρήστη κρυπτογραφείται με τον ειδικό αλγόριθμο κρυπτογράφησης PBKDF2. Ο αλγόριθμος αυτός δημιουργεί ένα επιπλέον πεδίο PasswordSalt που επίσης χρειάζεται για την αποκρυπτογράφηση του κωδικού.
- **UserLocation**: Σε αυτόν τον πίνακα αποθηκεύεται η τοποθεσία που δηλώνει ο χρήστης κατά την εγγραφή του. Η τοποθεσία αποθηκεύεται με την μορφή γεωγραφικού μήκους και γεωγραφικού πλάτους.
- **UserRating**: Εδώ αποθηκεύονται οι αξιολογήσεις που κάνει ο χρήστης για κάθε κατάσταση. Θα υπάρχουν τόσες εγγραφές για κάθε χρήστη όσο και τα καταστήματα που έχει αξιολογήσει. Στα πεδία RateScore και PreviousRateScore αποθηκεύονται ο βαθμός που έχει αξιολογήσει ο χρήστης μας ένα κατάστημα. Θυμίζουμε ότι κάθε χρήστης επιτρέπεται να αξιολογήσει κάθε κατάστημα το πολύ δύο φορές. Για να έχουμε πιο πλήρη πληροφορία αποθηκεύουμε και την ημερομηνία των αξιολογήσεων.

- UserPosts: Εδώ αποθηκεύονται τα σχόλια των χρηστών για κάθε κατάσταση.
- UserPrefAssoc: Αυτός είναι ένας κυρίως πίνακας για πού βοηθά στην αποθήκευση των προτιμήσεων του χρήστη.
- Stores: Σε αυτόν το πίνακα αποθηκεύουμε τα καταστήματα που είναι εγγεγραμμένα στην ιστοσελίδα μας. Αποθηκεύονται στοιχεία όπως το όνομα, η περιγραφή, τα στοιχεία επικοινωνίας καθώς και η ηλεκτρονική διεύθυνση του καταστήματος.
- StoreTypes: Ο πίνακας αυτός περιέχει τους τύπους των καταστημάτων που φιλοξενούνται στην ιστοσελίδα μας.
- StoreTypePreferences: Εδώ αποθηκεύονται οι τύποι των καταστημάτων που ο κάθε χρήστης προτιμά.
- StoreMusic: Στον πίνακα αυτόν έχουμε τους τύπους της μουσικής που κάθε κατάσταση έχει. Αυτός ο πίνακας είναι βοηθητικός, που σημαίνει ότι ένα κατάστημα μπορεί να έχει πάνω από ένα είδος.
- StoreLocation: Η τοποθεσία του καταστήματος είναι αποθηκευμένη σε αυτόν τον πίνακα με την μορφή γεωγραφικού μήκους και γεωγραφικού πλάτους.
- StoreDetails: Ο πίνακας αυτός είναι ένας συνδετικός πίνακας για τον τύπο του καταστήματος και την μουσική του καταστήματος.
- StoreComments: Αυτός ο πίνακας βοηθά την σύνδεση των σχολίων του καταστήματος με τους χρήστες της ιστοσελίδας.
- Music: Εδώ περιέχονται τα είδη της μουσικής που υπάρχουν στην ιστοσελίδα μας.
- MusicPreferences: Αυτός ο πίνακας αποθηκεύει την μουσική που κάθε χρήστης προτιμά.
- Invitations: Εδώ υπάρχουν τα αιτήματα φιλίας μεταξύ των χρηστών. Συγκεκριμένα αποθηκεύεται ο χρήστης που δημιουργεί το αίτημα φιλίας καθώς και ο χρήστης που δέχεται το αίτημα. Επίσης περιέχεται και η κατάσταση του αιτήματος το οποίο παίρνει τιμές από τον επόμενο πίνακα.
- InvitationStatus: Ο πίνακας αυτός περιέχει τις καταστάσεις των αιτημάτων φιλίας μεταξύ των χρηστών.
- Friends: Στον πίνακα Friends αποθηκεύουμε τους φίλους που έχουμε στην ιστοσελίδα μας. Για να γίνει κάποιος φίλος με κάποιον άλλο θυμίζουμε ότι πρέπει να αποστείλει ένα αίτημα και ο χρήστης στόχος να αποδεχθεί αυτό το αίτημα. Για κάθε ζευγάρι φίλων αποθηκεύονται δύο εγγραφές καθώς η σχέση αυτή είναι αμφίδρομη.

Στις εικόνες που ακολουθούν απεικονίζουμε σχηματικά τους πίνακες της βάσης δεδομένων Exodus_db με την βοήθεια σχεσιακών διαγραμμάτων (εικόνα 4-5, 4-6).



Εικόνα 4-5. Σχισιακό διάγραμμα Exodus_db, 1



Εικόνα 4-6. Σχισιακό διάγραμμα Exodus_db, 2

Παρακάτω παραθέτουμε αναλυτικά τους πίνακες της Exodus_db μαζί με τα χαρακτηριστικά τους.

1. *Users* (*UserID*, *FirstName*, *Email*, *Telephone*, *MobilePhone*, *NickName*, *PreferencesSet*)
 - *UserID*: είναι ο μοναδικός κωδικός που αντιστοιχεί σε κάθε εγγεγραμμένο χρήστη στην ιστοσελίδα Exodus. Ο κωδικός αυτός ανατίθεται στον κάθε χρήστη κατά την εγγραφή του.
 - *FirstName*: το όνομα του χρήστη
 - *LastName*: το επώνυμο του χρήστη
 - *Email*: η διεύθυνση ηλεκτρονικού ταχυδρομείου του χρήστη
 - *Telephone*: Το τηλέφωνο επικοινωνίας του χρήστη
 - *MobilePhone*: το κινητό τηλέφωνο επικοινωνίας του χρήστη
 - *NickName*: Το όνομα του χρήστη που θα εμφανίζεται στην ιστοσελίδα
 - *PreferencesSet*: Το πεδίο αυτό παίρνει τιμές 0 ή 1. Χρησιμοποιείται από το σύστημα για να διευκρινιστεί αν ο εγγεγραμμένος χρήστης έχει ανταποκριθεί στο ερωτηματολόγιο των προτιμήσεων του, όπως παραπέμπεται από το σύστημα. Σε περίπτωση που ο χρήστης δεν έχει συμπληρώσει το ερωτηματολόγιο, τότε δεν μπορεί να πλοηγηθεί στην ιστοσελίδα
2. *UserRatings* (*RatingId*, *UserId*, *RateScore*, *StoreId*, *RateDate*, *PreviousRateScore*, *PreviousRateDate*, *RatedTimes*)
 - *RatingId*: το μοναδικό αναγνωριστικό για τον πίνακα
 - *UserId*: το αναγνωριστικό του χρήστη
 - *StoreId*: το αναγνωριστικό του καταστήματος
 - *RateScore*: η τελευταία βαθμολόγηση του χρήστη
 - *RateDate*: η ημερομηνία της τελευταίας βαθμολόγησης
 - *PreviousRateScore*: η παλαιότερη βαθμολόγηση του χρήστη
 - *PreviousRateDate*: η ημερομηνία της τελευταίας βαθμολόγησης
 - *RatedTimes*: ο αριθμός των βαθμολογήσεων για το κατάστημα
3. *UserPrefAssoc* (*Id*, *UserId*)
 - *Id*: το μοναδικό αναγνωριστικό για τον πίνακα
 - *UserId*: το αναγνωριστικό του χρήστη
4. *UserPosts* (*PostId*, *UserId*, *DateCreated*, *TopicId*, *CommentText*)
 - *PostId*: το μοναδικό αναγνωριστικό για τον πίνακα
 - *UserId*: το αναγνωριστικό του χρήστη
 - *DateCreated*: η ημερομηνία καταχώρησης του σχολίου
 - *TopicId*: το αναγνωριστικό του θέματος του σχολίου
 - *CommentText*: το κείμενο του σχολίου που ο χρήστης παραθέτει
5. *UserLocation* (*UserId*, *Latitude*, *Longitude*)
 - *UserId*: το μοναδικό αναγνωριστικό χρήστη
 - *Latitude*: το γεωγραφικό πλάτος της δηλωθείσας τοποθεσίας του χρήστη
 - *Longitude*: το γεωγραφικό μήκος της δηλωθείσας τοποθεσίας του χρήστη
6. *UserCredentials* (*UserId*, *UserName*, *Password*, *ModifiedDate*, *PasswordSalt*)
 - *UserId*: το μοναδικό αναγνωριστικό χρήστη
 - *UserName*: το όνομα χρήστη για την διαδικασία της ταυτοποίησης
 - *Password*: ο κωδικός ασφαλείας του χρήστη
 - *ModifiedDate*: η ημερομηνία της τελευταίας αλλαγής των στοιχείων ταυτοποίησης του χρήστη
 - *PasswordSalt*: βοηθητικός κωδικός για την ασφάλεια των διαπιστευτηρίων του χρήστη

7. *StoreTypes* (*StoreTypeID*, *TypeId*, *TypeDescription*)
 - *StoreTypeID*: το μοναδικό αναγνωριστικό του πίνακα
 - *TypeId*: το μοναδικό αναγνωριστικό του τύπου του καταστήματος
 - *TypeDescription*: το λεκτικό του τύπου του καταστήματος
8. *StoreTypePreferences* (*Id*, *StoreTypeID*, *PreferencesId*)
 - *Id*: το μοναδικό αναγνωριστικό του πίνακα
 - *StoreTypeID*: το αναγνωριστικό του τύπου του καταστήματος
 - *PreferencesId* : το αναγνωριστικό της εγγραφής των προτιμήσεων του χρήστη
9. *Stores* (*StoreId*, *StoreDescription*, *StoreName*, *StorePhone*, *StoreEmail*, *StoreAddress*, *StoreExtendedDescription*, *StoreUrl*)
 - *StoreId*: το μοναδικό αναγνωριστικό του πίνακα
 - *StoreDescription*: η σύντομη περιγραφή του καταστήματος
 - *StoreName*: το όνομα του καταστήματος
 - *StorePhone*: το τηλέφωνο επικοινωνίας του καταστήματος
 - *StoreEmail*: η διεύθυνση ηλεκτρονικού ταχυδρομείου του καταστήματος
 - *StoreAddress*: η διεύθυνση του καταστήματος
 - *StoreExtendedDescription*: η αναλυτική περιγραφή του καταστήματος
 - *StoreUrl*: η ηλεκτρονική διεύθυνση του καταστήματος, Αυτή μπορεί να είναι η διεύθυνση ιστοσελίδας του καταστήματος ή η διεύθυνση σε ένα μέσο κοινωνική δικτύωσης
10. *StoreMusic* (*Id*, *StoreId*, *MusicId*)
 - *Id*: το μοναδικό αναγνωριστικό του πίνακα
 - *StoreId*: το αναγνωριστικό του καταστήματος
 - *MusicId*: το αναγνωριστικό της εγγραφής του τύπου μουσικής του καταστήματος
11. *StoreLocation* (*StoreId*, *Latitude*, *Longitude*)
 - *StoreId*: το μοναδικό αναγνωριστικό του καταστήματος
 - *Latitude*: το γεωγραφικό πλάτος της δηλωθείσας τοποθεσίας του χρήστη
 - *Longitude* το γεωγραφικό μήκος της δηλωθείσας τοποθεσίας του χρήστη
12. *StoreDetails* (*StoreDetailId*, *StoreTypeId*, *StoreId*, *StoreMusicId*)
 - *StoreDetailId*: το μοναδικό αναγνωριστικό πίνακα
 - *StoreTypeId*: το αναγνωριστικό του τύπου του καταστήματος
 - *StoreId*: το αναγνωριστικό του καταστήματος
 - *StoreMusicId*: το αναγνωριστικό της εγγραφής του τύπου της μουσικής του καταστήματος
13. *StoreComments* (*TopicId*, *StoreId*)
 - *TopicId*: το μοναδικό αναγνωριστικό πίνακα
 - *StoreId*: το αναγνωριστικό του καταστήματος
14. *MusicPreferences* (*Id*, *MusicId*, *PreferencesId*)
 - *Id*: το μοναδικό αναγνωριστικό πίνακα
 - *MusicId*: το αναγνωριστικό της εγγραφής της μουσικής του καταστήματος
 - *PreferencesId*: το αναγνωριστικό των προτιμήσεων του καταστήματος
15. *Music* (*MusicId*, *MusicDescription*)
 - *MusicId*: το μοναδικό αναγνωριστικό πίνακα
 - *MusicDescription*: το λεκτικό του τύπου μουσικής
16. *InvitationStatus* (*Id*, *StatusDescription*, *StatusCode*)
 - *Id*: το μοναδικό αναγνωριστικό πίνακα
 - *StatusDescription*: το λεκτικό της κατάστασης της πρόσκλησης
 - *StatusCode*: ο σύντομος κωδικός της κατάστασης της πρόσκλησης

17. *Invitations (Id, InviterId, InviteeId, InvitationStatusId)*

- *Id*: το μοναδικό αναγνωριστικό πίνακα
- *InviterId*: το αναγνωριστικό του χρήστη που κάνει την πρόσκληση
- *InviteeId*: το αναγνωριστικό του χρήστη που λαμβάνει την πρόσκληση
- *InvitationStatusId*: η κατάσταση του αιτήματος φιλίας

18. *Friends (Id, UserId, FriendId)*

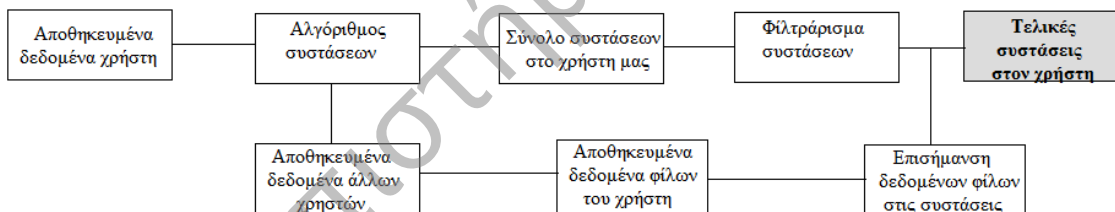
- *Id*: το μοναδικό αναγνωριστικό πίνακα
- *UserId*: το αναγνωριστικό του χρήστη
- *FriendId*: το αναγνωριστικό του φίλου

5. ΥΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ

5.1. Σύντομη περιγραφή αλγορίθμου συστάσεων

Ο αλγόριθμος που υλοποίησε η εργασία αυτή είναι ένας υβριδικός αλγόριθμος που βασίζεται στον συνδυασμό εύρεσης των K-κοντινότερων γειτόνων ως προς τον χρήστη μας και στην συνέχεια στο φιλτράρισμα των εξαχθέντων δεδομένων για προβολή προς αυτόν (εικόνα 5-1). Οι τελικές συστάσεις στον χρήστη εμπλουτίζονται με πληροφορία που αφορά το αν το προταθέν κατάσταση είναι έχει αξιολογηθεί θετικά (βαθμολογία > 6) από τους φίλους του χρήστη μας. Η πληροφορία αυτή παρουσιάζεται στον χρήστη οπτικά μέσω μίας ταμπέλας (label) η οποία ενημερώνει για την πληροφορία το χρήστη.

Η εύρεση των κοντινότερων γειτόνων προς τον χρήστη μας και το φιλτράρισμα αυτών, γίνεται με την ανάκτηση των δεδομένων του χρήστη μας, αλλά και των υπόλοιπων χρηστών της βάσης δεδομένων μας και την εφαρμογή του αλγορίθμου πάνω σε αυτά. Ο αλγόριθμος στηρίζει τις προτάσεις του στο προφίλ που έχει δημιουργήσει για τον χρήστη μας μέσω των προηγούμενων αξιολογήσεων του αλλά και μέσω της άμεσης δημιουργίας προφίλ κατά την δημιουργία του λογαριασμού του.



Εικόνα 5-1. Σχηματική περιγραφή αλγορίθμου συστάσεων.

5.1.1. Σύντομη αναφορά βημάτων αλγορίθμου

Παρακάτω θα παρουσιάσουμε τα βασικά βήματα του αλγορίθμου για την καλύτερη κατανόηση της διαδικασίας εξαγωγής δεδομένων συστάσεων από το χρήστη. Σε επόμενη παράγραφο θα παρουσιάσουμε αναλυτικά τα επιμέρους βήματα. Τα βήματα του αλγορίθμου παρουσιάζονται παρακάτω:

1. Ανάκτηση των αξιολογήσεων του χρήστη μας καθώς και των αξιολογήσεων των άλλων χρηστών και υπολογισμός του πίνακα ομοιότητας των χρηστών.
2. Εύρεση μη αξιολογημένων καταστημάτων από τον χρήστη μας αλλά αξιολογημένων από άλλους χρήστες.
3. Αναζήτηση και ανάκτηση καταστημάτων που έχουν αξιολογηθεί από τους παρόμοιους χρήστες με βαθμολογία μεγαλύτερη του 6/10. Εξαγωγή συστάσεων.
4. Εμπλουτισμός συστάσεων με την πληροφορία των φίλων του χρήστη.
5. Εξαγωγή συστάσεων σύμφωνα με τα φίλτρα που έχει ορίσει ο χρήστης.

5.1.2. Σύντομη περιγραφή ανατροφοδότησης αλγορίθμου και μεταβολής προφίλ χρήστη

Η αλληλεπίδραση του χρήστη μας με τα καταστήματα μέσω της αξιολόγησης τους επιφέρει αλλαγές στην ομάδα των συστάσεων που εξάγει ο αλγόριθμος. Η κατ'εξασκούθηση θετική βαθμολογία καταστημάτων με ένα συγκεκριμένο τύπο καταστήματος ή μία συγκεκριμένη μουσική επιφέρει αλλαγές στο προφίλ. Η διαδικασία της αλλαγής του προφίλ και των κριτηρίων που απαιτείται θα αναλυθεί σε επόμενη παράγραφο.

5.2 Ανάλυση βημάτων αλγορίθμου συστάσεων στον χρήστη

Σε αυτήν την παράγραφο θα αναλύσουμε την διαδικασία που ακολουθείται για την εξαγωγή των συστάσεων στον χρήστη. Η ανάλυση περιέχει την περιγραφή της διαδικασίας η οποία εμπλουτίζεται με την βοήθεια παράθεσης του πηγαίου κώδικα της ιστοσελίδας. Ο πηγαίος κώδικας συνοδεύεται από εκτεταμένα σχόλια για την εύκολη κατανόηση από τους αναγνώστες που δεν είναι απαραίτητα εξοικειωμένοι με την γλώσσα προγραμματισμού που χρησιμοποιήθηκε(C#).

5.2.1 Ανάκτηση δεδομένων χρηστών και εξαγωγή πίνακα ομοιότητας τους

Σε αυτό το στάδιο της διαδικασίας του αλγορίθμου αναζητούμε τις αξιολογήσεις του χρήστη μας οι οποίες ανακτώνται από την βάση δεδομένων(*Exodus_db*) που έχουμε δημιουργήσει και συγκεκριμένα από τον πίνακα *UserRatings*. Αναζητούμε από αυτόν τον πίνακα τις εγγραφές που έχουν στο πεδίο *UserId* την μοναδική συμβολοσειρά που έχει τεθεί σαν αναγνωριστικό για τον χρήστη μας. Έτσι πλέον έχουμε τις εγγραφές των αξιολογημένων από τον χρήστη μας καταστημάτων, δηλαδή πληροφορίες όπως τα καταστήματα που έχει αξιολογήσει ο χρήστης, το πλήθος των αξιολογήσεων, την ημερομηνία κάθε αξιολόγησης καθώς και την βαθμολογία των αξιολογήσεων. Από όλες αυτές τις διαδικασίες μπορούμε να εξάγουμε και τον μέσο όρο της αξιολόγησης του χρήστη μας για κάθε ένα από αυτά τα καταστήματα, μία πληροφορία που θα μας βοηθήσει-όπως θα δούμε στην συνέχεια- στην μέτρηση και την σύγκριση της ομοιότητας του χρήστη μας με τους υπόλοιπους.

Σε αυτό το σημείο είμαστε πλέον έτοιμοι να αναζητήσουμε τις αξιολογήσεις των άλλων χρηστών για το σύνολο των καταστημάτων που έχουμε εξάγει από την παραπάνω διαδικασία. Αναζητούμε λοιπόν του χρήστες εκείνους οι οποίοι έχουν αξιολογήσει τουλάχιστον τα ίδια καταστήματα με τον χρήστη μας, χωρίς στην παρούσα φάση να μας απασχολεί αν η βαθμολογία για το κάθε κατάστημα ήταν θετική(> 6) ή αρνητική(<=6). Επιλέγουμε να αγνοήσουμε τους χρήστες που δεν έχουν αξιολογήσει έστω και ένα κατάστημα από αυτά που έχει αξιολογήσει ο χρήστης μας για την βελτιστοποίηση του αλγορίθμου στο θέμα της ταχύτητας. Αυτό δεν επηρεάζει τα αποτελέσματα του αλγορίθμου μας διότι στον πίνακα της μέτρησης ομοιότητας των χρηστών η έλλειψη αξιολόγησης επιφέρει αξιολόγηση με βαθμό 0 και αγνοείται από τον μαθηματικό τύπο της μέτρησης ομοιότητας συνημίτονου. Με την παραπάνω διαδικασία έχουμε καταφέρει συγκεντρώσουμε τους υποψήφιους χρήστες για την μέτρηση αξιολόγησης. Στο παρακάτω πίνακα βλέπουμε τα αποτελέσματα τις μέχρι στιγμής εκτέλεσης του αλγορίθμου.

| Χρήστης | Υποψήφιοι χρήστες | | | |
|---------|-------------------|-------|-------|-----|
| MyUser | User1 | User2 | User3 | ... |

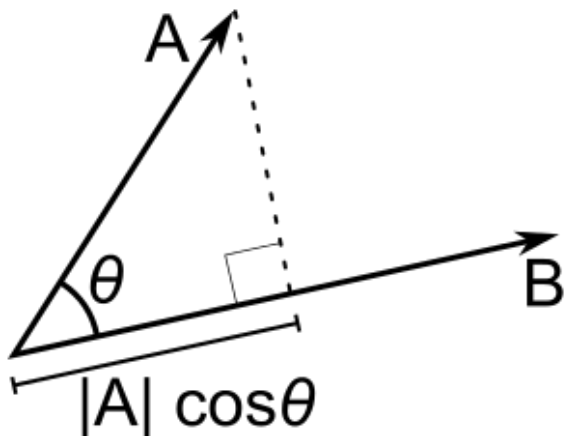
Πίνακας 5-1. Υποψήφιοι χρήστες για μέτρηση ομοιότητας.

Σε αυτό το σημείο έχουμε εντοπίσει τους υποψήφιους χρήστες και πάνω τους θα εφαρμόσουμε τον μαθηματικό τύπο υπολογισμού ομοιότητας συνημίτονου. Ο μαθηματικός τύπος συνημίτονου(Εικόνα 5-2) υπολογίζει την ομοιότητα δύο οντοτήτων βάσει της αριθμητικής τιμής των τιμών-αξιολογήσεων που κάθε χρήστης έχει αξιολογήσει κάθε κατάστημα.

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

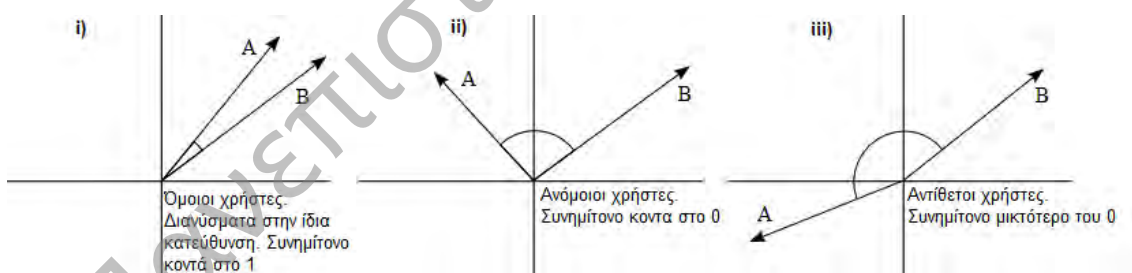
Εικόνα 5-2. Μαθηματικός τύπος ομοιότητας συνημίτονου.

Μπορούμε να αναπαραστήσουμε την ομοιότητα των βαθμολογήσεων δύο χρηστών(A,B) στο χώρο σαν δύο διανύσματα. Προβάλλοντας το διάνυσμα του χρήστη A στον διάνυσμα του χρήστη B και υπολογίζοντας την τιμή του συνημίτονου της συγκεκριμένης γωνίας λαμβάνουμε τον βαθμό ομοιότητας των δύο χρηστών(Εικόνα 5-3).



Εικόνα 5-3. Προβολή διανυσμάτων χρήστη A στο διάνυσμα του χρήστη B.

Όσο η τιμή του συνημίτονου της γωνίας των διανυσμάτων τείνει στην μονάδα τότε τόσο πιο όμοιοι είναι οι εξεταζόμενοι χρήστες. Μπορούμε να ορίσουμε τρεις περιπτώσεις διανυσμάτων από της οποίες προκύπτει η σχέση των δύο χρηστών όπως φαίνεται στην παρακάτω εικόνα 5-4.



Εικόνα 5-4. Περιπτώσεις διανυσμάτων στο χώρο και βαθμός ομοιότητας.

- i. Στην συγκεκριμένη περίπτωση παρατηρούμε ότι τα διανύσματα είναι πιο κοντά στην ταύτιση, κάτι που σημαίνει ότι το συνημίτονο της γωνίας τους πλησιάζει την μονάδα. Οι χρήστες A, B παρουσιάζουν μεγάλη ομοιότητα στις αξιολογήσεις των καταστημάτων και άρα ο υποψήφιος χρήστης B είναι ένας χρήστης που πρέπει να συμπεριληφθεί από τον αλγόριθμο προτάσεων.
- ii. Σε αυτήν την περίπτωση τα διανύσματα απομακρύνονται το ένα από το άλλο καθώς το συνημίτονο τις μεταξύ τους γωνίας πλησιάζει το 0. Ο χρήστης αυτός πρέπει να απορριφθεί από τον αλγόριθμο μας.

- iii. Τέλος σε αυτό το σχήμα η γωνία μεταξύ των διανυσμάτων του χρήστη A, B πλησιάζει κατά πολύ τις 180 μοίρες και άρα πρέπει να απορριφθεί από τον αλγόριθμο. Ο χρήστης B παρουσιάζει μία εντελώς αντίθετη συμπεριφορά από τον χρήστη A.

Στην συγκεκριμένη εργασία τίθεται το ερώτημα του ποιο πρέπει να είναι το κάτω όριο του συνημίτονου ώστε να κατατάξουμε την συμπεριφορά του χρήστη σαν «όμοια» με τον χρήστη μας. Για την επιλογή αυτή πρέπει να λάβουμε υπόψη το πόσο «αυστηρός» θα πρέπει να είναι ο αλγόριθμος μας στην επιλογή υποψηφίων για προτάσεις. Η μέγιστη τιμή που μπορεί να λάβει η ομοιότητα είναι η μονάδα.

Ας δούμε πιθανές τιμές για το κάτω όριο καθώς και την πιθανή συμπεριφορά του αλγορίθμου. Για παράδειγμα αν θέσουμε σαν κάτω όριο την τιμή 0.3 τότε ο αλγόριθμος θα επιστρέψει αρκετά αποτελέσματα ομοιότητας. Είναι αυτό όμως θεμιτό; Σε αυτήν την περίπτωση ο αλγόριθμος είναι ελαστικός ως προς την ομοιότητα και έτσι μπορεί να φτάσουμε σε αποτελέσματα τα οποία μας οδηγούν σε λανθασμένα συμπεράσματα σχετικά με την ομοιότητα του χρήστη μας με τον εξεταζόμενο χρήστη. Άρα επιλέγοντας μια τιμή κάτω ορίου κοντά στο μηδέν κινδυνεύουμε να προτείνουμε στον χρήστη μας καταστήματα τα οποία δεν ταιριάζουν με τις προτιμήσεις του χρήστη μας και άρα τον καθιστούν αναποτελεσματικό.

Ας επιλέξουμε τώρα μία τιμή κοντά στην μονάδα(1). Επιλέγοντας μία τέτοια τιμή ορίο καθιστούν τον αλγόριθμο μας ιδιαίτερα αυστηρό στην επιλογή υποψηφίων χρηστών για προτάσεις. Τα αποτελέσματα αυτής της επιλογής θα επιστρέψουν σαφώς τους πιο «κοντινούς γείτονες» σε θέματα συμπεριφοράς και προτιμήσεων με τον χρήστη μας. Έτσι είναι πιο πιθανό οι συστάσεις του αλγορίθμου να είναι ιδιαίτερα χρήσιμες για τον χρήστη μας, κάτι που καθιστά τον αλγόριθμο μας ιδιαίτερα επιτυχημένο. Τι αρνητικά έχει αυτό για τον αλγόριθμο μας όμως; Τι γίνεται στην περίπτωση που ο αριθμός των χρηστών δεν είναι ιδιαίτερα μεγάλος; Σε αυτήν την περίπτωση οι υποψήφιοι χρήστες για λήψη προτάσεων γίνεται τραγικά μικρός, ιδιαίτερα σε περίπτωση που το κάτω όριο προσεγγίζει την μονάδα και έτσι ο αλγόριθμος δεν θα επιστρέψει παρά ένα πολύ μικρό αριθμό προτάσεων. Αυτό μειώνει κατά πολύ την πιθανότητα να προτιμήσει θετικά ο χρήστης μας κάποιο από τα προταθέντα καταστήματα, όχι λόγω της αβεβαιότητας πρόβλεψης, αλλά λόγω του απειριστικά μικρού αριθμού των προτάσεων.

Λαμβάνοντας υπόψη τα παραπάνω συμπεράσματα επιλέγουμε ένα κάτω όριο της τάξης του 0.6. Η επιλογή μίας συγκρατημένης τιμής η οποία δεν πλησιάζει πάρα πολύ την μονάδα αλλά δεν πέφτει πολύ κάτω από αυτήν καθιστά το όριο αυτό ιδανικό. Ας μην ξεχνάμε ότι μιλάμε για κάτω όριο έτσι οι τιμές του συνημίτονου δεν μπορεί να πέσουν κάτω από αυτό ($0.6 < \text{similarity cosine} \leq 1.0$). Το γεγονός ότι είναι το κάτω όριο δεν εξαιρεί τις περιπτώσεις ομοιότητας που προσεγγίζουν την μονάδα. Κάτι τέτοιο επιτρέπει στα αποτελέσματα να αποτελούνται από εξαιρετικές επιλογές καταστημάτων για τον χρήστη αλλά και σε ένα ικανοποιητικό αριθμό αυτών.

Σε αυτό το σημείο μπορούμε να παραθέσουμε τον κώδικα του αλγορίθμου που επιτρέπει την εξαγωγή των κατάλληλων υποψηφίων χρηστών.

```
//ανάκτηση δεδομένων αξιολογήσεων του χρήστη μας
var curRatings = db.UserRatings.Where(f => f.UserId == curuserId).ToList();
//εύρεση αξιολογήσεων του χρήστη που έχουν συγκεντρώσει βαθμολογία μικρότερη του 6
var curRateLess6 = db.UserRatings.Where(f => f.UserId == curuserId &&
(f.PreviousRateScore != null ? ((f.PreviousRateScore + f.RateScore) / 2) : f.RateScore) < 6);
//δημιουργία προσωρινού πίνακα με τα αναγνωριστικά των καταστημάτων που ο χρήστης
μας έχει αξιολογήσει
var tempcurrateguid = curRatings.Select(ff => ff.StoreId.Value);

// cur users rating matrix
var curSumPow = 0.0;
// υπολογισμός του επάνω σκέλους του τύπου συνημίτονου για τον χρήστη μας
foreach (var rate in curRatings)
{
    var curRateAvg = rate.PreviousRateDate != null ? ((rate.PreviousRateScore +
rate.RateScore) / 2) : rate.RateScore;
    curSumPow += Math.Pow(curRateAvg.Value, 2);
}
```

```

}
var curRoot = Math.Sqrt(curSumPow);

// εύρεση βαθμολογιών των άλλων χρηστών που έχουν αξιολογήσει τα καταστήματα που
// έχει βαθμολογήσει και ο χρήστης μας
var othersRatings = from d in db.UserRatings.Where(f => f.UserId.Value != curuserId &&
tempcurrateguid.Contains(f.StoreId.Value))
                    group d by d.UserId.Value into gr select gr;

// μήτρα ομοιότητας χρηστών (guid-> UserID, double-> similarity value)
List<KeyValuePair<Guid, double>> result1 = new List<KeyValuePair<Guid, double>>();

// calculate similarity with cosine distance formula
foreach (var ouser in othersRatings.ToList())
{
    var upper = 0.0;
    var downotherSum = 0.0;
    //υπολογισμός του πίνακα ομοιότητας
    foreach (var orate in ouser)
    {
        var curUserStoreRate = curRatings.FirstOrDefault(f => f.StoreId ==
orate.StoreId);
        var curRateAvg = curUserStoreRate.PreviousRateDate != null ?
((curUserStoreRate.PreviousRateScore + curUserStoreRate.RateScore) / 2) :
curUserStoreRate.RateScore;
        var curotherAvg = orate.PreviousRateScore != null ?
((orate.PreviousRateScore + orate.RateScore) / 2) : orate.RateScore;
        upper += curotherAvg.Value * curRateAvg.Value;
        downotherSum += Math.Pow(curotherAvg.Value, 2);
    }
    var oRoot = Math.Sqrt(downotherSum);
    var restemp = upper / (curRoot * oRoot);
    //αν ο χρήστης αυτός παρουσιάζει ομοιότητα μεγαλύτερη του κάτω ορίου πρόσθεσε τον
    //στην λίστα με τους χρήστες για λήψη προτάσεων
    if(restemp > 0.6) result1.Add(new
KeyValuePair<Guid, double>(ouser.FirstOrDefault().UserId.Value, restemp));
}

```

5.2.2 Εύρεση μη αξιολογημένων καταστημάτων από τον χρήστη μας.

Σε αυτό το σημείο έχουμε ήδη επιλέξει τους χρήστες που θα χρησιμοποιήσουμε για την εύρεση προτάσεων. Αυτοί οι χρήστες είναι οι περισσότερο όμοιοι με τον χρήστη μας. Εδώ πρέπει να βρούμε τα καταστήματα τα οποία αυτοί οι χρήστες έχουν αξιολογήσει. Η αξιολόγηση πάνω σε αυτά τα καταστήματα θα μας δώσει το σύνολο των καταστημάτων που θα προταθούν στον χρήστη μας. Αυτές οι προτάσεις θα βασιστούν πάνω στις θετικές αξιολογήσεις στα εν λόγω καταστήματα. Εδώ αφαιρούμε τα καταστήματα που ο χρήστης μας έχει ήδη αξιολογήσει. Παρακάτω παραθέτουμε τον κώδικα ο οποίος μας επιτρέπει στην ανάκτηση των υποψήφιων για πρόταση καταστημάτων.

```

// Ανάκτηση των υποψήφιων χρηστών που έχουν επιλεχθεί για τις προτάσεις
var temp = result1.Select(ff => ff.Key).ToList();
// Ανάκτηση των αξιολογήσεων των υποψήφιων χρηστών
var otherSimilar = db.UserRatings.Where(f => temp.Contains(f.UserId.Value));
// Φιλτράρισμα των ήδη αξιολογηθέντων καταστημάτων από τον χρήστη μας
var result2 =otherSimilar.Where(f => !tempcurrateguid.Contains(f.StoreId.Value));

```

5.2.3 Εύρεση μη αξιολογημένων καταστημάτων από τον χρήστη μας.

Στο σημείο αυτό της εκτέλεσης του αλγορίθμου έχουμε ανακτήσει τις βαθμολογίες των παρόμοιων χρηστών. Έτσι έχουμε βαθμολογίες για καταστήματα οι οποίες μπορεί να κυμαίνονται μεταξύ 1-10. Εδώ τίθεται ένα ακόμα ερώτημα σχετικά με το ποιες βαθμολογίες πρέπει να χρησιμοποιήσουμε για να καταλήξουμε στην τελική λίστα με τα προταθέντα καταστήματα. Όπως αναφέραμε παραπάνω οι πιθανές βαθμολογίες μπορούν να κυμανθούν μεταξύ 1 έως και 10. Μπορούμε να χωρίσουμε την βαθμολογία ενός καταστήματος σε τρεις βασικές κατηγορίες όπως βλέπουμε στον παρακάτω πίνακα.

| Βαθμολογία καταστημάτων (εύρος [1,10]) | | |
|--|-------------------|------------------|
| Χαμηλή βαθμολογία | Μέτρια βαθμολογία | Υψηλή βαθμολογία |
| [1,3] | [4,6] | (6,10] |

Πίνακας 5-2. Κατάταξη βαθμολογίας χρηστών

Στον παραπάνω πίνακα χωρίζουμε την βαθμολόγηση σε τρεις κατηγορίες (χαμηλή, μέτρια, υψηλή). Στην χαμηλή βαθμολογία κατατάσσουμε τα καταστήματα τα οποία οι χρήστες έχουν βαθμολογήσει με το πλέον χαμηλό εύρος τιμών. Όπως είναι λογικό τα καταστήματα αυτά δεν προτιμώνται από χρήστες παρόμοιους με τον χρήστη μας. Έτσι επιλέγουμε όπως είναι λογικό να αποκλείσουμε τα καταστήματα αυτά από την λίστα των προτάσεων μας. Στην μέτρια βαθμολογία κατατάσσουμε τα καταστήματα τα οποία έχουν συγκεντρώσει βαθμολογία που κυμαίνεται λίγο κάτω και λίγο πάνω από τον μέσο όρο της βαθμολογικής κλίμακας. Αυτά τα καταστήματα μπορεί να δώσουν καταστήματα τα οποία συγκεντρώνουν πιθανότητες να ενδιαφέρουν τον χρήστη μας, όμως οι πιθανότητες αυτές είναι αρκετά μικρές. Η επιλογή καταστημάτων που ανήκουν σε αυτό το εύρος μπορεί μειώσουν δραστικά την αποτελεσματικότητα του αλγορίθμου μας. Κάτι τέτοιο θα μπορούσε να δώσει στον χρήστη μας την εντύπωση ότι ο αλγόριθμος δεν εξυπηρετεί τις ανάγκες του. Στην τελευταία κλίμακα βαθμολόγησης εντάσσουμε τα καταστήματα τα οποία έχουν συγκεντρώσει μία πολύ καλή έως και άριστη βαθμολογία από τους χρήστες. Αυτά τα καταστήματα είναι τα πλέον κατάλληλα να συμπεριληφθούν από τον αλγόριθμο μας αφού η υψηλή βαθμολογία που συγκεντρώνουν δίνουν πολύ υψηλές πιθανότητες επιλογής τους από τον χρήστη μας.

Βασίζομενοι στην παραπάνω λογική επιλέγουμε τα καταστήματα που έχουν βαθμολογηθεί τουλάχιστον με βαθμολογία μεγαλύτερη του 6, εξασφαλίζοντας ότι τα καταστήματα αυτά είναι τα πιο κατάλληλα για να προταθούν στον χρήστη μας. Ο παρακάτω κώδικας εκτελεί την περιγραφείσα διαδικασία και μας επιστρέφει τα τελικά καταστήματα που πρέπει να προταθούν στον χρήστη.

```
//ανάκτηση αξιολογήσεων με μέσο όρο βαθμολογίας μεγαλύτερο του 6
var result3i = result2.Where(f => (f.PreviousRateScore != null ? ((f.RateScore
+ f.PreviousRateScore) / 2) : f.RateScore) > 6).Select(ff =>
ff.Stores.StoreDetails.FirstOrDefault());

// ανάκτηση συνδετικού πίνακα προτιμήσεων χρήστη
var userPrefsMaster = db.UserPrefAssoc.FirstOrDefault(f => f.UserId ==
curuserId);
// δεδομένα προτιμήσεων χρήστη στην μουσική
var userPrefsMusic = userPrefsMaster.MusicPreferences.Select(f =>
f.MusicId.Value);
// δεδομένα προτιμήσεων χρήστη στο είδος του καταστήματος
var userPrefsCats = userPrefsMaster.StoreTypePreferences.Select(f =>
f.StoreTypeId.Value);

//τελική λίστα προτάσεων
IEnumerable<CollaborationStoresRaw> recommendedStores;
// εμπλουτισμός δεδομένων με στοιχεία για την μουσική και το είδος του
καταστήματος
recommendedStores = (from s in result3i
```

```

join sm in db.StoreMusic on s.StoreId.Value equals sm.StoreId
join m in db.Music on sm.MusicId equals m.MusicId
select new CollaborationStoresRaw
{
    StoreName = s.Stores.StoreName,
    MusicDescr = m.MusicDescription,
    MusicID = m.MusicId,
    CategoryId = s.StoreTypeId.Value,
    CatDescr =
s.StoreTypes.TypeDescription,
    StoreId = s.StoreId.Value,
    StoreDescr = s.Stores.StoreDescription
}).Distinct

```

5.2.4 Εμπλουτισμός προτάσεων με την πληροφορία των φίλων του χρήστη

Σε αυτό το σημείο έχουμε την λίστα με τις προτάσεις που θα παρουσιάσουμε στον χρήστη. Στην εργασία αυτή έχουμε δημιουργήσει ένα περιορισμένο κοινωνικό δίκτυο που το οποίο αποτελείται από χρήστες της ιστοσελίδας exodus. Η επιλογή του χρήστη μας για ένα κατάστημα καθίσταται πιο σίγουρη αν γνωρίζει ότι ένας «φίλος» του έχει αξιολογήσει το κατάστημα θετικά. Για να το επιτύχουμε αυτό πρώτα ανακτούμε τους φίλους του χρήστη μας. Επιλέγουμε τους φίλους που έχουν αξιολογήσει τα καταστήματα που βρίσκονται στην λίστα με τις προτάσεις μας. Στην συνέχεια υπολογίζουμε τον μέσο όρο των αξιολογήσεων του κάθε φίλου μας για τα καταστήματα. Επιλέγουμε τους φίλους που έχουν αξιολογήσει τα καταστήματα με μέσο όρο βαθμολογίας μεγαλύτερο του 6. Τέλος προσθέτουμε την πληροφορία των φίλων στην αρχική λίστα με τις προτάσεις των καταστημάτων. Το αποτέλεσμα από την παραπάνω πληροφορία περιέχει τώρα την πληροφορία των φίλων. Η πληροφορία αυτή εμφανίζεται οπτικά στον χρήστη μας με την βοήθεια ενός εικονιδίου που επιτρέπει σε αυτόν να εντοπίσει με την πρώτη ματιά ποιά καταστήματα από τα προταθέντα είναι αγαπημένα στους φίλους μας.

Το πλήθος των προτάσεων μπορεί να μην είναι αρκετά μεγάλο. Αυτό θα συμβαίνει όταν ο αριθμός των εγγεγραμμένων χρηστών στην ιστοσελίδα δεν είναι αρκετά μεγάλο ή οι προτιμήσεις του χρήστη μας είναι ιδιαίτερες. Στην περίπτωση αυτή τα δεδομένα που έχουμε για την δημιουργία προτάσεων είναι λίγα. Για να επιστρέψουμε ένα αρκετά ικανοποιητικό πλήθος προτάσεων όταν η λίστα των προτάσεων είναι μικρή επιλέγουμε να εμπλουτίσουμε την λίστα των προτάσεων ένα επιπλέον πλήθος καταστημάτων τα οποία ανταποκρίνονται στις προτιμήσεις του.

5.2.5. Εφαρμογή φίλτρων χρήστη

Θα πρέπει να φιλτράρουμε την λίστα των αποτελεσμάτων σύμφωνα με τα φίλτρα αναζήτησης που έχει επιλέξει ο χρήστης μας. Τα φίλτρα αυτά αφορούν το είδος του καταστήματος αλλά και το είδος της μουσικής του καταστήματος. Ο χρήστης μπορεί να επιλέξει παραπάνω από μια κατηγορία καταστημάτων για φίλτρο. Το ίδιο ισχύει και για την μουσική των καταστημάτων. Τα φίλτρα αυτά λειτουργούν συνδυαστικά όπως φαίνεται στο παρακάτω παράδειγμα.

| Μουσική | Κατηγορία |
|----------|-----------|
| M1) Jazz | C1) Bar |
| M2) Pop | C2) Club |

Πίνακας 5-3. Φίλτρα από τα είδη της μουσικής

Τα αποτελέσματα του φίλτρου μπορούν να είναι οποιοσδήποτε από τους παρακάτω συνδυασμούς.

M1-C1, M1-C2, M2-C1, M2, C2.

Ο κώδικας που εκτελεί την λογική των παραγράφων 5.2.4 και 5.2.5 παρατίθεται παρακάτω.

```
// αν το πλήθος των προτάσεων είναι μεγαλύτερο του 0
if (res != null && res.Count() > 0)
{
// εμπλουτισμός αποτελεσμάτων με τα στοιχεία των φίλων του χρήστη
dataResult = res.Where(f =>
filterOptions.SelectedStoreMusics.Contains(f.MusicID.ToString()) ||
filterOptions.SelectedStoreTypes.Contains(f.CategoryId.ToString()))
.ToList()
.Select(f => new AfterColabFilter
{
StoreDescription = f.StoreDescr,
StoreName = f.StoreName,
Id = f.StoreId,
ImageUrl = GetImgUrl2(f.StoreId),
FriendsLiked = friends.Where(s => s.UserRatings.Where(g =>
g.StoreId == f.StoreId && g.RateScore > 6 && s.UserId == g.UserId).Count() ==
1)
.Select(o => new FiendLikeStore { FirstName =
o.FirstName, LastName = o.LastName }).ToList()
}).ToList();
}

// αν το πλήθος των αποτελεσμάτων δεν είναι ικανοποιητικό
if (dataResult.Count == 0 || dataResult.Count < 5)
{
// αναζήτησε επιπλέον δεδομένα σύμφωνα με τα φίλτρα
var dataResultExtra = (from s in exodusDB.StoreDetails
join sm in exodusDB.StoreMusic on
s.StoreId equals sm.StoreId
join m in exodusDB.Music on sm.MusicId
equals m.MusicId
select new
store = s.Stores.StoreName,
music = m.MusicDescription,
musicId = m.MusicId,
categoryId = s.StoreTypeId,
category = s.StoreTypes.TypeDescription,
id = s.StoreId,
description = s.Stores.StoreDescription
}).AsEnumerable().Where(f =>
filterOptions.SelectedStoreMusics.Contains(f.musicId.ToString()) ||
filterOptions.SelectedStoreTypes.Contains(f.categoryId.ToString()))
.Select(f => new AfterColabFilter
{
StoreDescription = f.description,
StoreName = f.store,
Id = f.id.Value,
ImageUrl = GetImgUrl2(f.id),
FriendsLiked = friends.Where(s =>
s.UserRatings.Where(g => g.StoreId == f.id && g.RateScore > 6 && s.UserId ==
g.UserId).Count() == 1)
.Select(o => new FiendLikeStore { FirstName = o.FirstName, LastName =
o.LastName }).ToList()
}).Take(10).ToList();
var ls = dataResultExtra.Where(f => !dataResult.Select(ff =>
ff.Id).Contains(f.Id));
```

```

if (dataResultExtra.Count > 0)
    dataResult.AddRange(1s);
}

```

5.2.6. Ανατροφοδότηση αλγορίθμου και προφίλ χρήστη

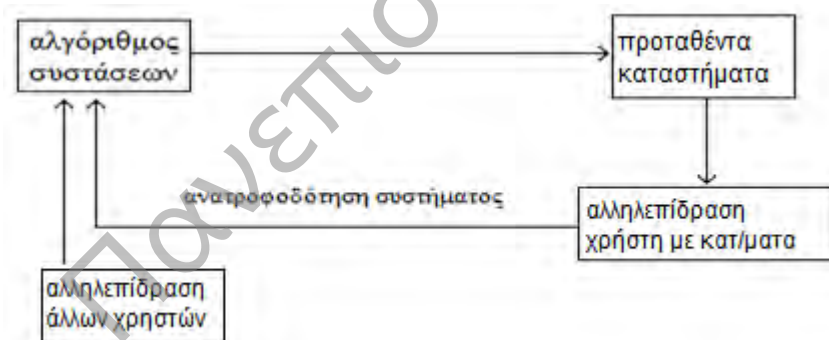
Σε αυτή την παράγραφο θα αναλύσουμε την ανατροφοδότηση του αλγορίθμου και την αλλαγή του προφίλ του χρήστη σαν αποτέλεσμα της αλληλεπίδρασης του με τα καταστήματα. Η αναπροσαρμογή του αλγορίθμου επιτρέπει την διαρκής μεταβολή των προτάσεων στον χρήστη με σκοπό να προσεγγίσει την συμπεριφορά του.

5.2.6.1. Ανατροφοδότηση αλγορίθμου

Ο αλγόριθμος προτάσεων που έχουμε υλοποιήσει χρησιμοποιεί την αλληλεπίδραση των χρηστών με σκοπό να εξάγει τα αποτελέσματα του. Η αλληλεπίδραση αυτή αφορά την αξιολόγηση των χρηστών για καταστήματα τα οποία είτε έχουν επισκεφτεί είτε έχουν σχηματίσει κάποια γνώμη για αυτά. Ο κάθε χρήστης μπορεί να αξιολογήσει τα καταστήματα χρησιμοποιώντας μία φιλική προς αυτόν διεπαφή. Όπως είδαμε και σε προηγούμενη παράγραφο (εικόνα 5-5) η αξιολόγηση αυτή κατατάσσεται σε τρεις κατηγορίες. Στην πρώτη κατηγορία στην οποία ο χρήστης έχει μια μάλλον κακή γνώμη για το κατάστημα, στην δεύτερη κατηγορία στην οποία βαθμολογεί το κατάστημα με μία μέτρια βαθμολογία και τέλος στην τελευταία κατηγορία που ο χρήστης δίνει την καλύτερη βαθμολογία.

Κάθε καινούργια αξιολόγηση από έναν χρήστη δίνει στην βάση γνώσης του αλγορίθμου παραπάνω δεδομένα για να επεξεργαστεί. Τα δεδομένα αυτά εμπλουτίζουν την βάση γνώσης για τα καταστήματα και έτσι μεταβάλλει την συμπεριφορά των προτάσεων που εξάγει. Αυτή η συνεχής μεταβολή της βάσης γνώσης μπορεί να αλλάξει τις προτάσεις προς ένα χρήστη δραστικά.

Η τακτική του αλγορίθμου να αναζητά τους κοντινότερους χρήστες προς τον χρήστη στόχο καταλήγει να αλλάζει κάθε φορά τους παρόμοιους χρήστες προς αυτόν. Αυτό έχει σαν συνέπεια να αλλάζει και να εξελίσσει τις προτάσεις προς τον χρήστη στόχο. Η διαδικασία αυτή φαίνεται καλύτερα στους παρακάτω πίνακες.



Εικόνα 5-5. Ανατροφοδότηση αλγορίθμου

Στον παρακάτω πίνακα βλέπουμε ένα στιγμιότυπο από την κατάταξη του αλγορίθμου και τα αποτελέσματα του αλγορίθμου ως προς την ομοιότητα των άλλων χρηστών με αυτόν. Ας θεωρήσουμε το σενάριο ότι οι χρήστες 3,6 και 5 πραγματοποιούν κάποιες αξιολογήσεις καταστημάτων. Ας θεωρήσουμε και ότι ο χρήστης μας βαθμολογεί τα ίδια καταστήματα με τους παραπάνω χρήστες.

| Χρήστης 1 (χρήστης στόχος) | |
|----------------------------|---------------|
| Παρόμοιοι χρήστες | Ξένοι χρήστες |
| Χρήστης 2 | Χρήστης 3 |
| Χρήστης 5 | Χρήστης 4 |
| Χρήστης 6 | |

Πίνακας 5-4. Στιγμιότυπο εκτέλεσης 1

Εκτελώντας ξανά τον αλγόριθμο για την ίδια ομάδα χρηστών και αναλύοντας τα αποτελέσματα από την σύγκριση ομοιότητας που προκύπτει το στιγμιότυπο που φαίνεται στον παρακάτω πίνακα.

| Χρήστης 1 (χρήστης στόχος) | |
|----------------------------|---------------|
| Παρόμοιοι χρήστες | Ξένοι χρήστες |
| Χρήστης 2 | Χρήστης 4 |
| Χρήστης 6 | Χρήστης 5 |
| Χρήστης 3 | |

Πίνακας 5-5. Στιγμιότυπο εκτέλεσης 2

Αναλύοντας το στιγμιότυπο του πίνακα 3-14 παρατηρούμε κάποιες μεταβολές στους συγγενείς χρήστες. Βλέπουμε ότι ο χρήστης 3 συγκαταλέγεται ανάμεσα στους παρόμοιους με τον χρήστη μας χρήστες. Επίσης βλέπουμε ότι ο χρήστης 5 δεν θεωρείται πλέον παρόμοιος χρήστης. Τέλος βλέπουμε ότι ο χρήστης 6 παραμένει στην λίστα με τους παρόμοιους χρήστες.

Από τα παραπάνω στιγμιότυπα βλέπουμε ότι με κάθε αλληλεπίδραση του χρήστη μας αλλά και των άλλων χρηστών πάνω στην βαθμολογία των καταστημάτων αλλάζει η συμπεριφορά του αλγορίθμου. Αυτό έχει σαν συνέπεια να αλλάξει η τελική λίστα με τους υποψήφιους για λήψη προτάσεων χρήστες. Κάτι τέτοιο αναπροσαρμόζει τα προταθέντα στον χρήστη καταστήματα καθιστώντας την διαδικασία προτάσεων απόλυτα δυναμική. Αν η παραπάνω διαδικασία δεν ήταν τόσο ευαίσθητη στις αλλαγές μέσω βαθμολόγησης καταστημάτων τότε ο χρήστης θα λάμβανε συνεχώς τα ίδια καταστήματα. Κάτι τέτοιο θα έκανε τον χρήστη να χάσει την εμπιστοσύνη του στον αλγόριθμο και κατά συνέπεια ο αλγόριθμος θα έχανε την χρηστικότητα του.

5.2.6.2. Προφίλ χρήστη

Τα συστήματα προτάσεων συνεργατικών φίλτρων εκτός από τα πλεονεκτήματα και την ευελιξία που παρέχουν παρουσιάζουν μία κύρια αδυναμία. Η αδυναμία αυτή ονομάζεται φαινόμενο cold start. Το φαινόμενο αυτό αναλύθηκε στην θεματική ενότητα της θεωρίας που παρουσιάζεται από αυτήν την εργασία. Αξίζει να το αναφέρουμε ξανά για διευκόλυνση του αναγνώστη και καλύτερη κατανόηση αυτής της παραγράφου.

Το φαινόμενο cold start αναφέρεται στην αδυναμία του αλγορίθμου να παρέχει προτάσεις σε έναν χρήστη που μόλις έχει κάνει εγγραφή στο σύστημα ή σε έναν χρήστη που δεν έχει προχωρήσει σε αρκετές αξιολογήσεις πάνω στα αντικείμενα που εφαρμόζεται ο αλγόριθμος, στην συγκεκριμένη περίπτωση στα καταστήματα. Έτσι ο αλγόριθμος δεν έχει επαρκή πληροφορία για να προχωρήσει σε επεξεργασία.

Για την παράκαμψη αυτού του προβλήματος δημιουργήσαμε το προφίλ του κάθε χρήστη. Το προφίλ αυτό είναι στην ουσία ένα στερεότυπο ως προς τα ενδιαφέροντα του χρήστη. Τα ενδιαφέροντα του χρήστη αυτού χωρίζονται σε δύο κατηγορίες, την κατηγορία της μουσικής και την κατηγορία του τύπου του καταστήματος.

Κατά την εγγραφή του χρήστη το σύστημα τον προτρέπει να ορίσει ποιες από τις κατηγορίες της μουσικής και ποια είδη καταστημάτων των ενδιαφέρουν. Ο χρήστης μπορεί να διαλέξει κατ'ελάχιστον δύο είδη μουσικής και δύο είδη καταστημάτων. Ο χρήστης πραγματοποιεί την δημιουργία του προφίλ το μέσο ενός απλού ερωτηματολογίου πολλαπλών επιλογών. Η παρουσίαση του ερωτηματολογίου γίνεται σε παρακάτω ενότητα. Πρέπει να τονίσουμε ότι δεν επιτρέπεται στον χρήστη να προχωρήσει στην πλοήγηση της ιστοσελίδας αν

δεν συμπληρώσει το ερωτηματολόγιο αυτό. Μετά την επιτυχής καταχώρηση του ερωτηματολογίου οι προτιμήσεις του χρήστη αποθηκεύονται από το σύστημα. Ο κώδικας που πραγματοποιεί την δημιουργία του προφίλ του χρήστη φαίνεται παρακάτω.

```
// ανάκτηση χρήστη από την βάση δεδομένων
var curUser = ControllerExtensions.GetCurrentUser(HttpContext);
    if(curUser == null)
        return new JsonResult() { Data = "ERROR", JsonRequestBehavior =
JsonRequestBehavior.AllowGet };

    try
    {
// δημιουργία συσχετιστικού πίνακα προφίλ
        var prefID = Guid.NewGuid();
        var prefAssoc = new UserPrefAssoc()
        {
            Id = prefID,
            UserId = curUser.UserId
        };
        db.UserPrefAssoc.Add(prefAssoc);
// δημιουργία αγαπημένων ειδών μουσικής του προφίλ
        foreach (var musicId in selectedPrefs.SelectedStoreMusics)
        {
            var musipref = new MusicPreferences()
            {
                Id = Guid.NewGuid(),
                MusicId = Guid.Parse(musicId),
                PreferenceId = prefID
            };
            db.MusicPreferences.Add(musipref);
        }
// δημιουργία αγαπημένων ειδών καταστημάτων του προφίλ
        foreach (var stId in selectedPrefs.SelectedStoreTypes)
        {
            var stPref = new StoreTypePreferences()
            {
                Id = Guid.NewGuid(),
                StoreTypeId = Guid.Parse(stId),
                PreferenceId = prefID
            };
            db.StoreTypePreferences.Add(stPref);
        }

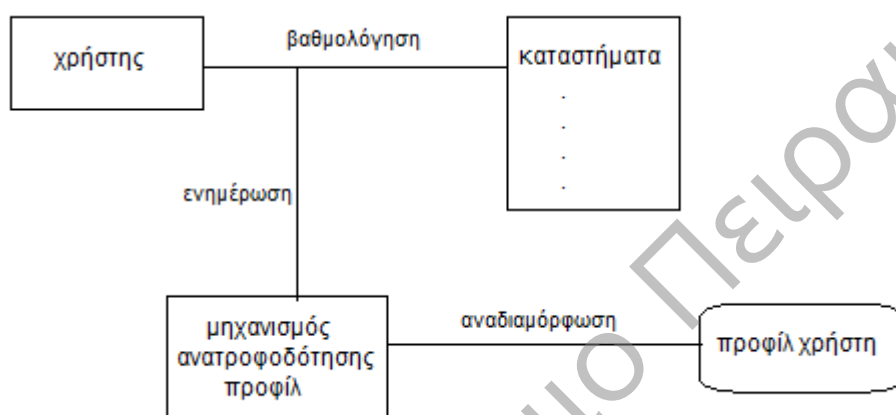
// ενημέρωση βάσης δεδομένων για την επιτυχή καταχώρηση
        var user = db.Users.FirstOrDefault(f => f.UserId ==
curUser.UserId);
        user.PreferencesSet = true;
        curUser.PreferencesSet = true;
        db.SaveChanges();
// ενημέρωση χρήστη για την κατάσταση της δημιουργίας του προφίλ
        return new JsonResult() { Data = "OK", JsonRequestBehavior =
JsonRequestBehavior.AllowGet };
    }
    catch (Exception ex)
    {
        return new JsonResult() { Data = "ERROR", JsonRequestBehavior =
JsonRequestBehavior.AllowGet };
    }
}
```


5.2.6.3. Ενημέρωση προφίλ χρήστη

Όπως είδαμε στην παραπάνω παράγραφο το προφίλ του χρήστη δημιουργείται κατά την εγγραφή του στο σύστημα. Το προφίλ αυτό αποτελεί το στερεότυπο του χρήστη και ενημερώνει το σύστημα για τις προτιμήσεις του χρήστη σχετικά με τα είδη της μουσικής και τα είδη των καταστημάτων που προτιμά.

Το προφίλ αυτό όμως δεν παραμένει στατικό κατά την επίσκεψη και διάδραση του χρήστη στην ιστοσελίδα. Αν συνέβαινε αυτό τότε ο λόγος ύπαρξης του-πέρα των πρώτων επισκέψεων- δεν θα είχε καμία χρησιμότητα. Αντ'αυτού στην εργασία μας έχουμε κατασκευάσει έναν απολύτως δυναμικό τρόπο μεταβολής του προφίλ του χρήστη. Η μεταβολή αυτή πραγματοποιείται και ανατροφοδοτείται μετά από κάθε βαθμολόγηση του χρήστη για κάποιο κατάστημα.

Κάθε φορά που ο χρήστης επιλέγει να βαθμολογήσει κάποιο κατάστημα τότε ενεργοποιείται ο μηχανισμός ανατροφοδότησης του προφίλ όπως φαίνεται στην παρακάτω εικόνα.

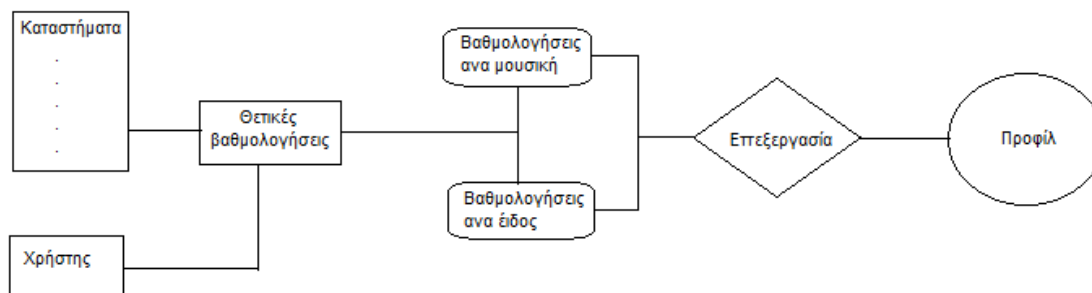


Εικόνα 5-6. Μηχανισμός αναδιαμόρφωσης προφίλ

Ο αλγόριθμος αυτός βασίζεται πάνω στις αξιολογήσεις του χρήστη μας και μόνο. Δεν λαμβάνονται υπόψη ούτε οι βαθμολογίες των παρόμοιων χρηστών ούτε οι βαθμολογίες που έχουν προκύψει μέσω της ομάδας των φίλων του χρήστη μας.

Η λειτουργία του αλγορίθμου περιγράφεται ως εξής. Όταν ο χρήστης μας επισκέπτεται ένα κατάστημα και στην συνέχεια μέσω της διεπαφής της ιστοσελίδας το αξιολογεί τότε ενεργοποιείται ο κώδικας που μπορούμε να δούμε στον παρακάτω πίνακα. Ο κώδικας αυτός εκτελείται αφού ο χρήστης έχει πραγματοποιήσει την βαθμολόγηση δηλαδή η βαθμολογία έχει καταχωρηθεί επιτυχώς στο σύστημα. Σε αυτό το σημείο αναζητούμε τις βαθμολογήσεις του χρήστη μας για τα καταστήματα που έχει αξιολογήσει. Σε περίπτωση που ο χρήστης έχει αξιολογήσει ένα κατάστημα μόνο μία φορά τότε λαμβάνεται η βαθμολογία αυτή σαν η τελική αξιολόγηση του χρήστη. Αν ο χρήστης μας έχει πραγματοποιήσει δύο αξιολογήσεις για το κατάστημα(μέγιστος αριθμός επιτρεπτών αξιολογήσεων) τότε δεν λαμβάνεται η τελευταία αξιολόγηση αλλά ο μέσος όρος των αξιολογήσεων του χρήστη.

Σε αυτό το σημείο επιλέγουμε μόνο τα καταστήματα που ο χρήστης μας έχει αξιολογήσει θετικά. Θυμίζουμε ότι η θετική αξιολόγηση αφορά βαθμολογήσεις της κλίμακας 7 έως 10. Με αυτήν την λογική έχουμε τώρα τα καταστήματα που είναι της αρεσκείας του χρήστη μας χωρίς να έχουμε καταστήματα που ο χρήστης για κάποιο λόγο έχει αξιολογήσει αρνητικά (1 έως 3) ή μέτρια (4 έως 6). Σε αυτό το σημείο πρέπει να πραγματοποιήσουμε μία ομαδοποίηση των βαθμολογήσεων αυτών με βάση τα κριτήρια αξιολόγησης που στην περίπτωση μας είναι η μουσική του καταστήματος και το είδος του καταστήματος. Η ομαδοποίηση αυτή φαίνεται στο παρακάτω διάγραμμα.



Εικόνα 5-7. Διαδικασία αναδιαμόρφωσης προφίλ χρήστη

Τώρα έχουμε την πληροφορία των θετικών βαθμολογιών του χρήστη μας σε ότι αφορά την μουσική και το είδος του καταστήματος. Με κατάλληλη επεξεργασία βρίσκουμε τις θετικές αξιολογήσεις για κάθε είδος μουσικής και για κάθε είδος καταστήματος και εξάγουμε το πλήθος αυτών για κάθε κατηγορία. Για παράδειγμα (Rock -> 2 θετικές αξιολογήσεις, Soul -> 6 θετικές αξιολογήσεις). Σε αυτό το σημείο πρέπει να ανακτήσουμε και τις υπάρχουσες προτιμήσεις του χρήστη όπως έχουν ήδη διαμορφωθεί. Για να ενταχθεί ένας καινούργιο είδος μουσικής ή ένα καινούργιο είδος στο προφίλ του χρήστη πρέπει το είδος αυτό να έχει αξιολογηθεί θετικά τουλάχιστον τρεις φορές από τον χρήστη. Επιλέγουμε τον αριθμό αυτό των ελάχιστων θετικών αξιολογήσεων γιατί ένα πλήθος αξιολογήσεων γιατί θεωρούμε ότι δηλώνει πλέον την συμπεριφορά και τις προτιμήσεις του χρήστη. Η κατ'εξακολούθηση θετική βαθμολόγηση για ένα είδος μουσικής ή κατάστημα δείχνει την προτίμηση του χρήστη για το συγκεκριμένο είδος. Κάτι τέτοιο κάνει τον αλγόριθμο να θεωρήσει ότι το συγκεκριμένο είδος είναι πλέον στις προτιμήσεις του χρήστη.

Η διαρκής ανατροφοδότηση του προφίλ του χρήστη πυροδοτούμενη από τις αξιολογήσεις του καθιστούν τον προφίλ απόλυτα δυναμικό. Το προφίλ, όπως είδαμε δεν είναι ευαίσθητο στον μικρό αριθμό θετικών αξιολογήσεων. Αυτό συμβαίνει γιατί ένας μικρός αριθμός αξιολογήσεων για ένα είδος μπορεί να είναι τυχαίος και έτσι δεν αποτελεί δείγμα τις συμπεριφοράς και των προτιμήσεων του χρήστη. Από την άλλη η επανάληψη της θετικής συμπεριφοράς προς ένα είδος θεωρούμε ότι πλέον αποτελεί μέρος των προτιμήσεων του χρήστη. Παρακάτω παραθέτουμε τον κώδικα της ανατροφοδότησης του αλγορίθμου.

```

// σύνδεση με την βάση δεδομένων
using (var db = new Exodus())
{
// ανάκτηση προτιμήσεων χρήστη από την βάση δεδομένων
var userPrefsAssoc = db.UserPrefAssoc.FirstOrDefault(f => f.UserId == userId);

    if (userPrefsAssoc == null)
        return;

// ανάκτηση προτιμήσεων στα είδη μουσικής
var musicPrefs = userPrefsAssoc.MusicPreferences;
// ανάκτηση προτιμήσεων στα είδη των καταστημάτων
var catPrefs = userPrefsAssoc.StoreTypePreferences;

// ανάκτηση βαθμολογήσεων και υπολογισμός μέσου όρου αυτών ανά κατάστημα
var rateFlat = (from d in db.UserRatings.Where(f => f.UserId == userId).ToList()
                select new
                {
// υπολογισμός μέσου όρου της βαθμολόγησης του καταστήματος
                    rateAvg = GetRateAvg(d),
                    musicTypeId = db.StoreMusic.FirstOrDefault(f => f.StoreId ==
d.StoreId).MusicId,
                    catTypeId = d.Stores.StoreDetails.FirstOrDefault().StoreTypeId
                }).ToList();

```

```

// φίλτράρισμα καταστημάτων με μέσο όρο βαθμολογίας μεγαλύτερη του 6 (7 έως 10)
var musicsRate = rateFlat.Where(f => f.rateAvg > 6).Select(f => f.musicTypeId);
// Ομαδοποίηση βάση των αξιολογήσεων στην μουσική και εύρεση πλήθους
βαθμολογήσεων ανά τύπο μουσικής
var mgr = musicsRate.GroupBy(f => f).Select(o => new
{
    musicID = o.Key,
    Count = o.Count()
});

foreach (var music in mgr)
{
    // Αν ο χρήστης έχει βαθμολογήσει θετικά το είδος μουσικής παραπάνω από 2 φορές
    if (music.Count > 2)
    {
        var s = musicPrefs.Where(f => f.MusicId == music.musicID);
        // Έλεγχος αν η προτίμηση έχει ήδη καταγραφεί
        if (s.Count() == 0)
        {
            // Αν δεν έχει καταγραφεί πρόσθεσε το είδος μουσικής στο προφίλ του χρήστη
            MusicPreferences newmusicPref = new MusicPreferences()
            {
                Id = Guid.NewGuid(),
                PreferenceId = userPrefsAssoc.Id,
                MusicId = music.musicID
            };

            db.MusicPreferences.Add(newmusicPref);
        }
    }
}

// φίλτράρισμα καταστημάτων με μέσο όρο βαθμολογίας μεγαλύτερη του 6 (7 έως 10)
var catRate = rateFlat.Where(f => f.rateAvg > 6).Select(f => f.catTypeId);

// Ομαδοποίηση βάση των αξιολογήσεων στην μουσική και εύρεση πλήθους
βαθμολογήσεων ανά τύπο του καταστήματος
var cgr = catRate.GroupBy(f => f.Value).Select(o => new
{
    catId = o.Key,
    Count = o.Count()
});

foreach (var cat in cgr)
{
    // Αν ο χρήστης έχει βαθμολογήσει θετικά το είδος καταστήματος παραπάνω από 2
    // φορές
    if (cat.Count > 2)
    {
        var c = catPrefs.Where(f => f.StoreTypeId == cat.catId);
        // Έλεγχος αν η προτίμηση έχει ήδη καταγραφεί
        if (c.Count() == 0)
        {
            // Αν δεν έχει καταγραφεί πρόσθεσε το είδος καταστήματος στο προφίλ του χρήστη
            StoreTypePreferences newcatPref = new StoreTypePreferences()
            {
                Id = Guid.NewGuid(),
                PreferenceId = userPrefsAssoc.Id,
                StoreTypeId = cat.catId
            };
        }
    }
}

```

```
        db.StoreTypePreferences.Add(newcatPref);
    }
}
// Αποθήκευση αλλαγής προφίλ στην βάση δεδομένων
db.SaveChanges();
}
```

6. Ανάλυση και περιγραφή της ιστοσελίδας exodus

6.1. Περιγραφή σελίδων

Παρακάτω θα αναλύσουμε τα κυριότερα μέρη της ιστοσελίδας Exodus με την βοήθεια εικόνων και την ανάλυση των κυριότερων λειτουργιών της. Θα αναφερθούμε στις διαδικασίες διάδρασης του χρήστη με την ιστοσελίδα και τα επιμέρους στοιχεία που την απαρτίζουν. Θα αναφέρουμε επίσης στα στοιχεία που αφορούν την ενημέρωση του χρήστη(notifications).

Η είσοδος στην ιστοσελίδα Exodus επιτρέπεται μόνο στους εγγεγραμμένους χρήστες σε αυτήν. Η ιστοσελίδα Exodus απαρτίζεται από τις εξής επιμέρους σελίδες:

- *Register*. Δημιουργία νέου χρήστη.
- *Login*. Σύνδεση χρήστη στην ιστοσελίδα.
- *Index*. Αρχική σελίδα Exodus.
- *Stores*. Σελίδα αναζήτησης καταστημάτων.
- *Special*. Σελίδα προτάσεων καταστημάτων στον χρήστη.
- *StoreDetails*. Σελίδα αναλυτικών πληροφοριών καταστήματος.
- *UserProfile*. Σελίδα πληροφοριών φίλων χρήστη.
- *ManageAccount*. Σελίδα τροποποίησης στοιχείων χρήστη.

6.1.1. Register, εγγραφή χρήστη

Όπως είδαμε στην εισαγωγή του παρόντος κεφαλαίου, ο χρήστης μπορεί να επισκεφτεί και να πλοηγηθεί στην ιστοσελίδα Exodus μόνο αν έχει εισαχθεί επιτυχώς στο σύστημα. Για αυτό τον σκοπό έχουμε δημιουργήσει μία σελίδα για την εισαγωγή των στοιχείων του στο σύστημα και την καταχώρησή τους σε αυτό. Ο χρήστης πρέπει να εισάγει μια σειρά από προσωπικά στοιχεία στα ειδικά πεδία που παρέχονται για τον σκοπό αυτό. Τα πεδία αυτά είναι τα εξής:

- *Username*. Ονομασία χρήστη.
- *Password*. Κωδικός χρήστη.
- *Confirm Password*. Επιβεβαίωση κωδικού.
- *First Name*. Όνομα χρήστη.
- *Last Name*. Επώνυμο χρήστη.
- *Nickname*. Φιλικό όνομα χρήστη.
- *Email*. Διεύθυνση ηλεκτρονικού ταχυδρομείου χρήστη.
- *Telephone*. Τηλέφωνο χρήστη.
- *Mobile*. Κινητό χρήστη.

Στην φόρμα αυτή υπάρχουν πεδία τα οποία είναι υποχρεωτικά και ο χρήστης πρέπει οπωσδήποτε να συμπληρώσει και πεδία τα οποία είναι προαιρετικά. Τα υποχρεωτικά πεδία είναι *Username*, *Password*, *Confirm Password*, *First Name*, *Last Name* και *Email*. Τα προαιρετικά πεδία είναι *Nickname*, *Telephone* και *Mobile*. Στην παρακάτω εικόνα φαίνεται η σελίδα της εγγραφής χρήστη.

The screenshot shows a registration form titled "Registration Form" for "exodus Athens guide". The form contains several input fields: Username, Password, Confirm password, First Name, Last Name, Nickname, Email, Telephone, and Mobile. A blue "Register" button is located at the bottom right of the form. On the right side of the page, there is a vertical list of red error messages, each preceded by a shield icon with an exclamation mark. The messages are: "The Email field is required.", "The Last Name field is required.", "The First Name field is required.", "The Password field is required.", and "The Username field is required." Additionally, red error messages are placed directly below the corresponding input fields: "The Username field is required." below the Username field, "The Password field is required." below the Password field, "The First Name field is required." below the First Name field, "The Last Name field is required." below the Last Name field, and "The Email field is required." below the Email field. The footer of the page contains the text "Copyright © Evangelos Tampakis All Rights Reserved" and "design by Me".

Εικόνα 6-1. Εγγραφή χρήστη και validation

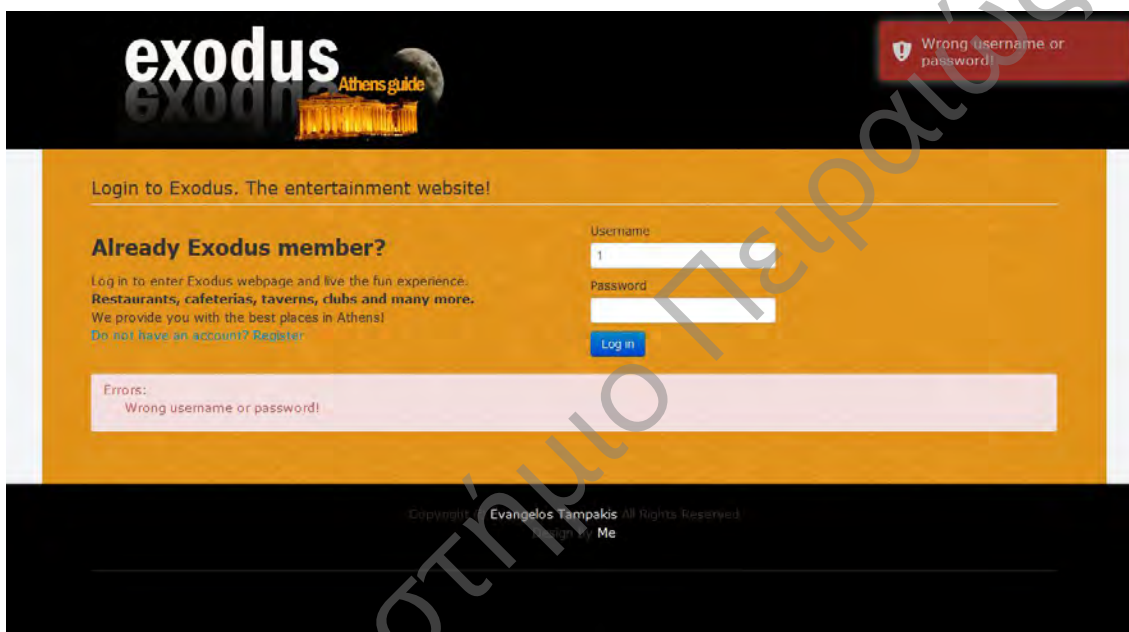
Στην εικόνα 6-1 που απεικονίζει την σελίδα της εγγραφής χρήστη βλέπουμε τα πεδία που ο χρήστης πρέπει να συμπληρώσει. Σε αυτήν την περίπτωση ο χρήστης δεν έχει συμπληρώσει κάποιο από τα πεδία. Έχουμε εφαρμόσει έναν μηχανισμό για την ενημέρωση του χρήστη για αυτήν την περίπτωση. Βλέπουμε τις ειδοποιήσεις πάνω δεξιά στην οθόνη οι οποίες με την βοήθεια μηνυμάτων ειδοποιούν τον χρήστη. Στην συγκεκριμένη περίπτωση βλέπουμε το μήνυμα "The Email field is required". Αυτό το μήνυμα ενημερώνει ότι το πεδίο Email είναι υποχρεωτικό και ότι ο χρήστης πρέπει να το συμπληρώσει για να ολοκληρώσει στην εγγραφή του. Παρόμοια μηνύματα βλέπουμε για κάθε υποχρεωτικό πεδίο που αναφέραμε παραπάνω.

Εκτός από τις ειδοποιήσεις για τα υποχρεωτικά πεδία υπάρχει και ένα πλήθος από άλλες ενημερώσεις. Αυτές αφορούν την εγκυρότητα του email, του password, του username και του πεδίου telephone. Ο μηχανισμός εγκυρότητας ειδοποιεί τον χρήστη για λάθος μορφή στο email που έχει εισάγει, έτσι για παράδειγμα αν ο χρήστης στο πεδίο email εισάγει μία τιμή της μορφής faultemai@l.gr το σύστημα θα τον ειδοποιήσει ότι το email που έχει εισάγει δεν είναι έγκυρο και άρα ο χρήστης πρέπει να εισάγει σωστά το email του. Στην περίπτωση το πεδίου telephone το σύστημα ειδοποιεί τον χρήστη για λάθος μορφή τηλεφώνου. Έτσι αν ο χρήστης εισάγει ένα τηλέφωνο της μορφής "210fd2254" το σύστημα θα ενημερώσει τον χρήστη για λανθασμένο τηλέφωνο. Στην περίπτωση του password το σύστημα δεν απαγορεύει από τον χρήστη την εισαγωγή οποιασδήποτε συμβολοσειράς. Όμως στο πεδίο confirm password ζητείται από τον χρήστη η επιβεβαίωση του password. Σε περίπτωση που τα πεδία password και confirm password δεν ταυτίζονται, προβάλλεται στον χρήστη το ανάλογο μήνυμα. Τέλος όσον αφορά το πεδίο username ο χρήστης μπορεί να θέσει όποια τιμή επιθυμεί, όμως και εδώ υπάρχει κάποιος περιορισμός. Όταν ο χρήστης ζητήσει την εγγραφή του με ένα username τότε το σύστημα ελέγχει αν το username είναι δεσμευμένο από άλλον χρήστη τότε ο χρήστης ενημερώνεται ότι πρέπει να διαλέξει κάποιο άλλο username.

Όταν ο χρήστης συμπληρώσει τουλάχιστον τα υποχρεωτικά πεδία σύμφωνα με τους περιορισμούς που είδαμε τότε μεταφέρεται στην σελίδα της σύνδεσης χρήστη όπως θα δούμε στην επόμενη παράγραφο.

6.1.2. Login, σύνδεση χρήστη στην ιστοσελίδα

Στην σελίδα login ο χρήστης μπορεί να συνδεθεί στην ιστοσελίδα Exodus. Για την σύνδεση του, απαιτούνται δύο πεδία, το πεδίο username και το πεδίο password. Τα πεδία αυτά είναι υποχρεωτικά. Σε περίπτωση που τα πεδία αυτά δεν συμπληρωθούν τότε το σύστημα τον ενημερώνει με μία ειδοποίηση υποχρέωση συμπλήρωσης πεδίου. Όταν ο χρήστης συμπληρώσει τα πεδία αυτά και επιλέξει το κουμπί Log in τότε το σύστημα ελέγχει την εγκυρότητα τους. Το σύστημα αναζητά τον χρήστη στην βάση δεδομένων και αν τα στοιχεία αυτά είναι έγκυρα τότε μεταφέρεται στην σελίδα Index που αποτελεί την αρχική σελίδα του ιστοτόπου Exodus. Στην περίπτωση που τα στοιχεία αυτά δεν ανήκουν σε κάποιον χρήστη τότε το σύστημα ενημερώνει για λάθος username και password. Οι ειδοποιήσεις εδώ είναι δύο, η μία είναι η ειδοποίηση πάνω δεξιά στην οθόνη αναφέροντας το λάθος, η άλλη ειδοποίηση εμφανίζεται στην περιοχή κάτω από την φόρμα των πεδίων. Η σελίδα login φαίνεται στην παρακάτω εικόνα 6.2 όπου βλέπουμε τα πεδία που πρέπει να συμπληρώσει ο χρήστης αλλά και τις ειδοποιήσεις λανθασμένης εισαγωγής στοιχείων.



Εικόνα 6-2. Σελίδα σύνδεσης χρήστη

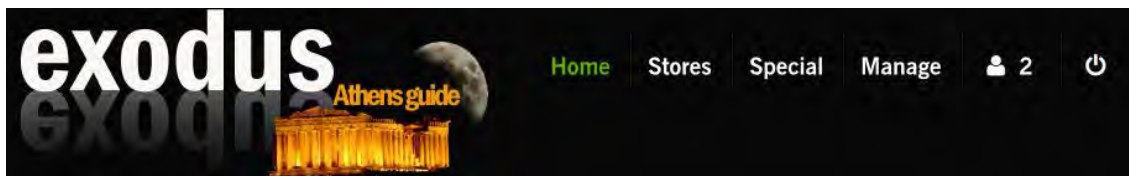
6.1.3. Index, αρχική σελίδα Exodus

Η σελίδα index αποτελεί την αρχική σελίδα του ιστοτόπου Exodus. Εδώ μεταφέρεται ο χρήστης μετά την επιτυχή σύνδεση και ταυτοποίηση του χρήστη από την σελίδα login που εξετάσαμε παραπάνω. Η σελίδα index φαίνεται στην παρακάτω εικόνα 6-3. Σε αυτήν την παράγραφο θα αναλύσουμε τα επιμέρους στοιχεία που αποτελούν την σελίδα αυτήν.



Εικόνα 6-3. Αρχική σελίδα Exodus

Αρχίζοντας την ανάλυση της ιστοσελίδας βλέπουμε την κεφαλίδα(header) της ιστοσελίδας μας. Η κεφαλίδα αυτή συναντάτε σε όλες τις σελίδες του ιστοτόπου Exodus για αυτό τον λόγο θα την αναλύσουμε μόνο σε αυτήν την παράγραφο. Η κεφαλίδα αποτελείται από το λογότυπο της ιστοσελίδας η οποία είναι σύνδεσμος που οδηγεί να πάσα στιγμή στην αρχική σελίδα index. Στο δεξί μέρος του λογότυπου ακολουθεί η μπάρα πλοήγησης.



Εικόνα 6-4. Κεφαλίδα Exodus

Η μπάρα πλοήγησης αποτελείται από μία σειρά συνδέσμων που βοηθούν τον χρήστη στην εύκολη μετάβαση του χρήστη από μία σελίδα της ιστοσελίδας σε μία άλλη. Εδώ έχουμε τις παρακάτω επιλογές.

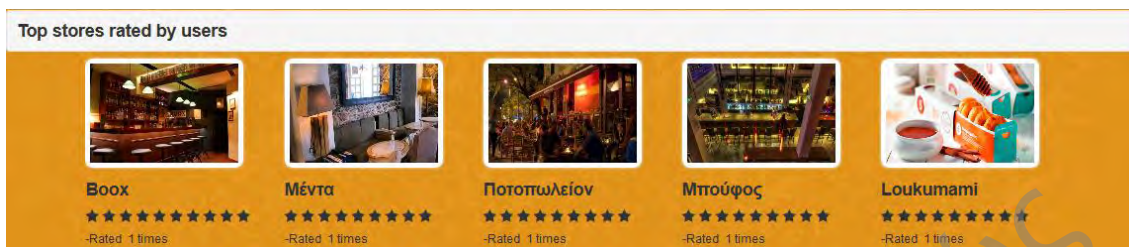
- Home. Με την επιλογή αυτού του συνδέσμου ο χρήστης μεταφέρεται από οποιοδήποτε σημείο της ιστοσελίδας, στην αρχική σελίδα Index.
- Stores. Επιλέγοντας ο χρήστης τον αυτόν τον σύνδεσμο πλοηγείτε στην σελίδα αναζήτησης καταστημάτων.
- Special. Εδώ ο χρήστης μεταφέρεται στην σελίδα των προτάσεων του αλγορίθμου που έχουμε υλοποιήσει.
- Manage. Η μεταφορά του χρήστη στην σελίδα τροποποίησης του λογαριασμού του γίνεται με την επιλογή αυτού του συνδέσμου
- Όνομα συνδεδεμένου χρήστη. Εδώ ο χρήστης μπορεί να δει το όνομα χρήστη με το οποίο έχει συνδεθεί στην ιστοσελίδα. Επιλέγοντας αυτόν τον σύνδεσμο μεταφέρεται στην σελίδα τροποποίησης του λογαριασμού του.
- Εικονίδιο On/Off. Επιλέγοντας ο χρήστης αυτό το εικονίδιο μεταφέρεται στην σελίδα login. Μετά την επιλογή αυτού του συνδέσμου ο χρήστης δεν είναι πια συνδεδεμένος στο σύστημα και για την επανεισαγωγή του απαιτείται η δια διαδικασία σύνδεσης όπως την είδαμε σε προηγούμενη παράγραφο .

Αμέσως μετά ακολουθεί η παρουσίαση των κατηγοριών των καταστημάτων μέσω εναλλασσόμενων εικόνων. Κατά την παραμονή του χρήστη στην σελίδα αυτή οι εικόνες των κατηγοριών των καταστημάτων εναλλάσσονται και ο χρήστης μπορεί να πλοηγηθεί στα καταστήματα της κατηγορίας που επέλεξε μέσω συνδέσμων που βρίσκονται σε αυτές τις εικόνες. Το κομμάτι αυτό παρουσιάζεται στην εικόνα 6-5.



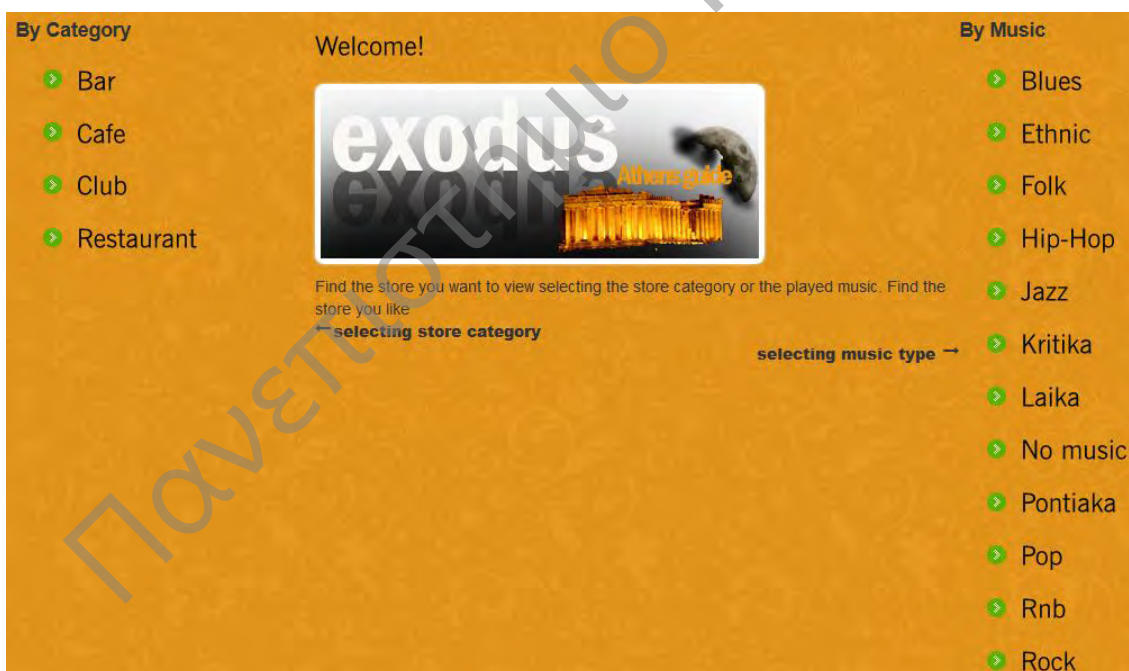
Εικόνα 6-5. Παρουσίαση κατηγοριών καταστημάτων

Παρακάτω στην αρχική σελίδα βλέπουμε το πεδίο με τα πέντε υψηλότερα βαθμολογημένα καταστήματα Εικόνα 6-6. Τα καταστήματα αυτά έχουν βαθμολογηθεί με τον μεγαλύτερο μέσο όρο αξιολόγησης από όλους τους χρήστες της ιστοσελίδας. Αυτά τα καταστήματα δεν είναι αποτέλεσμα προτάσεων του αλγόριθμου μας. Η παρουσίαση τους στον χρήστη μας αποσκοπούν στο να του δείξουν την τάση των χρηστών απέναντι στα καταστήματα. Τα καταστήματα αυτά κατατάσσονται ανάμεσα στα δημοφιλέστερα και πιο αγαπημένα των χρηστών της ιστοσελίδας.



Εικόνα 6-6. Παρουσίαση υψηλότερα βαθμολογημένων καταστημάτων

Ο χρήστης εδώ βλέπει την εικόνα του καταστήματος ακολουθούμενο από το όνομα του. Λίγο πιο κάτω φαίνεται η βαθμολογία του καταστήματος με την μορφή αστεριών. Ο μέγιστος αριθμός αστεριών που μπορεί να συγκεντρώσει ένα κατάστημα είναι δέκα και η χαμηλότερη 1. Στην συνέχεια παρουσιάζουμε τον αριθμό των αξιολογήσεων που έχουν λάβει. Η σειρά της εμφάνισης των καταστημάτων δεν είναι τυχαία αλλά είναι διατεταγμένα από την μεγαλύτερη βαθμολογία στην χαμηλότερη. Ο χρήστης επιλέγοντας την εικόνα ή το όνομα του καταστήματος, μεταφέρεται στην σελίδα με τις λεπτομέρειες του καταστήματος.

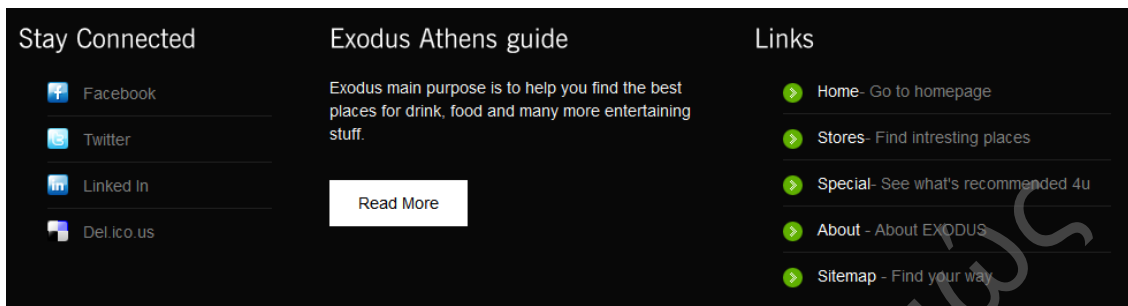


Εικόνα 6-7. Παρουσίαση ειδών μουσική και ειδών καταστημάτων

Κατεβαίνοντας στην ιστοσελίδα παρατηρούμε ένα κείμενο καλωσορίσματος στην ιστοσελίδα Εικόνα 6-7. Εκατέρωθεν αυτού βλέπουμε παρατεταγμένα σε λίστα μία σειρά από επιλογές. Οι επιλογές αυτές δείχνουν στον χρήστη με μια πρώτη ματιά τις επιλογές των καταστημάτων που είναι διαθέσιμες στον χρήστη. Στα αριστερά παρατίθεται η λίστα με τις κατηγορίες των καταστημάτων που φιλοξενεί η ιστοσελίδα. Στα δεξιά παρατίθενται τα είδη μουσικής που απαντώνται από τα καταστήματα που φιλοξενούνται. Ο χρήστης με την επιλογή είτε μίας κατηγορίας είτε ενός είδους μουσική μπορεί να πλοηγηθεί στην σελίδα αναζήτησης

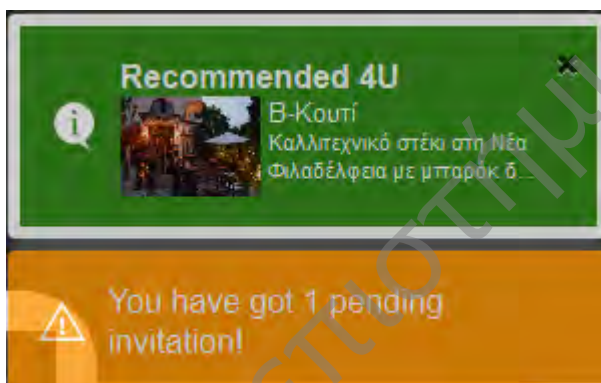
καταστημάτων βλέποντας τα καταστήματα που ανταποκρίνονται στην κατηγορία που έχει επιλέξει από τις δύο αυτές λίστες.

Στο κάτω μέρος της ιστοσελίδας βλέπουμε το υποσέλιδο(footer) όπως φαίνεται στην παρακάτω εικόνα 6-8. Εδώ βλέπουμε μία σειρά από συνδέσμους που βοηθούν τον χρήστη στην πλοήγηση της σελίδας. Επίσης βλέπουμε συνδέσμους για τα κυριότερα μέσα κοινωνικής δικτύωσης όπου ο χρήστης μπορεί να συνδεθεί. Το υποσέλιδο είναι ορατό από όλες τις σελίδες της ιστοσελίδας Exodus όπως έχουμε αναφέρει και για την κεφαλίδα.



Εικόνα 6-8. Υποσέλιδο Exodus

Στην αρχική σελίδα προσφέρουμε στον χρήστη ένα σύνολο ειδοποιήσεων ενημέρωσης όπως φαίνεται στην παρακάτω εικόνα 6-9. Εδώ βλέπουμε δύο ειδών ειδοποιήσεις, τις ειδοποιήσεις προτεινόμενων καταστημάτων και τις ειδοποιήσεις προσκλήσεων φιλίας από άλλους χρήστες.



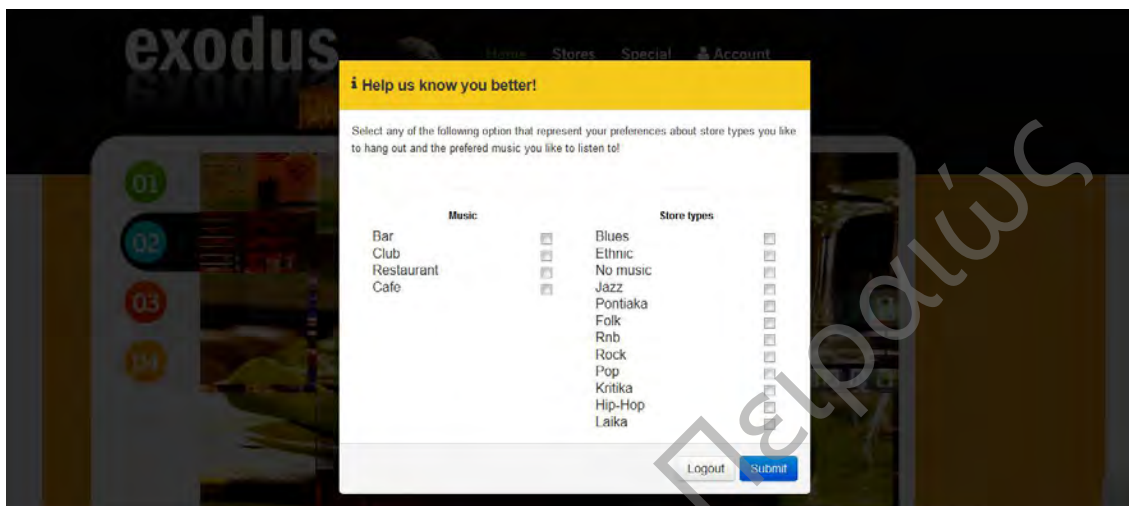
Εικόνα 6-9. Ειδοποιήσεις ενημέρωσης χρήστη

Ο αλγόριθμος προτάσεων παρέχει στον χρήστη μία σειρά από προτεινόμενα καταστήματα όπως έχουμε ήδη αναλύσει σε προηγούμενη παράγραφο. Επιλέγουμε ένα από αυτά τα καταστήματα για να προτείνουμε στον χρήστη ο οποίος ειδοποιείται από ένα ευδιάκριτο κουτί ειδοποίησης. Εδώ φαίνεται το ο τίτλος της ειδοποίησης η οποία ενημερώνει ότι πρόκειται για ένα κατάστημα που προτείνεται αποκλειστικά στον χρήστη και είναι πιθανόν να τον ενδιαφέρει. Το κατάστημα αυτό δεν έχει τύχει βαθμολόγησης του χρήστη στο παρελθόν και για αυτό το λόγο εντάσσεται στα προτεινόμενα καταστήματα. Στην ειδοποίηση περιλαμβάνεται η εικόνα του καταστήματος ακολουθούμενη από το όνομα του και μία σύντομη περιγραφή. Επιλέγοντας ο χρήστης την ειδοποίηση αυτή μεταφέρεται στην σελίδα λεπτομερειών του καταστήματος στην οποία ο χρήστης μπορεί να δει περισσότερες πληροφορίες για αυτό αλλά και να το αξιολογήσει.

Σε περίπτωση που κάποιος χρήστης έχει αποστείλει ένα αίτημα φιλίας στον χρήστη μας εμφανίζεται το μήνυμα ειδοποίησης προσκλήσεων. Με αυτήν την ειδοποίηση ο χρήστης μας ενημερώνεται ότι εκκρεμεί ένα πλήθος προσκλήσεων φιλίας στις οποίες ο πρέπει να απαντήσει είτε θετικά, είτε αρνητικά. Επιλέγοντας ο χρήστης την ειδοποίηση αυτήν μεταφέρεται στην

σελίδα προτάσεων καταστημάτων(Special) στην οποία ο χρήστης μας μπορεί να δει τους φίλους του και να απαντήσει στα αιτήματα φιλίας που έχει δεχθεί. Περισσότερα θα δούμε σε επόμενη παράγραφο στην οποία θα αναλύσουμε την συμπεριφορά και δημιουργία του κοινωνικού δικτύου του χρήστη.

Σε αυτό το σημείο έχουμε αναλύσει το μεγαλύτερο μέρος της αρχικής σελίδας Exodus. Πότε δημιουργείται το προφίλ του χρήστη; Το προφίλ του χρήστη το οποίο έχουμε αναλύσει σε προηγούμενο κεφάλαιο δημιουργείται όταν ο χρήστης εισέρχεται στην ιστοσελίδα για πρώτη φορά. Στην περίπτωση αυτή εμφανίζεται ένα παράθυρο επικοινωνίας με τον χρήστη το οποίο φαίνεται στην παρακάτω εικόνα 6-10.



Εικόνα 6-10. Δημιουργία προφίλ χρήστη

Σε αυτό το παράθυρο διαλόγου ο χρήστης καλείται να επιλέξει ανάμεσα στις κατηγορίες καταστημάτων και στα είδη μουσικής που φιλοξενούνται στην ιστοσελίδα μας. Παρέχονται σε αυτόν δύο λίστες με τις παραπάνω επιλογές και ο χρήστης μπορεί να τις επιλέξει μέσω των κουτιών επιλογής(checkboxes) που υπάρχουν δίπλα από κάθε κατηγορία. Στο τέλος του παραθύρου αυτού δίνονται στον χρήστη η επιλογή δημιουργίας του προφίλ και η επιλογή της αποσύνδεσης από την σελίδα. Ο χρήστης δεν μπορεί να παρακάμψει την δημιουργία του προφίλ καθώς αποτελεί βασική μονάδα του αλγορίθμου προτάσεων. Ο χρήστης θα πρέπει να επιλέξει τουλάχιστον δύο κατηγορίες καταστημάτων και δύο είδη μουσικής. Σε περίπτωση που ο χρήστης δεν επιλέξει τον ελάχιστο αριθμό κατηγοριών ειδοποιείται από το σύστημα. Όταν ο χρήστης επιτυχώς καταχωρήσει το προφίλ του είναι ελεύθερος να συνεχίσει την περιήγηση του στην ιστοσελίδα.

6.1.4 Stores. Σελίδα αναζήτησης καταστημάτων

Στην σελίδα αναζήτησης καταστημάτων Stores ο χρήστης έχει την δυνατότητα να αναζητήσει τα καταστήματα που φιλοξενούνται στην ιστοσελίδα μας. Η αναζήτηση γίνεται με την βοήθεια φίλτρων όπως φαίνονται στην παρακάτω εικόνα της σελίδας Εικόνα 6-11. Ο χρήστης μπορεί να επιλέξει ανάμεσα σε φίλτρα τα οποία αφορούν το είδος των καταστημάτων αλλά και το είδος της μουσικής οι οποίες απαντώνται σε αυτά. Η επιλογή των φίλτρων μπορεί να είναι πολλαπλή, δηλαδή ο χρήστης να επιλέξει παραπάνω από δύο είδη μουσικής και κατηγορίες καταστημάτων. Τα φίλτρα αυτά λειτουργούν συνδυαστικά. Για παράδειγμα αν ο χρήστης επιλέξει το είδος καταστήματος Bar και επιλέξει σαν φίλτρα μουσικής Jazz και Rock τότε τα αποτελέσματα που θα προκύψουν μπορεί να είναι του συνδυασμού Bar-Jazz ή Bar-Rock.

Τα καταστήματα που θα επιστρέψει η αναζήτηση παρουσιάζονται στο κεντρικό τμήμα με την μορφή grid. Κάθε κατάστημα εμφανίζεται με την εικόνα του καταστήματος ακολουθούμενοι από το όνομα του. Παρακάτω εμφανίζεται μια σύντομη περιγραφή του καταστήματος καθώς και ένα κουμπί το οποίο οδηγεί στην σελίδα της περιγραφής του καταστήματος.



Εικόνα 6-11. Σελίδα αναζήτησης καταστημάτων

6.1.5. Special. Σελίδα προτάσεων καταστημάτων στον χρήστη

Στην σελίδα Special ο χρήστης μπορεί να δει τα προτεινόμενα προς αυτόν καταστήματα Εικόνα 6-12. Εισερχόμενος στην σελίδα αυτή ο αλγόριθμος προτάσεων ενεργοποιείται και σύμφωνα με τα όσα είδαμε στην παράγραφο ανάλυσης του, επιστρέφει μία ομάδα καταστημάτων που επιλέγονται αποκλειστικά για τον χρήστη.



Εικόνα 6-12. Σελίδα προτάσεων στον χρήστη

Στην αριστερή πλευρά της σελίδας βλέπουμε τα φίλτρα αναζήτησης των καταστημάτων. Αυτή την φορά τα φίλτρα αυτά είναι συμπληρωμένα για λογαριασμό του χρήστη αφού όπως βλέπουμε είναι επιλεγμένα ορισμένα είδη μουσικής και είδη καταστημάτων. Η συμπλήρωση των φίλτρων προκύπτει από το δυναμικό προφίλ του χρήστη. Στην παραπάνω εικόνα βλέπουμε πως έχουν συμπληρωθεί τα φίλτρα μουσικής Blues, Folk και τα φίλτρα κατηγορίας καταστημάτων Café, Restaurant. Αυτό σημαίνει ότι το προφίλ που έχει δημιουργήσει το σύστημα για τον χρήστη μας έχει καταλήξει ότι ο χρήστης μας έχει αρέσκεια σε αυτά τα είδη μουσικής και κατηγορίες καταστημάτων. Το ότι τα φίλτρα αυτά είναι ήδη συμπληρωμένα δεν αφαιρούν από τον χρήστη να ψάξει για προτεινόμενα καταστήματα με βάσει άλλα φίλτρα. Ο αλγόριθμος θα αναζητήσει για προτεινόμενα καταστήματα σύμφωνα με τα νέα κριτήρια. Σε περίπτωση που ο αλγόριθμος δεν φέρει αποτελέσματα ή τα αποτελέσματα αυτά είναι πολύ λίγα τότε το σύστημα εμπλουτίζει τα αποτελέσματα με καταστήματα τα οποία πληρούν τα κριτήρια που έχει θέσει μέσω των φίλτρων.

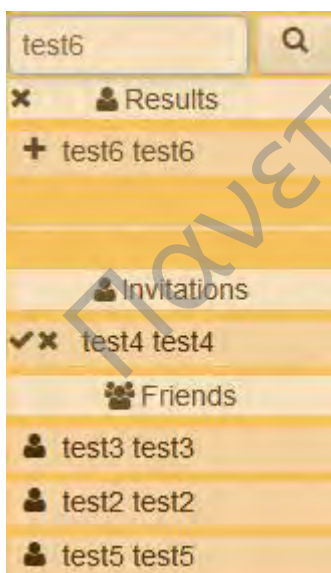
Στην παρακάτω εικόνα 6-13 το προτεινόμενο κατάστημα παρουσιάζεται όπως έχει αναλυθεί και στην προηγούμενη παράγραφο που αναφέρεται στην σελίδα Stores. Παρατηρούμε όμως ότι στο πάνω δεξιά μέρος της εικόνας του καταστήματος υπάρχει μια ειδική σήμανση. Η

σήμανση αυτή δείχνει ότι το συγκεκριμένο κατάστημα έχει χαρακτηριστεί ως αγαπημένο από έναν ή περισσότερους φίλους του χρήστη. Αυτή η σήμανση μπορεί να προτρέψει τον χρήστη μας να επισκεφτεί το συγκεκριμένο κατάστημα και να το αξιολογήσει.



Εικόνα 6-13. Προτεινόμενο κατάστημα το οποίο αρέσει σε φίλους

Στα δεξιά της σελίδας βλέπουμε μία ειδική φόρμα που μας δείχνει τους φίλους του χρήστη μας στο κοινωνικό δίκτυο της σελίδας μας Εικόνα 6-14. Στο πάνω μέρος της φόρμας ο χρήστης μπορεί να αναζητήσει άλλους εγγεγραμμένους χρήστες στην ιστοσελίδα μας. Αυτό γίνεται εισάγοντας το όνομα του χρήστη που αναζητά και πατώντας στο σύμβολο του μεγεθυντικού φακού ή πατώντας το πλήκτρο enter. Ο ελάχιστος αριθμός χαρακτήρων για αναζήτηση είναι πέντε και είναι μια δικλείδα που περιορίζει τα αποτελέσματα για βελτίωση των αποτελεσμάτων. Τα αποτελέσματα παρουσιάζονται κάτω από το φίλτρο της αναζήτησης. Από τα αποτελέσματα αφαιρούνται οι χρήστες που είναι ήδη φίλοι με τον χρήστη μας. Δίπλα στα αποτελέσματα της αναζήτησης εμφανίζεται το σύμβολο + το οποίο είναι ένας σύνδεσμος για αποστολή αιτήματος φιλίας σε αυτόν τον χρήστη. Ο χρήστης αυτός θα λάβει ένα αίτημα φιλίας το οποίο αν γίνει αποδεκτό θα εντάξει τον χρήστη στην ομάδα των φίλων του χρήστη μας.



Εικόνα 6-14. Φόρμα φίλων χρήστη

Πιο κάτω στην ίδια φόρμα υπάρχει η ενότητα των εκκρεμωσών αιτημάτων φιλίας προς τον χρήστη μας. Οι χρήστες που εμφανίζονται εδώ έχουν αποστείλει ένα αίτημα φιλίας στον χρήστη μας. Δίπλα στο όνομα των χρηστών αυτών εμφανίζονται τα εικονίδια απάντησης

αίτηματος φιλίας. Αν ο χρήστης απαντήσει θετικά στο αίτημα φιλίας τότε θα δημιουργηθεί μία σύνδεση φιλίας ανάμεσα στους δύο χρήστες. Ο χρήστης αυτός θα μεταφερθεί στο κομμάτι των φίλων του χρήστη μας. Αν ο χρήστης μας αρνηθεί το αίτημα φιλίας τότε ο χρήστης θα αφαιρεθεί από την λίστα των αιτημάτων.

Στην επόμενη ενότητα της φόρμας αυτής παρουσιάζονται οι φίλοι του χρήστη μας. Επιλέγοντας κάποιον από τους φίλους του ο χρήστης μας μεταφέρεται στην σελίδα του προφίλ τους χρήστη αυτού. Την σελίδα αυτή θα περιγράψουμε στην επόμενη παράγραφο.

6.1.6. UserProfile. Σελίδα πληροφοριών φίλων χρήστη

Η σελίδα UserProfile είναι η σελίδα πληροφοριών των φίλων του χρήστη μας Εικόνα 6-15. Επιλέγοντας ο χρήστης μας κάποιον φίλο του μεταφέρεται στην σελίδα αυτή στην οποία μπορεί να δει αρκετές πληροφορίες για τον φίλο του και τις προτιμήσεις του. Στα αριστερά της σελίδας αυτής παρουσιάζονται ορισμένα προσωπικά στοιχεία του χρήστη αυτού. Βλέπουμε την εικόνα του χρήστη αυτού καθώς και πληροφορίες όπως το όνομα, το επώνυμο, το nickname και την διεύθυνση ηλεκτρονικού ταχυδρομείου του χρήστη. Η σελίδα αυτή εμφανίζεται μόνο σε φίλους του χρήστη μας δηλαδή χρήστες που έχουν αποδεχθεί την φιλία με τον χρήστη μας και άρα μπορούμε να δείξουμε αυτές τις προσωπικές πληροφορίες.

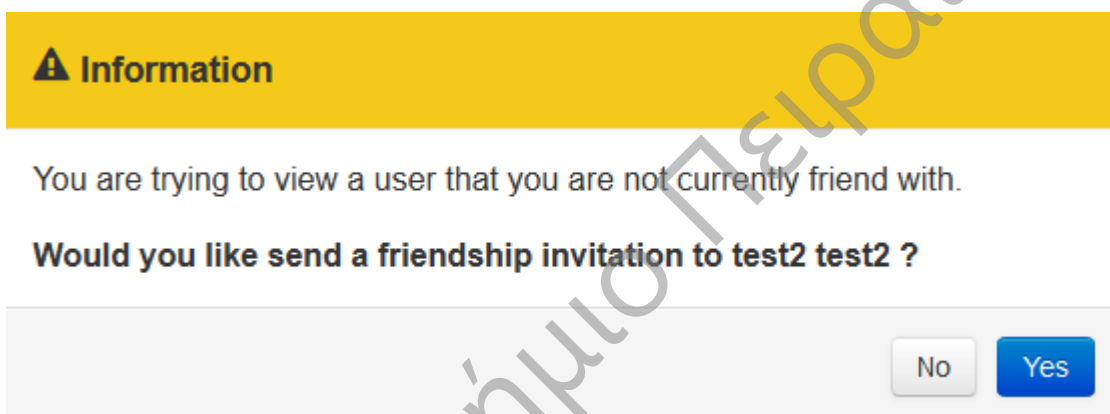


Εικόνα 6-15. Προφίλ φίλου χρήστη

Στο κεντρικό μέρος της σελίδας παρουσιάζεται η λίστα με τα πέντε πιο αγαπημένα καταστήματα στον φίλο του χρήστη μας αλλά και τα πέντε καταστήματα που έχουν λάβει βαθμολογία κάτω του μετρίου. Τα καταστήματα αυτά παρουσιάζονται με σειρά βαθμολογίας από το υψηλότερα βαθμολογημένο στο χαμηλότερα βαθμολογημένο. Αυτά τα καταστήματα μπορεί να αρέσουν στον χρήστη μας αφού σαν φίλοι μπορεί να μοιράζονται τις ίδιες συνήθειες και τις ίδιες προτιμήσεις.

Στα δεξιά της σελίδας αυτής παρουσιάζεται μία φόρμα με τους φίλους του χρήστη που έχουμε επιλέξει. Αν ένας φίλος του χρήστη είναι και φίλος του δικού μας χρήστη τότε δίπλα στο όνομα του εμφανίζεται η ένδειξη (mutual). Αν ένας από τους φίλους του χρήστη μας έχει αποστείλει ένα αίτημα φιλίας τότε εμφανίζεται η ένδειξη tick(check) η οποία ενημερώνει τον χρήστη μας για αυτό. Στην περίπτωση που ένας φίλος του χρήστη που έχουμε επιλέξει δεν ανήκει στις παραπάνω περιπτώσεις τότε δίπλα στο όνομα του εμφανίζεται η ένδειξη + με την οποία μπορούμε να στείλουμε ένα αίτημα φιλίας σε αυτόν τον χρήστη.

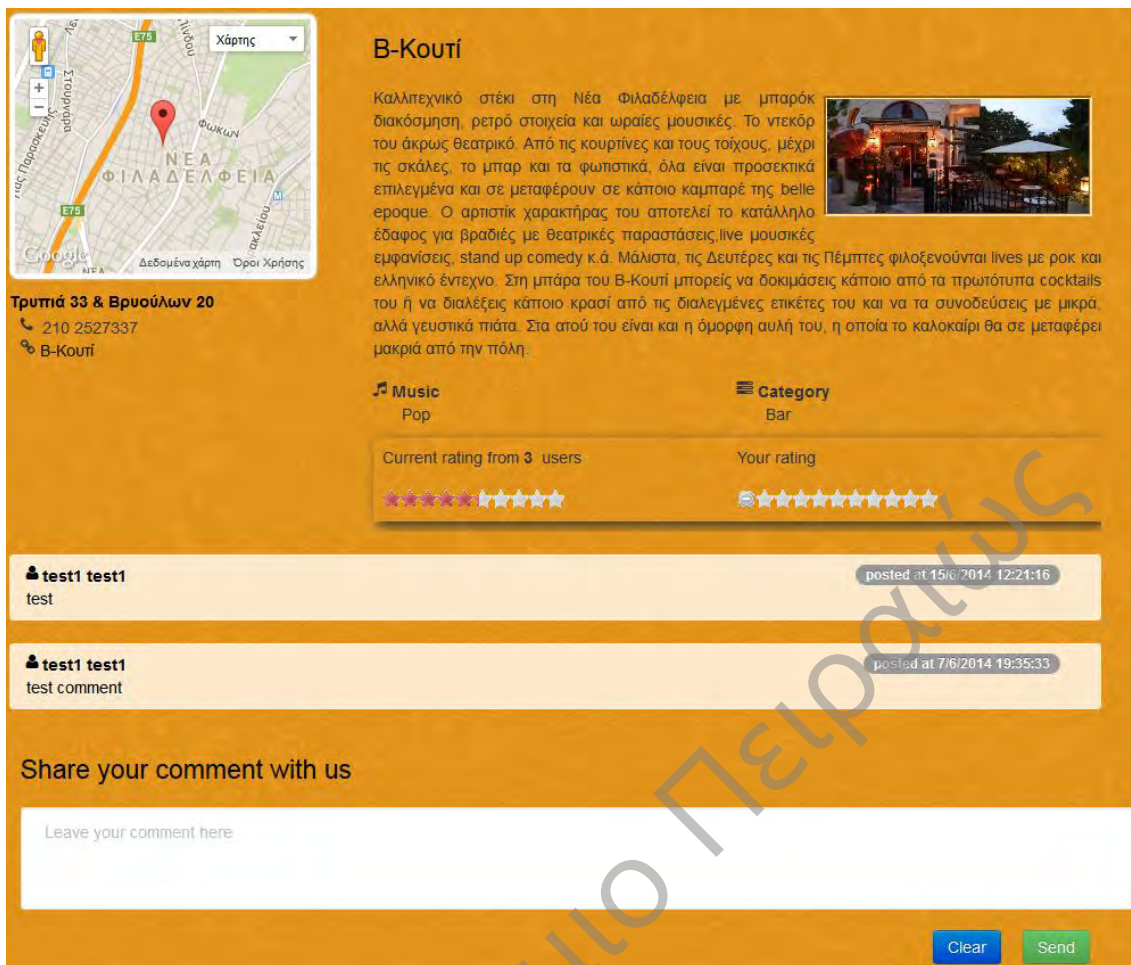
Αν ο χρήστης μας επιλέξει να δει το προφίλ ενός “ξένου” χρήστη τότε εμφανίζεται σε αυτόν ένα παράθυρο διαλόγου που φαίνεται στην παρακάτω εικόνα 6-16. Το παράθυρο αυτό ενημερώνει τον χρήστη μας ότι έχει επιχειρήσει να δει το προφίλ ενός χρήστη που δεν έχουμε σχέση φιλίας. Έτσι δίνεται η επιλογή στον χρήστη μας να αποστείλει ένα αίτημα φιλίας σε αυτόν τον χρήστη.



Εικόνα 6-16. Παράθυρο διαλόγου αποστολή αιτήματος φιλίας

6.1.7. StoreDetails. Σελίδα αναλυτικών πληροφοριών καταστήματος

Στην σελίδα StoreDetails ο χρήστης μας μπορεί να δει αναλυτικές πληροφορίες για το κατάστημα που έχει επιλέξει. Η σελίδα αυτή φαίνεται στην παρακάτω εικόνα 6-17.



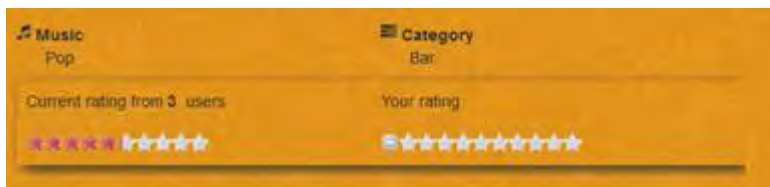
Εικόνα 6-17. Σελίδα εμφάνισης πληροφοριών καταστήματος

Σε αυτήν την σελίδα ο χρήστης έχει την δυνατότητα να δει αναλυτικές πληροφορίες σχετικά με το κατάστημα που έχει επιλέξει. Στο αριστερά πάνω μέρος της σελίδας ο χρήστης μπορεί να δει την τοποθεσία του καταστήματος με την βοήθεια ενός χάρτη(google maps). Στον χάρτη αυτό παρουσιάζεται το ακριβές σημείο του καταστήματος με την βοήθεια μίας πινέζας η οποία καταδεικνύει την ακριβή τοποθεσία του. Όταν ο χρήστης επιλέξει την πινέζα αυτή εμφανίζεται ένα ενημερωτικό παράθυρο στο οποίο αναγράφεται η διεύθυνση το καταστήματος. Ο χρήστης μπορεί να κάνει ζουμ στον χάρτη αυτόν για να κατατοπιστεί για την ακριβή θέση του καταστήματος στον χώρο.

Ακριβώς κάτω από τον χάρτη, εμφανίζονται στοιχεία όπως η διεύθυνση του καταστήματος καθώς και στοιχεία επικοινωνίας όπως το τηλέφωνο του καταστήματος. Επίσης εμφανίζεται και ο σύνδεσμος του καταστήματος ο οποίος μεταφέρει τον χρήστη μας στην σελίδα του καταστήματος ανοίγοντας ξεχωριστή καρτέλα στον φυλλομετρητή. Η σελίδα του καταστήματος μπορεί να είναι η επίσημη ιστοσελίδα του καταστήματος αλλά μπορεί να είναι και η σελίδα σε κάποιο μέσο κοινωνικής δικτύωσης όπως το facebook.

Στο δεξί μέρος της σελίδας StoreDetails ο χρήστης βλέπει το όνομα του καταστήματος που έχει επιλέξει καθώς και την περιγραφή του καταστήματος η οποία έχει καταχωρηθεί από το κατάστημα στην βάση δεδομένων μας. Το κείμενο της περιγραφής του καταστήματος συνοδεύεται από την εικόνα του καταστήματος έτσι ώστε ο χρήστης να έχει και μία “γεύση” του χώρου αυτού.

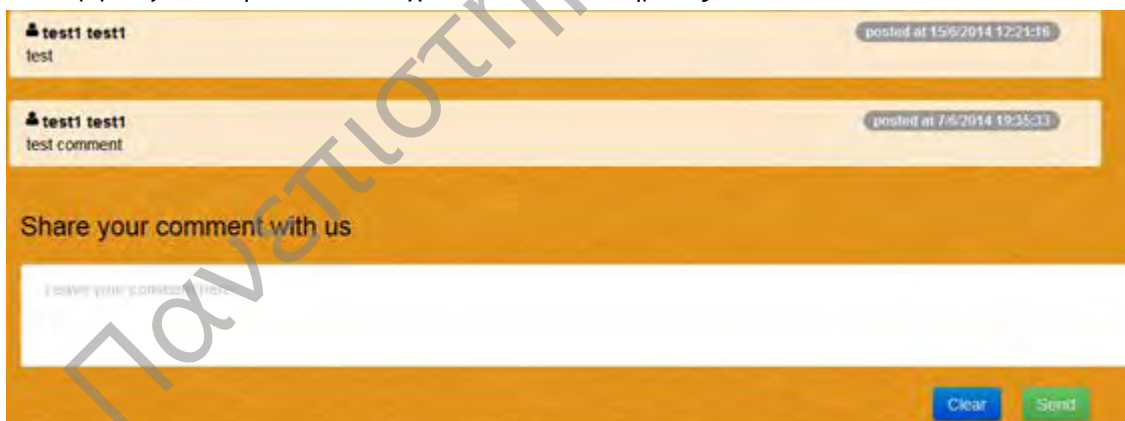
Λίγο πιο κάτω από την περιγραφή του καταστήματος εμφανίζεται η περιγραφή της κατηγορίας της μουσικής που ανήκει το κατάστημα καθώς και το είδος του καταστήματος. Κάτω από αυτές τις πληροφορίες εμφανίζεται η φόρμα της βαθμολόγησης του καταστήματος Εικόνα 6-18.



Εικόνα 6-18. Φόρμα βαθμολόγησης καταστήματος

Στο αριστερό μέρος αυτής της φόρμας ο χρήστης μπορεί να δει την βαθμολογία που έχει συγκεντρώσει το συγκεκριμένο κατάστημα καθώς και τον αριθμό των χρηστών που το έχουν αξιολογήσει. Η βαθμολογία του καταστήματος παρουσιάζεται σε κλίμακα 1-10 με την μορφή δέκα αστεριών. Η βαθμολογία που έχει συγκεντρώσει το κατάστημα φαίνεται με τα χρωματισμένα αστέρια στην κλίμακα αυτή. Στην συγκεκριμένη εικόνα το κατάστημα έχει λάβει πέντε αστέρια δηλαδή βαθμολογία 5/10 από σύνολο τριών χρηστών. Στο δεξί μέρος της φόρμας παρουσιάζεται η βαθμολόγηση του χρήστη μας για το συγκεκριμένο κατάστημα. Θυμίζουμε ότι ο χρήστης μπορεί να βαθμολογήσει ένα κατάστημα το πολύ δύο φορές και η τελική βαθμολογία είναι ο μέσος όρος των βαθμολογιών που έχει θέσει. Ο χρήστης μπορεί να βαθμολογήσει το κατάστημα απλά επιλέγοντας το αστέρι που επιθυμεί στην κλίμακα 1-10. Αφού επιλέξει την βαθμολογία του καταχωρείται στο σύστημα και ένα παράθυρο επιτυχούς βαθμολόγησης τον ενημερώνει. Σε περίπτωση που ο χρήστης έχει ήδη βαθμολογήσει το κατάστημα δύο φορές τότε τα αστέρια αυτά είναι ανενεργά μην επιτρέποντας επιπλέον βαθμολόγηση ενώ παρουσιάζεται στον χρήστη μήνυμα που τον ενημερώνει ότι έχει φτάσει το μέγιστο αριθμό αξιολογήσεων.

Στο τέλος της συγκεκριμένης σελίδας φαίνονται τα σχόλια των χρηστών για το συγκεκριμένο κατάστημα Εικόνα 6-19. Το κάθε μήνυμα περιέχεται σε ξεχωριστό τομέα στο οποίο αναφέρονται κάποιες πληροφορίες για τον συντάκτη καθώς και την ημερομηνία και ώρα της δημοσίευσης. Τα μηνύματα για τα καταστήματα βοηθούν τους χρήστες ανταλλάξουν απόψεις για τα καταστήματα και να συζητήσουν για τα αρνητικά και τα θετικά σχετικά με την ατμόσφαιρα το σέρβις και ότι άλλο θέλουν να αναφέρουν. Η εισαγωγή ενός νέου σχολίου από τον χρήστη γίνεται με την ειδική φόρμα στο τέλος της σελίδας. Ο χρήστης μπορεί να αφήσει το μήνυμά του το οποίο μπορεί να έχει μέγεθος το πολύ 200 χαρακτήρες. Για να καταχωρήσει το σχόλιο του πρέπει να πατήσει το κουμπί send. Μετά την επιτυχή καταχώρηση του σχολίου, αυτό εμφανίζεται στην λίστα των σχολίων του καταστήματος.



Εικόνα 6-19. Σχόλια χρήστη για το κατάστημα

6.1.8. ManageAccount. Σελίδα αλλαγή προφίλ χρήστη

Σε αυτή τη σελίδα ο χρήστης μπορεί να αλλάξει κάποια στοιχεία που αφορούν τον λογαριασμό του. Τα στοιχεία αυτά είναι το όνομα, το επώνυμο το nickname, το email, το τηλέφωνο και τον αριθμό του κινητού τηλεφώνου. Η αλλαγή των στοιχείων γίνεται με πεδία παρόμοια με τα πεδία που είδαμε στην σελίδα εγγραφής του χρήστη. Σε αυτά τα πεδία ισχύουν οι περιορισμοί που είδαμε στην σελίδα αυτή.



Εικόνα 6-20. Αλλαγή προφίλ χρήστη

7. Συμπεράσματα

7.1. Εισαγωγή

Σε αυτήν την εργασία αναλύσαμε το θεωρητικό υπόβαθρο που διέπει τους αλγορίθμους συνεργατικού φιλτραρίσματος. Μελετήσαμε την προέκτασή τους στα κοινωνικά δίκτυα και δημιουργήσαμε το δικό μας περιορισμένο κοινωνικό δίκτυο. Αναλύσαμε την εφαρμογή του αλγορίθμου προτάσεων και την ενσωμάτωσή του στο κοινωνικό δίκτυο. Τέλος αναλύσαμε την αλληλεπίδραση του χρήστη με τον αλγόριθμο αλλά και με τους άλλους χρήστες του κοινωνικού δικτύου.

7.2. Συμπεράσματα

Οι εξέλιξη των συνεργατικών αλγορίθμων συστάσεων έχουν δημιουργήσει μία τάση για δημιουργία ολοένα και περισσότερων ιστοσελίδων που αποζητούν την αλληλεπίδραση του χρήστη έτσι ώστε να τον τροφοδοτήσουν με στοχευμένα δεδομένα. Παράλληλα με την ανάπτυξη αυτών των αλγορίθμων έχει ανθίσει τα κοινωνικά δίκτυα στα οποία ο χρήστης εκφράζει πιο ελεύθερα και απρόσκοπτα τις προτιμήσεις του.

Από τα παραπάνω καταλήγουμε στο συμπέρασμα ότι η ενσωμάτωση των κοινωνικών δικτύων στην διαδικασία προτάσεων είναι αναγκαία. Αυτό προκύπτει από τα πλεονεκτήματα που μας δίνουν τα κοινωνικά δίκτυα στο “κυνήγι της πληροφορίας”. Μερικά από τα πλεονεκτήματα είναι τα εξής:

- Αμεσότητα με τον χρήστη. Τα κοινωνικά δίκτυα παρέχουν στον χρήστη μία αμεσότητα στην αλληλεπίδραση του με άλλους χρήστες και την συμπεριφορά τους.

- Οικειότητα του χρήστη. Ο χρήστης τείνει να καταστήσει την χρησιμοποίηση των κοινωνικών δικτύων ρουτίνα της ημέρας του. Αυτό επιτρέπει στα μέσα κοινωνικής δικτύωσης να κερδίσουν την εμπιστοσύνη του, κάτι που σημαίνει ότι ο χρήστης αισθάνεται πιο άνετα σε ένα περιβάλλον που είναι συσχετισμένο με ένα τέτοιο μέσο. Σε αυτό το περιβάλλον ο χρήστης προβάλλει από μόνος του στοιχεία της καταναλωτικής του συμπεριφοράς και κατά συνέπεια βοηθά τους αλγορίθμους προτάσεων στην να εξάγουν πιο εύκολα συμπεράσματα για τις προτιμήσεις του.
- Εύκολη πρόσβαση. Ο χρήστης μπορεί να χρησιμοποιήσει τα κοινωνικά δίκτυα από παντού έχοντας μαζί του ένα φορητό υπολογιστή, μία ταμπλέτα ή ακόμα και από το κινητό του τηλέφωνο καθώς και μία σύνδεση στο διαδίκτυο.

Τα παραπάνω είναι μόνο μερικά από τα πλεονεκτήματα που τα κοινωνικά δίκτυα παρέχουν στο στις εταιρίες που στηρίζονται στους αλγορίθμους προτάσεων. Με την ενσωμάτωση των κοινωνικών δικτύων σε ένα τέτοιο σύστημα προτάσεων τα εξαχθέντα αποτελέσματα είναι κατά πολύ ποιοτικότερα σε σχέση με την μεμονωμένη λειτουργία των αλγορίθμων.

7.3. Εξέλιξη στο μέλλον

Όπως έχουμε διαπιστώσει από την εργασία αυτή η συνεργασία των κοινωνικών δικτύων με τους αλγόριθμους προτάσεων στον χρήστη καθιστού επιτακτική την ανάγκη στις εταιρίες να επενδύσουν όχι μόνο στην εξέλιξη των αλγορίθμων αλλά και στα ίδια τα κοινωνικά δίκτυα.

Η χρησιμοποίηση των μέσων αυτών επιτρέπει στις εταιρίες να προσεγγίσουν ακόμα περισσότερο του υποψήφιους πελάτες τους αφού με την δημιουργία πιο επιτυχημένων προφίλ μπορούν να προσεγγίσουν σε μεγάλο βαθμό την συμπεριφορά τους. Η συμπεριφορά των χρηστών μπορεί να εξαχθεί με επιτυχημένες τεχνικές data mining για την εξαγωγή αυτών των δεδομένων. Τα δεδομένα που προκύπτουν βρίσκονται σε μία πολύ απλή και αδόμητη μορφή. Σε αυτό το σημείο η κατάλληλη επεξεργασία τους με την βοήθεια του collaborating filtering μπορεί να δώσει δεδομένα τα οποία είναι κατά ένα πολύ μεγάλο ποσοστό έτοιμα για πρόταση στον χρήστη.

Ολοκληρώνοντας θεωρούμε ότι οι εταιρίες που παρέχουν τις προτάσεις στον χρήστη σαν μέσο για μία επιτυχημένη συμπεριφορά πρέπει να ακολουθήσουν τα βήματα των μέσων κοινωνικής δικτύωσης όχι μόνο για να επιτύχουν μία καλή στρατηγική marketing αλλά για να μετατρέψουν τους χρήστες σε υποψήφιους πελάτες. Άρα αυτές οι εταιρίες πρέπει να συνεργαστούν και να εναρμονιστούν με την ιδέα και την δομή των κοινωνικών δικτύων για να εξασφαλίσουν μεγαλύτερες πωλήσεις άρα και μεγαλύτερα κέρδη.

Αυτό όμως πολλές φορές δεν είναι αρκετό. Θα πρέπει οι ίδιες εταιρίες να δημιουργήσουν ένα περιορισμένο κοινωνικό δίκτυο που θα επιτρέπει στους χρήστες να ενταχθούν σε αυτό και να δημιουργήσουν το δικό τους προφίλ. Το προφίλ αυτό θα αποτελέσει ένα μεγάλο δείγμα της συμπεριφοράς τους παρέχοντας στην εταιρεία πληροφορίες που μπορεί να δείξουν τον δρόμο στο πρόβλημα των προτάσεων προϊόντων και υπηρεσιών στον χρήστη. Το προφίλ του χρήστη σε αυτήν την περίπτωση θα πρέπει να είναι η κύρια βάση στην οποία ένα επιτυχημένο σύστημα προτάσεων πρέπει να στηριχθεί.

8 Βιβλιογραφία

- [1] Μ. Κ. Βίρβου. Αντικειμενοστραφής τεχνολογία λογισμικού. Πανεπιστήμιο Πειραιώς, 1993.
- [2] Γ. Βασιλακόπουλος, Β. Χρυσικόπουλος. Πληροφοριακά Συστήματα Διοίκησης. Εκδόσεις Σταμούλη, 1990.
- [3] Silberschatz, Korth, Sudarshan Συστήματα Βάσεων Δεδομένων 4^η έκδοση 2007
- [4] Hart, C.A., Doherty, N. and Ellis-Chadwick, F. (2000). Retailer adoption of the Internet - implications for retail marketing. *European Journal of Marketing* 24, pp.954–974.
- [5] Reynolds, J. (2000). E-Commerce: a critical review. *International Journal of Retail and Distribution Management* 28, pp. 414–444.
- [6] Bakos, Y. (2001). The emerging landscape for retail e-commerce. *Journal of Economic Perspectives* 15(1), pp. 69–80.
- [7] Fengkun Liu, Hong Joo Lee (2009). Use of social network information to enhance collaborative filtering performance.
- [8] Robert M. Bell and Yehuda Koren AT&T Labs – Research 180 Park Ave, Florham Park, NJ 07932 Improved Neighborhood-based Collaborative Filtering.
- [9] Rowley, J.A. (2000). Product searching with shopping bots. *Internet Research* 10 (3), pp. 203–214.
- [10] Schwab, I., Kobsa, A. and Koychev, I. (2001). Learning User Interests through Positive Examples Using Content Analysis and Collaborative Filtering. Internal Memo, GMD, St. Augustin, Germany.
- [11] Zukerman, I. and Albrecht, D. (2001). Predictive Statistical Models for Modelling. *User Modelling and User-Adapted Interaction* 11(1-2), pp. 5-18.
- [12] Alderferer, M.S. and Blashfield, R.K. (2004). *Cluster Analysis*. Beverly Hills, CA., Sage Publications.
- [13] Bo Yang Yu Lei Dayou Liu Jiming Liu. Social Collaborative Filtering by Trust.
- [14] Joong Hee Jeong and Jong Woo Kim (2012). Personalized Recommendation Based on Collaborative Filtering with Social Network Analysis.
- [15] Xiwang Yang, Yang Guo, Yong Liu, Harald Steck. A Survey of Collaborative Filtering Based Social Recommender Systems.
- [17] Han, J. and Kamber, M. (2000). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- [18] Liao, S. (2003). Knowledge Management Technologies and Applications – Literature Review from 1995 to 2002. *Expert Systems with Applications* 25, pp. 155-164.

- [19] Terveen, L. and Hill, W. (2001). Human-Computer Collaboration in Recommender Systems. Human Computer Interaction in the New Millenium. New York: Addison-Wesley, pp. 487-509.
- [20] Sarwar B.M., Karypis G., Konstan J.A. and Riedl J. (2001). Item-Based Collaborative Filtering Recommendation Algorithms. Proceedings of the 10th International World Wide Web Conference, Hong Kong, ACM Press.
- [21] Rosenstein, M. and Lochbaum, C. (2000). Recommending from Content: Preliminary Results from an E-Commerce Experiment. Proceedings of the CHI'00: Conference on Human Factors in Computing, The Hague, Netherlands.
- [22] Burke, R. (2000). Knowledge-based Recommender Systems. A. Kent (ed.): Encyclopedia of Library and Information Systems 69, Supplement 32.
- [23] Kautz, Selman, Shah (1997). Combining Social Networks & Collaborating Filtering.
- [24] www.wikipedia.org.

9. Πίνακας εικόνων

Εικόνα 2-1. Αναπαράσταση συστήματος Collaborative Filtering

Εικόνα 2-2. Λειτουργία συστήματος γνώσης

Εικόνα 2-3. Λειτουργία συστήματος τύπου «Καταρράκτη»

Εικόνα 4-1. Visual studio 2012 IDE

Εικόνα 4-2. MS Sql Management studio

Εικόνα 4-3. Αρχιτεκτονική server

Εικόνα 4-4. Διάγραμμα σχέσεων Exodus_db

Εικόνα 4-5. Σχεσιακό διάγραμμα Exodus_db, 1

Εικόνα 4-6. Σχεσιακό διάγραμμα Exodus_db, 2

Εικόνα 5-1. Σχηματική περιγραφή αλγορίθμου συστάσεων.

Εικόνα 5-2. Μαθηματικός τύπος ομοιότητας συνημίτονου.

Εικόνα 5-3. Προβολή διανυσμάτων χρήστη A στο διάνυσμα του χρήστη B.

Εικόνα 5-4. Περιπτώσεις διανυσμάτων στο χώρο και βαθμός ομοιότητας.

Εικόνα 5-5. Ανατροφοδότηση αλγορίθμου

Εικόνα 5-6. Μηχανισμός αναδιαμόρφωσης προφίλ

Εικόνα 5-7. Διαδικασία αναδιαμόρφωσης προφίλ χρήστη

Εικόνα 6-1. Εγγραφή χρήστη και validation

Εικόνα 6-2. Σελίδα σύνδεσης χρήστη

Εικόνα 6-3. Αρχική σελίδα Exodus

Εικόνα 6-4. Κεφαλίδα Exodus

Εικόνα 6-5. Παρουσίαση κατηγοριών καταστημάτων

Εικόνα 6-6. Παρουσίαση υψηλότερα βαθμολογημένων καταστημάτων

Εικόνα 6-7. Παρουσίαση ειδών μουσική και ειδών καταστημάτων

Εικόνα 6-8. Υποσέλιδο Exodus

Εικόνα 6-9. Ειδοποιήσεις ενημέρωσης χρήστη

Εικόνα 6-10. Δημιουργία προφίλ χρήστη

Εικόνα 6-11. Σελίδα αναζήτησης καταστημάτων

Εικόνα 6-12. Σελίδα προτάσεων στον χρήστη

Εικόνα 6-13. Προτεινόμενο κατάστημα το οποίο αρέσει σε φίλους

Εικόνα 6-14. Φόρμα φίλων χρήστη

Εικόνα 6-15. Προφίλ φίλου χρήστη

Εικόνα 6-16. Παράθυρο διαλόγου αποστολή αιτήματος φιλίας

Εικόνα 6-17. Σελίδα εμφάνισης πληροφοριών καταστήματος

Εικόνα 6-18. Φόρμα βαθμολόγησης καταστήματος

Εικόνα 6-19. Σχόλια χρήστη για το κατάστημα

Εικόνα 6-20. Αλλαγή προφίλ χρήστη

Πίνακας 4-1. Πίνακες Exodus_db

Πίνακας 5-1. Υποψήφιοι χρήστες για μέτρηση ομοιότητας.

Πίνακας 5-2. Κατάταξη βαθμολογίας χρηστών.

Πίνακας 5-3. Φίλτρα από τα είδη της μουσικής.

Πίνακας 5-4. Στιγμιότυπο εκτέλεσης 1

Πίνακας 5-5. Στιγμιότυπο εκτέλεσης 2

Πανεπιστήμιο Πειραιώς